European Organisation
for Astronomical
Research in the
Southern Hemisphere

Organisation Européenne
pour des Recherches
Astronomiques
dans l'Hémisphère Austral

Europäische Organisation
für astronomische
Forschung in der
südlichen Hemisphäre

E-ELT PROGRAMME

THE E-ELT DESIGN REFERENCE MISSION

E-TRE-ESO-080-0717 Issue 2

18 January 2011

Owner	J. Liske		
Project scientist	M. Kissler-Patig		
Head of Programme Office	R. Gilmozzi		
	Name	Date	Signature

Authors

Name	Affiliation
J. Liske	ESO (EScO)
A. Ardeberg	Lund Observatory (SWG)
G. Battaglia	ESO (EScO)
J. Bergeron	Paris Institute of Astrophysics (SWG)
A. Calamida	ESO (EScO)
A. Cimatti	Bologna University (SWG)
M. Cappellari	University of Oxford
F. Comerón	ESO (SWG)
X. Dumusque	Geneva Observatory and CAUP, Porto
S. Feltzing	Lund Observatory (SWG)
M. Franx	Leiden Observatory (SWG)
P. François	GEPI, Paris-Meudon Observatory
W. Freudling	ESO (SWG)
S. Gładysz	ESO (EScO)
R. Gratton	INAF, Astronomical Observatory of Padova (SWG)
M. Haehnelt	IoA, Cambridge (SWG)
V. Hill	Côte d'Azur Observatory (SWG)
I. Hook	University of Oxford & University of Rome (SWG)
G. James	ESO (EScO)
H.-U. Kaeufl	ESO (SWG)
M. Kissler-Patig	ESO
A. Küpcü Yoldaş	ESO (EScO)
M. Lehnert	GEPI, Paris-Meudon Observatory (SWG)
B. Leibundgut	ESO (SWG)
C. Lovis	Geneva Observatory (SWG)
M. McCaughrean	ESA (SWG)
D. Naef	Geneva Observatory (EScO)
M. Nonino	INAF, Astronomical Observatory of Trieste
L. Pasquini	ESO
M. Puech	GEPI, Paris-Meudon Observatory (EScO)
D. Queloz	Geneva Observatory (SWG)
R. Rebolo	IAC (SWG)
J.M. Rodriguez Espinosa	IAC (SWG)
P. Rosati	ESO (SWG)
S. Toft	Dark Cosmology Centre, Copenhagen (EScO)
E. Tolstoy	Kapteyn Astronomical Institute, Groningen (SWG)
S. Udry	Geneva Observatory
B. Venemans	ESO (EScO)
D. Villegas	ESO (EScO)
H. Zinnecker	Astrophysical Institute Potsdam (SWG)

Change record

Issue	Date	Section / Paragraph affected	Reason / Initiation Documents / Remarks
1	01 Oct 2010	All	Initial version
2	18 Jan 2011	Section 8	Some revision

Contents

1 Scope

This document describes the E-ELT Design Reference Mission. It includes detailed descriptions of the work carried out for each of the DRM cases. The original DRM proposals are included in the Appendix.

2 Applicable documents

2.1 Applicable documents

AD1 E-ESO-SPE-313-0066 Issue 4, Common definitions and acronyms

2.2 Reference documents

RD1 E-TRE-ESO-080-0718 Issue 1, The E-ELT Design Reference Mission: technical data used for simulations
RD2 Report of the ELT Science Working Group: Science Cases and Requirements for the ESO ELT, April 2006
RD3 E-TRE-ESO-080-0806 Issue 1, E-ELT Science Case
RD4 EFP7-TRE-ESO-04000-0002 Issue 1, Report on impact of instrumentation on the DRM

3 Executive summary

The E-ELT Design Reference Mission (DRM) encompasses a detailed, hands-on exploration of a selected sample of science cases through the analysis of simulated E-ELT data. The purpose of this exercise was (i) to provide a quantitative assessment of the extent to which the E-ELT will be capable of addressing key scientific questions, (ii) to assist the project in making critical trade-off decisions by quantifying their consequences in terms of scientific gains and losses, and (iii) to support the development of the E-ELT Science Case. The overarching aim of the DRM was to help ensure that the E-ELT will meet the scientific aspirations of its community. The science cases studied by the DRM were selected by the E-ELT Science Working Group for being prominent examples highlighting the key capabilities of the telescope, and they cover a wide range of science topics (see Table 4.1). We now briefly describe each of these in turn.

The detection, characterisation and evolution of giant and terrestrial exoplanets is a key science goal of the E-ELT and this topic was also a focus of the DRM. The DRM case **"Direct imaging of terrestrial and giant exoplanets"** investigated which types of planets the E-ELT will be able to detect in reflected light using high-contrast imaging techniques. The simulations have confirmed that the E-ELT will open an entirely new discovery space in the field of exoplanet imaging. Not only will it be possible to detect old Jupiter-like and smaller giant gaseous planets out to distances of 30 pc from the Sun, but it may even be feasible to detect a terrestrial planet, or at least a super-Earth, in the habitable zone of its parent star within a distance of 10 pc if certain stringent instrument requirements can be met.

Similarly, the DRM case **"Earth twins in the habitable zone of solar-type stars"** pursued the question of whether the E-ELT will be able to uncover Earth-like planets by detecting and measuring the reflex motion of the host stars using the radial velocity method. The issue here is whether it is possible to detect the small radial velocity signal due to an Earth-like planet in the 'noise' generated by various stellar phenomena. However, the DRM simulations have shown that the large collecting area of the E-ELT enables the design of observing strategies that will allow us to average out the stellar noise, such that the detection of planets with masses as small as 1–2 Earth masses in the habitable zone of their parent star appears feasible.

The DRM case **"Imaging the planet-forming regions of circumstellar disks"** examined the E-ELT's capability of studying the process of planet formation by investigating the spatial structure of young, low-

mass circumstellar disks via direct near and mid-infrared imaging. In particular, the goal was to test whether the E-ELT will be able to detect gaps in such disks resulting from a forming planet clearing its orbit. The simulations have shown that a young, low-mass disk surrounding a T Tauri star can easily be detected out to distances of at least 1 kpc. Detecting planet formation features in such a disk will be much more challenging and will require excellent knowledge of the imaging system's point spread function. Nevertheless, the simulation results indicate that a 2 AU-wide gap at a distance of 10 AU from the central star will be detectable at the distances of the nearest star forming regions.

The aim of the DRM case **"Characterising the lowest mass freely floating objects in star forming regions"** was to understand how well the E-ELT will be able to spectroscopically characterise the low-mass end of the sub-stellar initial mass function in the nearest star forming regions. The simulations for this case have shown that it will be possible to extract the diagnostic atomic and molecular features relevant for comparisons with models of ultracool atmospheres from intermediate resolution spectra with sufficient accuracy for objects with masses as small as 3 M_{Jup}.

Showing how the E-ELT will help to understand the formation and early dynamical evolution of massive stars embedded in dense, highly obscured Galactic proto-cluster clouds was the topic of the DRM case **"The centres of massive dense young clusters: deep E-ELT infrared imaging and 3D spectroscopy"**. The issue at hand is whether the E-ELT will be capable of penetrating the extreme dust extinction that is encountered in the dense molecular clumps where massive stars are born. The DRM simulations have shown that by observing in the mid-infrared the E-ELT will indeed be able to obtain a complete census of the high-mass end of the initial mass function in young massive clusters as far away as the Galactic centre, even when the stars are hidden behind up to 200 mag of visual extinction.

The DRM case **"Giant planet-mass objects in the Large Magellanic Cloud"** (LMC) appraised the E-ELT's ability of constraining the low-mass end of the sub-stellar initial mass function, and possibly determining its lower mass limit, in the low-metallicity environment of the LMC through near-infrared imaging. Will the E-ELT have enough sensitivity and resolution to be able to detect the faintest objects in a low-mass star forming region, given the large dynamic range of the stars in such regions and their high density when observed at the distance of the LMC? The answer provided by the simulations is that the E-ELT will indeed be able to probe the initial mass function in the LMC down to masses as small as 11 M_{Jup}.

Studying the stellar populations of galaxies by resolving them into individual stars is another key science driver of the E-ELT. The first DRM case to address this topic, **"The resolved stellar populations of elliptical galaxies"**, was concerned with the E-ELT's ability to perform photometry in extremely crowded stellar fields and to construct accurate colour-magnitude diagrams for elliptical galaxies, the nearest true examples of which are only found in the relatively distant Virgo cluster. For the central Virgo elliptical the simulation results indicate that it will be possible to probe at least the tip of its red giant branch with high-quality photometry all the way into the galaxy's core, down to just 0.5 effective radii.

Complementary to the above, the DRM case **"The chemo-dynamical structure of galaxies"** considered the prospect of obtaining accurate metallicity distributions and kinematics from intermediate resolution spectroscopy of individual red giant branch stars in galaxies as distant as the Virgo cluster. As shown by the simulations, this will definitely be possible for Centaurus A but it will be extremely challenging at the distance of Virgo, even when targeting only the outer, less crowded regions of a galaxy.

Finally on this topic, the DRM case **"First stars relics in the Milky Way and satellites"** addressed the question to what extent the E-ELT will be able to help us understand the nature and nucleosynthesis of the earliest chemical enrichment processes by obtaining detailed abundance patterns of extremely metal-poor stars in the Milky Way and other Local Group galaxies by way of high-resolution spectroscopy. The conclusion from the simulations is that studies encompassing $\sim$100 Milky Way stars and dozens of stars in each of four satellite galaxies are entirely feasible, albeit requiring hundreds of hours of observing time.

The relation between the mass of the supermassive black holes found at the centres of early type galaxies and the velocity dispersion of their surrounding stellar spheroid is widely acknowledged to be of central importance to our understanding of galaxy evolution. The goal of the DRM case **"A survey of Black Holes in different environments"** was to quantify to what extent the E-ELT will be able to shed light on both the low and high-mass ends of this relation through spatially resolved spectroscopy of the central

regions of galaxies. The simulations have shown that the smallest supermassive black holes (with masses of order $10^6\ M_\odot$) will be detectable out to the distance of the Virgo cluster, while the largest ones (with masses of order $10^{10}\ M_\odot$) will be observable out to a redshift of $\sim$0.3.

Unravelling the complexities of galaxy evolution by surveying large numbers of high-redshift galaxies with spatially resolved spectroscopy, delivering dynamical masses, ages, metallicities, star formation rates, kinematics, etc., will undoubtedly be a key task for the E-ELT. The DRM case **"The physics and mass assembly of galaxies out to z$\sim$6"** explored the huge parameter space of such a programme. One of the key results of the simulations was that even at a redshift of 4 it will still be possible to recover the large-scale, internal motions of galaxies, enabling the distinction between rotating disks and major mergers, down to stellar masses of $\sim 10^{10}\ M_\odot$.

The DRM case **"High-resolution imaging of high-redshift galaxies"** examined the role of the E-ELT in studying the morphologies of high-redshift galaxies. According to the simulations, diffraction-limited images of high enough quality to allow the extraction of morphological parameters will be attainable for galaxies of characteristic stellar mass up to redshifts of $\sim$2.5. At somewhat lower spatial resolutions the E-ELT will be able to probe the stellar mass function up to a redshift of $\sim$7.5, although with its superior sensitivity the JWST will be an important competitor above a redshift of $\sim$4.

Moving to even higher redshifts, the DRM case **"First light – the highest redshift galaxies"** tested the E-ELT's ability to push spectroscopic studies of galaxies beyond a redshift of 7. The simulations found that it will not only be feasible to measure accurate redshifts for these galaxies, enabling clustering and dynamical studies, but even to determine physical parameters from absorption lines, at least for the brightest part of the luminosity function.

Finally, the DRM case **"A dynamical measurement of the expansion history of the Universe"** explored the fascinating possibility that the E-ELT might be able to directly observe the evolving expansion rate of the Universe by detecting, for the first time, the so-called redshift drift. Indeed, the simulations have shown that a two decade-long spectroscopic monitoring campaign of high redshift QSO absorption lines with the E-ELT, requiring hundreds of nights of observing time in total, would provide a measurement of the redshift drift of sufficient accuracy to unequivocally prove the existence of a period of accelerated expansion in the Universe's history.

In summary, the DRM simulations have provided quantitative evidence of the transformational nature of the science programme envisioned for the E-ELT: the E-ELT will undoubtedly revolutionise several of the fields mentioned above. However, the simulations have also demonstrated that a fair fraction of this programme lies at the edge of feasibility, leading to the conclusion that much of the science encapsulated by the DRM will indeed require a 42 m telescope. In addition, the simulations have verified that the site chosen for the E-ELT conforms to the science plans, and that almost none of the capabilities required by these science cases are missing from current adaptive optics and instrumentation plans.

4 Introduction

At the end of 2005 ESO established the ELT Science Working Group (SWG) together with four other Working Groups consisting of experts from the community and ESO's staff to help ESO prepare for the establishment of the E-ELT project. These Working Groups were charged to synthesise and collate ELT capabilities in their specified topic areas, to propose a basis for prioritising these and to provide a list of key trade-offs and an initial prioritisation of ELT capabilities. The reports from these Working Groups were subsequently used by the E-ELT Telescope Project Office as a toolkit to distill the first version of a Baseline Reference Design of the E-ELT in the latter half of 2006.

Much of the work of the SWG in particular focused on re-aligning earlier work on the science case for a previously envisaged 50–100 m class telescope (e.g., Hook, 2005, "The Science Case for the European Extremely Large Telescope: The next step in mankind's quest for the Universe") to the range of 30–60 m considered for the E-ELT at the time. This work resulted in a report [RD2], published in April 2006, which has served as a valuable basis for all E-ELT science case development since, including [RD3].

Table 4.1: Science cases studied by the DRM. The letter/number combinations refer to the science case designations in [RD2].

Planets & Stars	S3: From giant to terrestrial exoplanets: detection, characterisation and evolution ■ Direct imaging of terrestrial and giant exoplanets ■ Earth twins in the habitable zone of solar-type stars S9: Circumstellar disks ■ Imaging the planet-forming regions of circumstellar disks S5: Young stellar clusters ■ Characterising the lowest mass freely floating objects in star forming regions ■ The centres of massive dense young clusters: deep E-ELT infrared imaging and 3D spectroscopy ■ Giant planet-mass objects in the Large Magellanic Cloud
Stars & Galaxies	G4: Imaging and spectroscopy of resolved stellar populations in galaxies ■ The resolved stellar populations of elliptical galaxies ■ The chemo-dynamical structure of galaxies ■ First stars relics in the Milky Way and satellites G9: Black holes / AGN ■ A survey of black holes in different environments
Galaxies & Cosmology	C10: The physics of high redshift galaxies ■ The physics and mass assembly of galaxies out to z$\sim$6 ■ High resolution imaging of high redshift galaxies C4: First light – the highest redshift galaxies C2: A dynamical measurement of the expansion history of the Universe

Unlike the other Working Groups, the SWG was not disbanded after the completion of their report. Instead, in May 2006 it was merged with the OPTICON ELT networking activity and charged with further developing the science case for the E-ELT and advising the project on all scientific aspects of its development. In particular, one of the SWG's main tasks was the establishment of the E-ELT Design Reference Mission (DRM).

The DRM encompasses a detailed, hands-on exploration of a selected sample of science programmes. The purpose of this exercise was (i) to provide a quantitative assessment of the extent to which the E-ELT will be capable of addressing key scientific questions, (ii) to assist the project in making critical trade-off decisions by quantifying their consequences in terms of scientific gains and losses, and (iii) to support the development of the E-ELT Science Case. This process aimed to ensure that the E-ELT will be aligned with the scientific aspirations of its community as much as possible.

In Table 4.1 we list the set of prominent science cases that the SWG selected from [RD2] to be studied by the DRM. Although these cases were considered by the SWG to be among the highlights of the E-ELT science case, they were not intended to be an exhaustive list of the science that the E-ELT will do. They were rather chosen to encompass a wide range of different science topics and goals, and to exemplify cases which exploit and highlight the key capabilities of the telescope.

The DRM process began in 2007 with members of the SWG (with help from the community) drafting a 'DRM proposal' for each of the DRM science cases. These proposals were written somewhat in the manner of a regular ESO observing proposal (see Section 20): they briefly summarised the science

case and then described a more or less well-defined set of E-ELT observations designed to address the scientific question at hand.

The key task of the DRM was to answer, for each science case, the following two questions:

1. What precisely can be achieved with the observations described in the proposal in a given amount of observing time, or, vice versa, how much observing time is needed to achieve a given set of science goals?

2. How do these results depend on the properties of the telescope, the instrument, the AO performance and/or the site? Which features of the E-ELT system are critical to the success of the proposal? What is the limiting factor?

followed by:

3. Identification of key requirements for each case and quantification of the scientific losses in case the requirements cannot be met, thus helping the project to understand the scientific consequences of any trade-off decisions.

In addition to the generic questions above, the proposals also formulated some that were specific to each science case, such as: What is the impact of a given astrophysical effect on the final result? Are there any trade-offs to be considered, e.g. in terms of filter combinations, type of AO and field of view?

The nature of the chosen DRM cases and/or of the questions was such that they could not usually be answered by a straightforward signal-to-noise consideration. Instead, they required an appropriately realistic simulation of the data 'requested' in the proposals and subsequent analysis. This then constituted the main body of work carried out within the framework of the DRM: the provision of a suite of simulated E-ELT data, their quantitative scientific analysis, and the exploration of scientific and technical parameter spaces in order to identify the critical aspects of each science case.

In order to acquire resources to assist the SWG in implementing the DRM a dedicated DRM work package was included in the FP7 Programme "Preparing for construction of the E-ELT". This funding (over the period 01 Jan 2008 – 30 Jun 2010) led to the formation of a team of postdocs, working at ESO in the E-ELT Science Office (EScO) under the auspices of the E-ELT Project Scientist, which was tasked with carrying out the simulations and the analyses.

Each DRM case was assigned to one of the members of the EScO team, who then worked on the simulations and the analysis in collaboration with the SWG member responsible for this case (i.e. the author of its proposal) and, where necessary, with other experts from the community. In the first step of this collaboration the SWG member was asked to refine the questions posed in the proposal and to precisely define the goals of the DRM for this case. Next, one or more metrics or figures of merit had to be defined and agreed upon with which to evaluate and compare the results of different simulations with different input parameters. Where appropriate, the SWG member was also asked to specify (more or less arbitrarily) a particular value of the figure of merit to be used as the definition of 'success': any simulations for which the figure of merit exceeded this value were deemed to have met the goals of the proposal. It was obviously of particular interest to identify all simulation parameters whose variation within reasonable limits caused the figure of merit to cross this boundary.

The simulations required both scientific and technical input. The scientific input was provided by, or agreed upon with, the responsible SWG member, and essentially consisted of the distribution of astrophysical source photons as a function of position on the sky, wavelength and time (polarisation was not considered in any of the cases). The input might consist of relatively simple quantities like the brightness of target objects, their distribution on the sky, spectral characteristics, internal flux distribution, etc., and/or of more complicated quantities, such as physical models of the objects in question. On the technical side, a large number of inputs regarding the characteristics of the telescope, the site, the AO performance, etc., were required. In order to create as much homogeneity across the whole DRM as possible these data were compiled during the first stages of the DRM and collected in a central technical data repository (where they were also available to the community). These data have since been published in [RD1]. The simulations used these common data whenever possible. However, in a small number of cases the core of the simulation work had been carried out before the establishment of the technical data repository. In

these cases the technical assumptions may differ from [RD1]. In addition, since the E-ELT instrumentation suite was ill-defined at the time this work was performed (phase A instrumentation studies were carried out concurrently) the simulators were given some freedom in defining the parameters of their assumed instruments. Hence, the DRM is mostly, but not completely, consistent across all of its cases in terms of its technical assumptions.

Both the simulations and the analyses were carried out using existing software as much as possible, such as e.g. various IRAF tasks or tried-and-tested photometry packages. However, inevitably all simulators were required to write at least some amount of specialised software to complete their tasks.

In the following sections we describe in detail the work carried out for each of the DRM cases. Each section begins by summarising the science case, laying out the goals of the simulations and defining the metric or figure of merit used to evaluate the results. This is followed by a description of the simulations and the analysis procedures, an account of the scientific and technical inputs and assumptions, and finally a detailed discussion of the results. Lastly, we summarise all of the results and draw our overall conclusions in Section 19.

5 S3-1: Direct imaging of terrestrial and giant exoplanets

Authors: S. Gladysz, R. Rebolo

5.1 The science case

We propose to perform a direct imaging search for terrestrial and giant planets around 150 FGK stars within 20 pc from the Sun. Using the E-ELT, coronagraphy and image post-processing we intend to achieve a contrast of 10^{-10} beyond an angular distance of 8 λ/D (>50 mas) from each parent star. The observations will be conducted to ensure detection (S/N = 5) of the reflected light from Earth-like planets in the habitable zone of the parent stars (i.e. at 1 AU orbital distance of the solar-type stars in the sample). This will allow a rather extensive demographic study of terrestrial and giant exoplanets over a wide range of separations leading to a better understanding of the formation mechanism of planetary systems.

5.2 Goals of the DRM simulations

The primary goal of this work is to test whether the E-ELT will be capable of directly imaging Earth-like planets in the habitable zones of the parent stars (i.e. at approximately 1 AU from solar-type stars, and at smaller distances for K and M-type stars). This task is extremely challenging because it requires suppressing stellar light down to 10^{-10} of the peak PSF value at very small separations, e.g. 0.1″ for an Earth-like planet located at a distance of 10 pc.

A slightly less challenging secondary goal of the simulations is to verify whether contrasts of the order of 10^{-9} could be reached on the E-ELT. Such a capability would allow the detection of giant, Jupiter-type planets at 5 AU. In both cases we assume that detections could be considered reliable if they are obtained at the S/N = 5 level.

Direct imaging and spectroscopy of exoplanets requires, for the most challenging and interesting objects, extreme adaptive optics (XAO). The XAO instrument for detection and characterisation of exoplanets on the E-ELT is EPICS (Kasper et al. 2008). The study of exoplanets motivates the design of HARMONI (Tecza et al. 2009) and METIS (Brandl et al. 2008) too, although these two instruments will rely on the more modest adaptive optics (AO) correction provided only by the telescope (METIS will additionally rely on long wavelengths to deliver high Strehl ratio). In this report we do not focus on any specific instrument design. The goal of this work is to sample a large parameter space, and not to have a full end-to-end model of one of the planned E-ELT instrument. We want to know how parameters such as: level of static aberrations, efficiency of coronagraphic diffraction suppression, and distance to the target impact the final contrast curves.

5.3 Metrics / figures of merit

We make use of the most common metric in the high-contrast-imaging field, i.e. the contrast curve. This curve shows the level of suppression of stellar light as a function of angular distance from the star. It gives 5 σ detection limits as a function of angular separation from a star, and tells us which exoplanets are detectable at a given separation from the star.

5.4 DRM simulations

5.4.1 Methodology

We generally follow the work of Cavarroc et al. (2006) who analysed limitations for the detection of exoplanets with a general model of an ELT with a 30 m mirror. However, here we will focus on the specific design of the E-ELT.

All simulations are polychromatic and they assume the same spectral range: 0.9–1.7 μm. Planetary spectra are not simulated as part of this project but we make use of 100 channels in the afore-mentioned spectral range to calibrate and subtract residual stellar contributions after AO and coronagraphy, thus aiding the detection of the faintest planets (Section 5.4.2.6). The spectral range was pre-determined by the additional objective of this project: a cross-check with the EPICS simulation results. Originally, the 0.9–1.7 μm range was chosen by the EPICS team because it contains the most interesting molecular bands: O_2, CO_2, H_2O, and CH_4. It should also be mentioned that the planet finder instrument for the VLT, SPHERE, will have an integral field spectrograph (IFS) working in the similar spectral range (Fusco et al. 2006).

We achieve high dynamic range in the final, single image by simulating the following three techniques:

1. Image sharpening through turbulence compensation with XAO (0.2 m inter-actuator spacing). The impact of using the telescope's AO (0.5 m inter-actuator spacing) is also explored (Section 5.5.4.1). In Section 5.4.2.1 we explain how the action of the AO on the turbulent atmospheric phase can be modelled in an analytic fashion.

2. Diffraction suppression through coronagraphy. It should be mentioned here that efficient coronagraphic starlight suppression is helped by first reducing the image to a narrow spot with AO. Hence the second step in our simulations is intrinsically linked to the first. If AO works as expected then the diffraction-limited spot hits the focal-plane mask and most of the light is blocked. We consider a 'perfect' coronagraph, which is a theoretical construct removing the coherent portion of the stellar light (Fusco et al. 2006), the classical Lyot coronagraph (Lyot 1939), and the apodized pupil Lyot coronagraph (APLC; Soummer 2005). In Section 5.4.2.5 we explain the differences between these chronographs, and their impact on the achievable contrast is quantified in Section 5.5.4.2.

3. Finally we implement a post-processing technique to remove the residual stellar light from the spectral data cube (Section 5.4.2.6). Here we use the method called spectral deconvolution (Sparks & Ford 2002) which has been successfully tested on spectral data cubes recorded with the SINFONI IFS on the VLT (Thatte et al. 2007). In all our simulations we assume an IFS is present but we do not simulate spectra and we do not have a detailed model of light propagation through a spectrograph. Such a model depends on the design of an IFS (BIGRE or TIGER configuration, see Antichi et al. 2009) and is certainly outside the scope of this work. For the same reason we do not model Fresnel propagation, which depends on the number and relative separations of optical surfaces within an instrument (Marois et al. 2008).

5.4.2 Pipeline

5.4.2.1 Adaptive optics

For the wavefront sensor (WFS) we implemented the Shack-Hartmann device. This is the most widely-used WFS concept. The main component of this WFS is a lenslet array in the conjugate pupil plane. Each lenslet produces an image on the detector. The displacements of the images from the reference positions can be translated to average slopes of the incoming wavefront over individual sub-apertures (lenslets). The two main shortcomings of the Shack-Hartmann WFS relevant for high-contrast imaging are aliasing and noise propagation. Aliasing is the result of the finite sampling provided by the Shack-Hartmann WFS: the wavefront is sampled with a grid of lenslets of linear dimension d, and in theory perfect wavefront reconstruction should be possible for aberrations with spatial frequencies up to 0.5 d if the aberrated phase has no spatial-frequency content above 0.5 d. If aberrations are not band-limited, and this is always the case with atmospheric phase perturbations, aliasing occurs, i.e. high-frequency, 'unseen' content is folded back into the low-frequency range by the WFS. The control system will mistakenly try to correct these phase components which will lead to more power in the low-frequency range of the corrected wavefront. Since in the Fraunhofer approximation the image formed by a lens is given as the Fourier transform of the electric field in the pupil plane, these aliased low-order aberrations will accumulate near the centre of the PSF. This is certainly undesired because this region of the image is particularly interesting in high-contrast imaging: this is where Earth-like planets are expected to be found. The solution to the aliasing problem is relatively simple: use a low-pass filter in the focal plane before the

Shack-Hartmann WFS, thus removing the high-frequency content and making the phase band-limited. This concept is called the spatially filtered Shack-Hartmann WFS (Poyneer & Macintosh 2004) and it is used in our simulations.

All WFS work with light coming from a guide star and therefore wavefront reconstruction will be imperfect for fainter stars when photon noise starts playing a role. The second disadvantage to using the Shack-Hartmann WFS is that its capability to correctly sense low-order aberrations is weak because of photon noise (Verinaud et al. 2005). The pyramid WFS (Ragazzoni 1996) has a better response to photon noise for large-scale aberrations, and in theory could provide a gain in contrast of up to ten magnitudes for the smallest separations (Verinaud et al. 2005). However, end-to-end simulations have shown that the performance of the pyramid WFS is comparable, and sometimes poorer than that of the Shack-Hartmann device (Fusco et al. 2006). Because of that, and because of the existing simple and well-tested Shack-Hartmann implementation in the PAOLA simulation package used throughout this project it was decided that an aliasing-free Shack-Hartmann model could approximate the behaviour of a more favourable WFS to be developed during the construction of the E-ELT.

The choice of the WFS wavelength calls for a trade-off study (Fusco et al. 2006). Briefly: using visible wavelengths for sensing and near-IR for imaging causes chromatic aberrations to appear in the corrected wavefronts. On the other hand no science photons are lost when distinct spectral ranges are used for sensing and science. Also, fast, low-noise detectors for the visible band are already available on the market, while similar detectors for the near-IR are currently in the development phase. The higher background noise for near-IR wavelengths also impacts the choice, as does the fact that stars of spectral types from G5 to M5 – which are of interest in exoplanet surveys – produce less photons in the near-IR portion of the spectrum compared to the visible (Fusco et al. 2006). For these reasons the Shack-Hartmann WFS simulated in this study uses visible wavelengths in the range 0.6–0.9 μm.

The action of AO is modelled in an analytic fashion. For example, the AO fitting error can be described as the result of a high-pass transfer function acting on the Kolmogorov atmospheric power spectrum of phase. Stars of various magnitudes (m_V = 2.5 to 10.5, covering spectral types F, G, K and M at the distance of 10 pc, or G-type stars at distances 10 to 50 pc) are assumed to be observed on-axis, and therefore we consider the transfer functions corresponding to the fitting error, the WFS noise error, dispersion error, and the AO loop delay error, but we neglect the anisoplanatic transfer function. The WFS aliasing effect is also neglected because we assume a 'perfect' spatial filter in front of the WFS. The PAOLA simulation package (Jolissaint et al. 2006) was used to model the action of the AO on the atmospheric power spectrum. The modelled atmosphere consists of ten layers. The heights, fractional strengths and wind speeds in the layers were taken from the DRM technical database. The seeing was set to 0.7 arcsec at 500 nm and the zenith angle was assumed to be zero. Scintillation was neglected because this effect will be much smaller than the other effects considered in this report (Fusco et al. 2006). Generally, no amplitude errors – such as the effect of dust on the optical surfaces – have been considered here. The parameters of the simulations are summarised in Table 5.1.

The AO-filtered phase spectrum was used to generate independent realisations of the AO residual phase using a classic, Fourier-domain filtering algorithm (McGlamery 1976). To this phase two static components have been added: static aberrations from the main mirror's segments, and the non-common-path aberrations not sensed by the AO and modelled with an f^{-2} spectrum (with f being the spatial frequency) (Cavarroc et al. 2006). The next section describes the nature of the optical aberrations to be expected on the E-ELT and the simplifications used in the simulations.

5.4.2.2 Static aberrations

The E-ELT has a five-mirror design and naturally all five mirrors will introduce some aberrations. Nevertheless in our simulations we consider only M1 aberrations because these will be by far the most significant. The same simplification has been used in the thirty Meter Telescope (TMT) XAO simulations (Troy et al. 2006; Crossfield & Troy 2007). Below we describe the strategies for phasing and control of M1 which are currently being investigated at ESO. This is by no means a complete description, and the reader is encouraged to consult the relevant documents referenced in this section. The main procedures are defined here as follows:

- M1 phasing through optical sensors, segment position actuators and warping harnesses,
- M1 control through edge sensors and position actuators.

M1 phasing Phasing of the main mirror serves two purposes: (i) to remove piston and tip/tilt errors between segments, and shape errors on the individual segments, thus increasing the coherency of the wavefront; (ii) to calibrate set points of the edge sensors (ES).

The baseline concept for the phasing sensor is a Shack-Hartmann WFS with 19 inner sub-apertures covering the segment, and 12 phasing sub-apertures straddling the inter-segment edges (Bonnet 2009a). The phasing sub-apertures will sense the distortion of the diffraction pattern due to the inter-segment step errors. To do so the sensor must be operated in the diffraction-limited regime which will be the case since the sub-aperture size is of the order of the Fried parameter r_0 at the operating wavelength of 800 nm (Bonnet 2009a). The edge measurements have two main advantages over the traditional Shack-Hartmann slope measurements (Chanan 2009). Firstly, they are only sensitive to wavefront discontinuities which can only come from the segmented primary mirror, and therefore they are not affected by the aberrations on the other mirrors. This removes the complication of having to disentangle which aberrations came from which mirror. Secondly, for the same reason they are insensitive to the atmospheric aberrations. Since the distortion of the PSF is controlled by the relative edge height, the phasing algorithm compares the measured PSF shape to the one predicted by diffraction theory. Actuator commands deduced from the piston values are then sent to the segment position actuators (PACTs). Three PACTs control piston and tip/tilt of one segment.

The inner sub-apertures will be used to sense the shape errors of the segments, for example: defocus and astigmatism aberrations which could be produced by the spatial inhomogeneities of the coefficient of thermal expansion within one segment (Swat, priv. comm.). The measurements will be transformed to signals for the warping harnesses (WH) which will control the shapes of the segments with 24 actuators (Müller 2008). There is some ambiguity regarding the WH reduction factors (rms output aberration / rms input aberration) and the spatial structure of the resulting segment error (Troy et al. 2006; Müller 2008). Here we will assume the WH reduction factors predicted by simulations carried out at ESO (Müller 2008). They translate to nm-level aberrations after correction of realistic input loads due to gravity or thermal effects. Regarding the spatial frequency content of the output error we will follow ESO predictions and assume the output segment shape to be independent of the input shape. For example, correction of the simple tilt due to gravity results in a complicated, high-spatial-frequency shape which bears no resemblance to the input aberration (Müller 2008). Since the simulated output segment shapes are characterised by relatively high spatial frequencies, between approximately 5 and 10 cycles per segment, the resulting speckles in the PSF will cover the outer regions of the image which are less interesting from the high-contrast point of view (see beginning of Section 5.4.2.1). In our simulations we therefore ignore correction limitations of the WH.

As regards the second function of the phasing procedure, the segment alignment process described above calibrates the edge sensors on the segments. An edge sensor is a combination of one emitter and one receiver belonging to two adjacent segments. In the current E-ELT design there are two ES pairs per segment edge. ES will be used to maintain the figure of M1 which will be perturbed by wind buffeting, thermal effects and gravity. ES are only sensitive to inter-segment piston variations. Because they deliver a relative measurement they have to be calibrated. At the E-ELT this will be done using the optical phasing sensor described above. The role of ES in M1 figure control is explained below. Here we will only mention that their accuracy can reach sub-nm level (Dimmler 2009; Bonnet 2009b). As explained below, it will be the accuracy of the optical phasing with a WFS, rather than the ES measurement precision, that will dominate the error budget of the aligned, off-line telescope.

The level of phasing precision to be expected at the E-ELT is open for debate. On the one hand results from phasing of the Keck telescope point to the possibility of achieving an accuracy of tens of nm (Crossfield & Troy 2007). This seems to be supported by the results of the Active Phasing Experiment at ESO (Gonte 2009). On the other hand simulations of ELT alignment yield post-correction rms values on the order of hundreds of nm (Chanan 2009), or μm (Piatrou & Chanan 2010), of wavefront error on M1. Importantly, the aberrations – or rather mirror modes – left over after alignment are all of the low-order type: global tip/tilt, defocus, astigmatism, etc. (Piatrou & Chanan 2010). They are left uncorrected on purpose

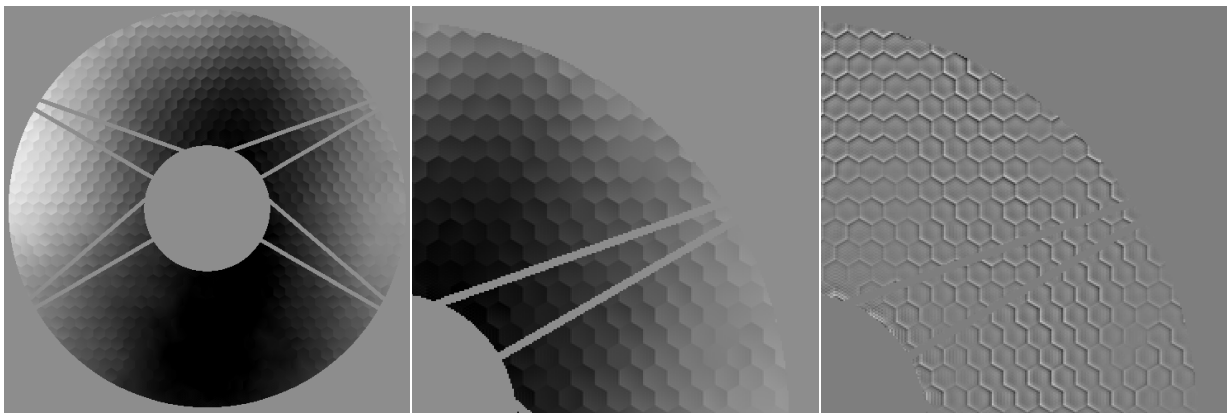

Figure 5.1: Left: One instance of M1 phase aberrations due to wind buffeting before XAO correction. Middle: Close-up of one quadrant of M1. Right: same part after XAO correction. The rms error has been reduced from 0.3 μm to 8 nm. The first two images are displayed on the same linear scale, between -4π and 4π. The last image is displayed on the linear scale between -0.5π and 0.5π.

because their 'observability' is low: attempts to measure them would result in a noisy reconstruction and in oscillations of the global shape of M1, which would have a very strong effect on image quality. Instead they are allowed to propagate through the main optics, up to the AO system which in principle should have excellent response to low-order aberrations. As discussed in the next section the simulations presented in this report implement low-order, high-amplitude M1 aberrations.

In the currently adopted strategy phasing of the E-ELT primary mirror will be performed about once a month. Provisions are made for a back-up strategy where phasing is done also during the science observations. At the TMT phasing will be carried out at monthly intervals outside the science queue. The procedure is expected to last no longer than two hours (Piatrou & Chanan 2010).

M1 control During the science operations dynamical perturbation of the M1 figure will be measured by ES and corrected with PACTs with a bandwidth of 1–3 Hz. Naturally, the resulting accuracy should be better for slow aberrations. These are due to temperature variations (current ESO assumptions: seasonal amplitude of ambient temperature variations of $\pm$15 K; ambient temperature change during tracking of $\pm$1 K/h), and gravity-load changes during tracking (E-ELT's operational pointing range is 0 to 70° zenith angle). In simulations, after execution of the ES-PACT loop for these quasi-static perturbations the main mirror assumes shapes characterised by low-spatial-frequency aberrations (Bonnet 2009c). The same is true for fast perturbations due to wind buffeting (ESO assumptions: wind speed = 10 m/s outside, and 1.6 m/s inside the dome), although the degree of correction is not spectacular. Similar to the phasing loop, the M1 controller does not attempt to remove low-order aberrations – these are to be corrected by the AO. The role of the M1 controller is to compensate mainly for high-spatial-frequency errors which have small amplitudes and therefore their contribution to the rms wavefront error is also small.

It is expected that the wind-induced dynamical aberrations will dominate over aberrations generated by gravity and thermal effects. Hence for our simulations we adopt a model of the segments' piston, tip and tilt from a scenario of a closed-loop M1 control system responding to wind disturbances alone (Dimmler 2009). ES noise is included in this scenario and perfect telescope phasing is assumed. The full model contains piston and tip/tilt values for every segment during the interval of 30 s, sampled every 5 ms. From this set we take only one 'instantaneous' M1 shape for our simulations. What is important here is the spatial structure of the closed-loop aberrations shown in Fig. 5.1. These low-order distortions are a good representation for a class of slowly-varying perturbations to be expected to leak through M1 control. It is important to have them in an XAO simulation since their residuals after AO will focus light close to the PSF core, which is the most interesting region in the image.

As will be explained later, high-contrast imagery is dominated by quasi-static speckles which are the result of quasi-static aberrations. While the wind-induced perturbations have timescales less than one minute, which is the timescale adopted in our simulations, and temporal noise averaging in the final

image would be in practice quite efficient, we argue that the complete set of controllable aberrations (gravity, temperature variations and wind) could be assumed constant within the time-span of one minute. In the future we plan to integrate into the simulations models of these aberrations under the action of the M1 controller, when such models become available. Currently we assume that M1 aberrations are completely static on the time-scale of one minute.

5.4.2.3 Post-AO aberrations

The spatially-filtered WFS employed in the simulations has a 'perfect' transfer function, i.e. it sets to zero all aberrations up to its cut-off frequency, which by the Shannon sampling theorem is 0.5 d for the inter-actuator spacing d. In the XAO case this means no PSF deformations (speckles) should be present up to 0.45″ from the PSF centre. Still, in real systems speckles do appear within the AO control region: they are due to non-common-path aberrations, which are not sensed and therefore not compensated for by the AO (Cavarroc et al. 2006; Soummer et al. 2007). These instrumental errors are usually slowly-varying and they generate quasi-static speckles, which are the biggest obstacle in detecting planets. They are too slow to efficiently averaged out in long exposures, and too fast to be completely subtracted out through PSF calibration (Marois et al. 2006; Racine et al. 1999).

We simulate these aberrations with an f^{-2} spectrum which has been empirically obtained on optical surfaces. The spectrum is flat in the range $0 < f < 0.5\ d$. Beyond the cut-off frequency it decreases towards zero. Similarly to the case of atmospheric aberrations this spectrum is used to filter white noise to give one realisation of non-common-path aberrations, i.e. a phase-screen (McGlamery 1976). The total rms power in this phase-screen is 30 nm. This adopted value is consistent with commonly assumed levels of non-common-path aberrations (Cavarroc et al. 2006; Marois et al. 2008; Verinaud 2010). This phase remains static during one simulation run.

Here we come to the problem of choosing the time-scale for the quasi-static aberrations. In this work we obtain images corresponding to one minute of integration time, compute contrast curve, and then scale this curve to 10 hours. The total variance of speckle noise at some separation from the PSF centre in an image corresponding to 1 m is:

$$\sigma_f^2 = \sigma_{\rm qs}^2 + \frac{\tau_{\rm qs}}{\tau_{\rm a}}\sigma_{\rm a}^2, \tag{5.1}$$

where $\sigma_{\rm qs}$ is the standard deviation of quasi-static speckles at the location under investigation, $\sigma_{\rm a}$ is one realisation of AO-filtered atmospheric speckles, and $\tau_{\rm qs}$ and $\tau_{\rm a}$ are the corresponding time-scales (1 m and 0.5 s, respectively – see Section 5.4.2.4). In 10 h the noise variance will be:

$$\sigma_f^2 = \frac{10\,{\rm h}}{\tau_{\rm qs}}\left(\sigma_{\rm qs}^2 + \frac{\tau_{\rm qs}}{\tau_{\rm a}}\sigma_{\rm a}^2\right). \tag{5.2}$$

What happens when $\tau_{\rm qs}$ is in reality twice as long as the chosen value?

$$\sigma_f^2 = \frac{10\,{\rm h}}{2\tau_{\rm qs}}\left(4\sigma_{\rm qs}^2 + \frac{2\tau_{\rm qs}}{\tau_{\rm a}}\sigma_{\rm a}^2\right) = \frac{10\,{\rm h}}{\tau_{\rm qs}}\left(2\sigma_{\rm qs}^2 + \frac{\tau_{\rm qs}}{\tau_{\rm a}}\sigma_{\rm a}^2\right), \tag{5.3}$$

(The same argument could be applied to $\tau_{\rm a}$). This means that the results of simulations based on too short time-scales will under-estimate the noise in very long exposures. Unfortunately there is very little information in the literature about the temporal properties of long-lived speckles. The reported values range from a few seconds to a few hundred seconds (Hinkley et al. 2007). As mentioned earlier we adopt a value of 60 s, which is probably optimistic. To illustrate the effect of two different time-scales of quasi-static aberrations we also ran one simulation for an integration time of 30 s. Fig. 5.2 shows the impact of the choice of $\tau_{\rm qs}$ on the final contrast curve, after coronagraphy (Section 5.4.2.5) and spectral deconvolution (Section 5.4.2.6). As expected the gain in contrast caused by assuming a shorter $\tau_{\rm qs}$ of 30 s is slightly smaller than the factor of two (cf. equations 5.2 and 5.3). Since we expect our order-of-magnitude guess for $\tau_{\rm qs}$ to be optimistic but reasonable, the results reported here will also have an order-of-magnitude correspondence to real observations. In the future we plan to investigate the auto-correlation of quasi-static speckles induced by a full model of M1 aberrations (dynamical wind perturbations, quasi-static PACT response to gravity while tracking, etc.) similarly to work that has been carried out for filtered atmospheric aberrations after XAO (Macintosh et al. 2005; Poyneer & Macintosh 2006).

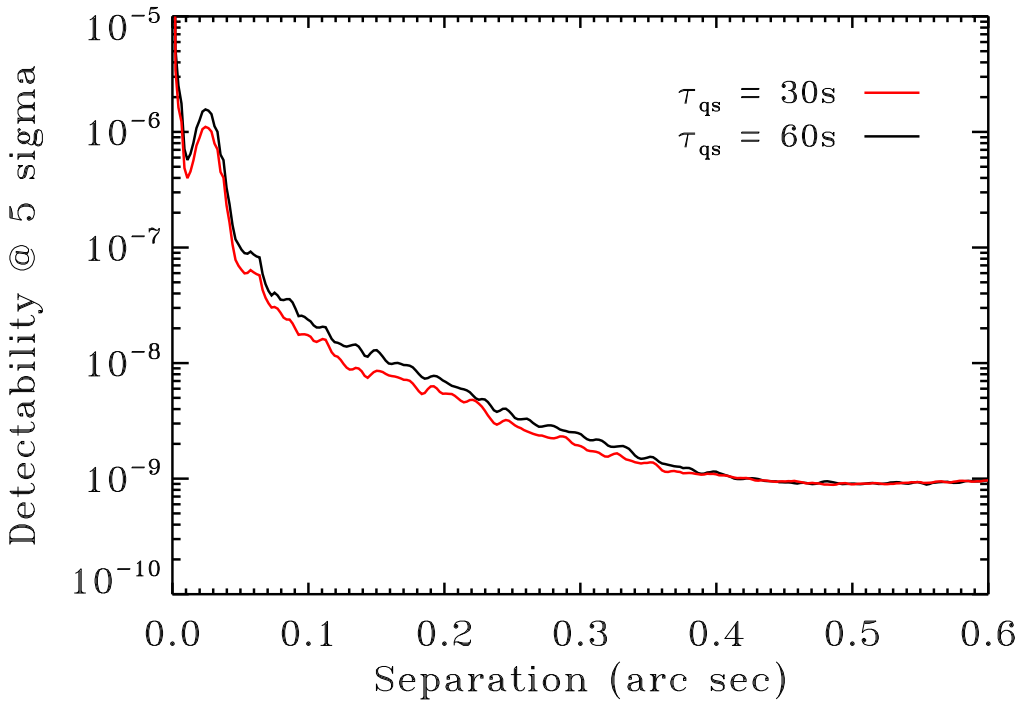

Figure 5.2: Comparison of two simulation runs corresponding to periods of 30 and 60 s. Final contrast is better for the shorter simulation. Star G2 (m_V = 4.7 mag), distance 10 pc, APLC coronagraph, 10 h integration time.

5.4.2.4 Integral field spectrograph

To simulate an IFS we needed to have the wavelength-scaled wavefronts. After one AO-corrected wavefront is generated for the first wavelength of 0.9 μm, the simulation loops over the 100 channels in the spectral range, scaling the stored phase by each channel's wavelength:

$$\phi(\lambda_i) = \phi(\lambda_0)\lambda_0/\lambda_i, \tag{5.4}$$

where ϕ stands for phase of the wavefront, and λ is wavelength (λ_0 = 0.9 μm). In the simulation this loop over wavelength is contained inside the loop over time: 120 iterations over instantaneous wavefronts. The resulting instantaneous PSFs are scaled by the flux expected for an integration time of 0.5 s. We take this as a guess for an XAO speckle lifetime. It has been shown (Macintosh et al. 2005; Poyneer & Macintosh 2006) that second-order speckles, which dominate coronagraphic images, are significantly slower than pure atmospheric speckles. These XAO coronagraphic speckles have timescales set by the atmospheric clearing time of the telescope, $\tau \approx D/\nu$, where ν is the wind speed. The value of 0.5 s, while consistent with estimates reported for simulated XAO systems on 8 m-class telescopes (Macintosh et al. 2005), is probably optimistic for a 42 m aperture. Nevertheless, the characteristics of the contrast curves in the region of interest are dominated by the quasi-static aberrations and not by the smooth halo generated by AO-filtered atmospheric perturbations.

Equation (5.4) excludes differential chromatic aberrations between the channels. M1 was sampled with 1024 $\times$ 1024 pixels inside a 2048 $\times$ 2048 zero-filled array. To keep the constant plate scale ($\lambda/(2D)$ at 0.9 μm) the wavefronts for consecutive wavelengths were embedded in ever-increasing arrays of zeros before Fourier-transforming. For an m_V = 4.7 mag star the wavefronts containing AO-filtered atmospheric and M1 aberrations, together with non-common-path-errors resulted in Strehl ratios between 85 and 96 % in the spectral range 0.9–1.7 μm. In the SCAO simulations these Strehl ratios were 55 and 85 %.

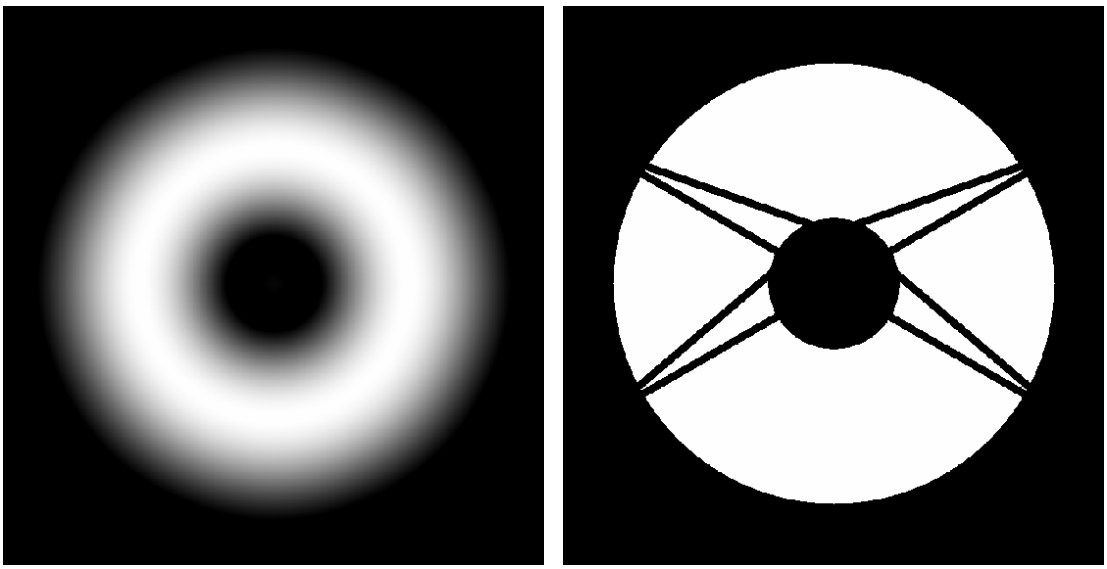

Figure 5.3: Left: APLC apodizer. Right: APLC pupil-plane stop.

5.4.2.5 Coronagraphs

Firstly we consider the 'perfect' coronagraph in order to explore the theoretical limitations to high-contrast imaging at the E-ELT. One can model such a coronagraph as a subtraction in the pupil plane of the perfect, flat wavefront (Fusco et al. 2006). Since coronagraphs can only remove the coherent part of the incoming electric field, as quantified by the Strehl ratio, this idealised subtraction is less efficient for less coherent wavefronts:

$$\mathrm{PSF} = \left| FT\left(Pe^{i\phi} - \sqrt{\mathrm{SR}} Pe^{i0} \right) \right|^2 , \tag{5.5}$$

where ϕ stands for phase, P is the binary pupil function, and SR stands for the Strehl ratio. This 'perfect' coronagraph is completely achromatic and therefore the resulting multi-wavelength images can be very efficiently 'cleaned' by spectral deconvolution (Section 5.4.2.6).

The action of a more realistic Lyot-style coronagraph is simulated in the following way. For each wavelength the electric field is Fourier-transformed to give the amplitude PSF in the focal plane. The result is multiplied by the representation of the focal-plane mask, taken to be 7.5 λ/D in diameter. This value has been optimised for the inner working angle of 3.9 λ/D and the current design of the E-ELT main mirror, i.e. 30 % linear obscuration ratio (Martinez et al. 2008). The electric field after the focal-plane mask is inverse Fourier-transformed and multiplied by the model of the Lyot stop. For the Lyot coronagraph this stop has an over-sized central obscuration and undersized outer edge. Diffraction patterns induced by the spiders are also covered with masks of twice the size, i.e. 0.8 m. The throughput of the Lyot coronagraph is 63 %. The corresponding attenuation of the peak of the planetary PSF is included in the computation of the contrast curve.

The more suitable APLC coronagraph has an apodizing mask in the first pupil-conjugate plane after XAO. This mask suppresses the Airy rings in the subsequent focal plane, and scatters non-coherent light outside the next pupil-conjugate plane where the stop is located. That is why for the APLC coronagraph the stop is not undersized. Only the diffraction from the spiders is blocked with masks of twice the size. Transmission of the APLC coronagraph is determined by the apodizer and reaches 55 %. The design was optimized for the current E-ELT configuration and the inner working angle of 2.4 λ/D (Martinez et al. 2008). Fig. 5.3 shows the apodizer and the pupil-plane stop for the APLC configuration.

The Fourier transform of the field after the Lyot stop gives the electric field in the detector plane, and the squared modulus of this quantity yields the final on-axis instantaneous PSF. For the images of the test planets the electric field is propagated only through the Lyot stop (and the apodizer for APLC), giving the off-axis PSF. We scale and shift this off-axis PSF to simulate seven planets located at various separa-

tions from the star, within the range of 0.1–0.5″. The contrast ratios are between 10^6 and 10^7. These instantaneous images of planetary systems are stored for each channel and then co-added, resulting in a spectral data cube with 100 images, each 60 s long. Finally, we loop over this cube adding Poisson, background (15 mag/arcsec^2) and read-out noise (10 e$^-$).

Spatial dependence of the coronagraphic transmission has to be taken into account when spatially-dependent metrics, such as contrast curves, are considered. For this, one has to simulate off-axis propagation with tilted wavefronts. We do this be adding a phase term in Fourier space to a perfectly flat wavefront. This tilted wavefront is then propagated through a coronagraph. This is done for a range of tilts. Transmission as a function of the angular separation in the focal plane is then stored for both the Lyot and the APLC designs and it is used in the computation of the contrast curves.

5.4.2.6 Post-processing

Spectral deconvolution is applied to the IFS data cube to increase the contrast by two to three orders of magnitude and recover the signal of the faintest test planets. Briefly: in the simulated data cube the residuals of the Airy pattern expand with wavelength, and the planet stays fixed. Each image in the cube is spatially scaled by Fourier interpolation to fit a common grid, so that the Airy rings are now aligned and the planet's position changes through the cube: it gets closer to the star with increasing wavelength. A 2nd-order polynomial is fitted for each spaxel (spectral vector in the scaled cube). This particular choice of the polynomial order is justified by the following argument: the wavefront expressed in units of length, which to a first order can be assumed to have the same optical path difference across the 100 channels, is scaled by $2\pi/\lambda$ to give the phase. Theoretically, the response of intensity for one constant location on an Airy ring to a change in wavelength of the wavefront should be $\propto (2\pi/\lambda)^2$, because image intensity is the power spectrum of the electric field. This argument ignores changes in PSF morphology which naturally occur for a wide wavelength range where the Strehl ratio changes significantly. Nevertheless, these changes are smooth when one considers many channels, as we do here, and can be fitted with a moderately smooth polynomial. After the first fitting step, any outliers at the 2 σ level (possible planets) are removed from the vector, and this procedure is repeated until no value in the sample exceeds the fit by two standard errors. Once the fit is not skewed anymore by possible signals, it is subtracted from the spaxel under consideration – this step removes most of the stellar modulation, but leaves possible planetary signal intact. The images are then re-scaled to the original grid, so that now a planet would be aligned on the same pixel again. Finally the cube is collapsed to produce one final image.

5.4.3 Inputs

5.4.3.1 Scientific data

Young and hot exoplanets are the primary targets for the SPHERE instrument (Boccaletti et al. 2008). XAO at the E-ELT will be geared towards the detection of old and cold planets, similar to the ones in the Solar System. These targets are significantly more challenging. The contrasts to be expected are between 10^7 and 10^{10}, depending on the size of the planet and its physical distance from the star. Old planets do not emit their own light; they only reflect light generated by the stars. Therefore the planet/star flux ratio, F_p/F_s, is (Brown 2004):

$$\frac{F_p}{F_s} = p\,(\lambda)\,\Phi\,(\beta)\left(\frac{R}{r}\right)^2, \tag{5.6}$$

where p is the geometric albedo of a planet, which depends on wavelength, R stands for the radius of an exoplanet, r is its separation from the star in AU, and $\Phi\,(\beta)$ is the phase function of the planet which depends on the phase angle β (angle at the planet between star and observer):

$$\Phi\,(\beta) = \frac{\sin(\beta) + (\pi - \beta)\cos(\beta)}{\pi}, \tag{5.7}$$

Assuming an optimistic case of a planet observed at opposition ($\Phi(\beta)$ = 1) and a near-IR albedo of 0.3 we get for Jupiter (R = 0.00048 AU, r = 5.2 AU) the expected contrast of 2.5 $\times$ 10^{-9}. For Earth we get

5.6×10^{-10}. In both cases we assumed distances between planets and their parent star identical to the ones in our system. What happens when one considers smaller stars?

The habitable zone is expected to be found closer for smaller stars (Selsis et al. 2007). For M-type stars it could be located as close as 0.1 AU. Assuming the same spatial scaling applies to other, heavier planets in this hypothetical planetary system around an M-type star we get contrasts of 2.5×10^{-7} for 'Jupiter' and 5.6×10^{-8} for 'Earth'. 'Super-Earths', i.e. planets with masses several times greater than the Earth's mass would have contrasts around 2.6×10^{-7} in this M-type star's system.

5.4.3.2 Technical data

This DRM science case requires more technical data than is provided in the DRM technical database. The additional parameters and their values are listed in Table 5.1.

Table 5.1: Values of the parameters used in the simulations.

Input parameters	XAO simulation	SCAO simulation[a]
	Telescope	
Telescope diameter	42 m	
Central obscuration	12.5 m	
Thickness of the spiders	0.4 m	
Number of segments	984	
M1 aberrations due to wind[b]	0.3 μm rms piston 0.006″ rms tip 0.006″ rms tilt	
	Atmosphere	
Seeing at 0.5 μm	0.7 arcsec	
Outer scale	25 m	
	Sky	
Background noise in H-band	15 mag/arcsec^2	
	AO system	
Type of WFS	Spatially filtered Shack-Hartmann	Spatially filtered Shack-Hartmann
Number of WFS sub-apertures	200×200	84×84
WFS integration time	0.33 ms	2 ms
AO lag[c]	1 ms	6 ms
WFS read-out noise	0.5 e$^-$	3 e$^-$
WFS spectral range	0.6–0.9 μm	0.6–0.9 μm
Non-common-path aberrations[d]	30 nm rms	30 nm rms
	Integral field spectrograph	
Spectral range	0.9–1.7 μm	
Number of spectral channels	100	
Optical transmission to the detector	0.1	
Read-out noise	10 e$^-$	
Pixel scale	$\lambda/(2D)$ at 0.9 μm = 0.002″	

[a] SCAO refers here to the AO correction provided by the telescope.
[b] rms wavefront: these are correlated, low-order aberrations on the main mirror, left uncorrected by the M1 control system, see Section 5.4.2.2.
[c] AO lag equals WFS read-out time + wavefront reconstruction time.
[d] Simulated with an f^{-2} power spectrum: flat in the AO working range, and decreasing beyond the AO cut-off frequency.

Table 5.2: Simulation runs.

Run number	Values of the investigated parameters (cf. Table 5.1)
1	APLC, SCAO
2	APLC, XAO
3	Perfect coronagraph, XAO
4	Lyot coronagraph, XAO
5	APLC, G2, 10 pc
6	APLC, G2, 30 pc
7	APLC, G2, 50 pc
8	APLC, F2, 10 pc
9	APLC, K2, 10 pc
10	APLC, M2, 10 pc
11	APLC, G2, 10 pc, $\tau_{\rm qs}$ = 30 s
12	Perfect coronagraph, XAO, G2, M1 aberrations = 0.15 μm rms
13	Perfect coronagraph, XAO, G2, M1 aberrations = 0.9 μm rms
14	APLC, XAO, G2, M1 aberrations = 0.15 μm rms
15	APLC, XAO, G2, M1 aberrations = 0.9 μm rms
16	Perfect coronagraph, XAO, G2, post-AO aberrations = 10 nm rms
17	Perfect coronagraph, XAO, G2, post-AO aberrations = 60 nm rms
18	APLC, XAO, G2, post-AO aberrations = 10 nm rms
19	APLC, XAO, G2, post-AO aberrations = 60 nm rms
20	Perfect coronagraph, SCAO
21	Perfect coronagraph, XAO400
22	APLC, XAO400
23	Perfect coronagraph, XAO400, M2

5.4.4 Outputs

The single image resulting from spectral deconvolution (see right panel in Fig. 5.4) is processed to yield azimuthally averaged contrast curves. The contrast curve is simply a linear function of the noise standard deviation, normalised by the peak of the on-axis un-obscured PSF integrated over all wavelengths. It gives 5 σ detection limits as a function of angular separation from a star. Fig. 5.5 shows the contrast curves that correspond to the images shown in Fig. 5.4.

5.5 Results of simulations

5.5.1 Simulation runs

In Table 5.2 we summarise the various simulation runs performed for this DRM science case.

5.5.2 Analysis

In Figs. 5.6 and 5.7 we show the contrast curves obtained with the standard set of simulation parameters presented in Table 5.1 and the APLC coronagraph for a G-type and an M-type star located at a distance of 10 pc, respectively. We over-plot expected fluxes from the three classes of planets discussed in Section 5.4.3.1. If a symbol representing an exoplanet is above a contrast curve it means that this particular type of planet would be detected at the S/N = 5 level.

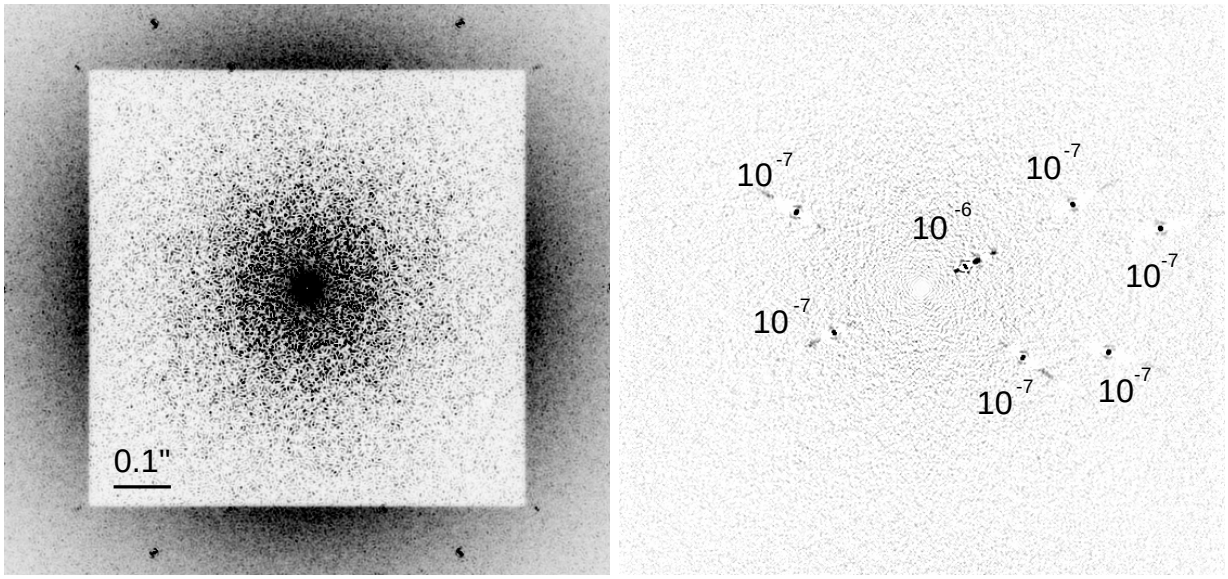

Figure 5.4: Left: A single 60 s long exposure at 0.9 μm. Right: Image formed by collapsing the full spectral data-cube after processing using spectral deconvolution; all seven planets are now clearly visible. Star G2 (m_V = 4.7 mag), distance 10 pc, 'perfect' coronagraph, 1 m integration time. Images are displayed on a linear scale, and artificially saturated to bring out the faint details. Colours are inverted for better contrast.

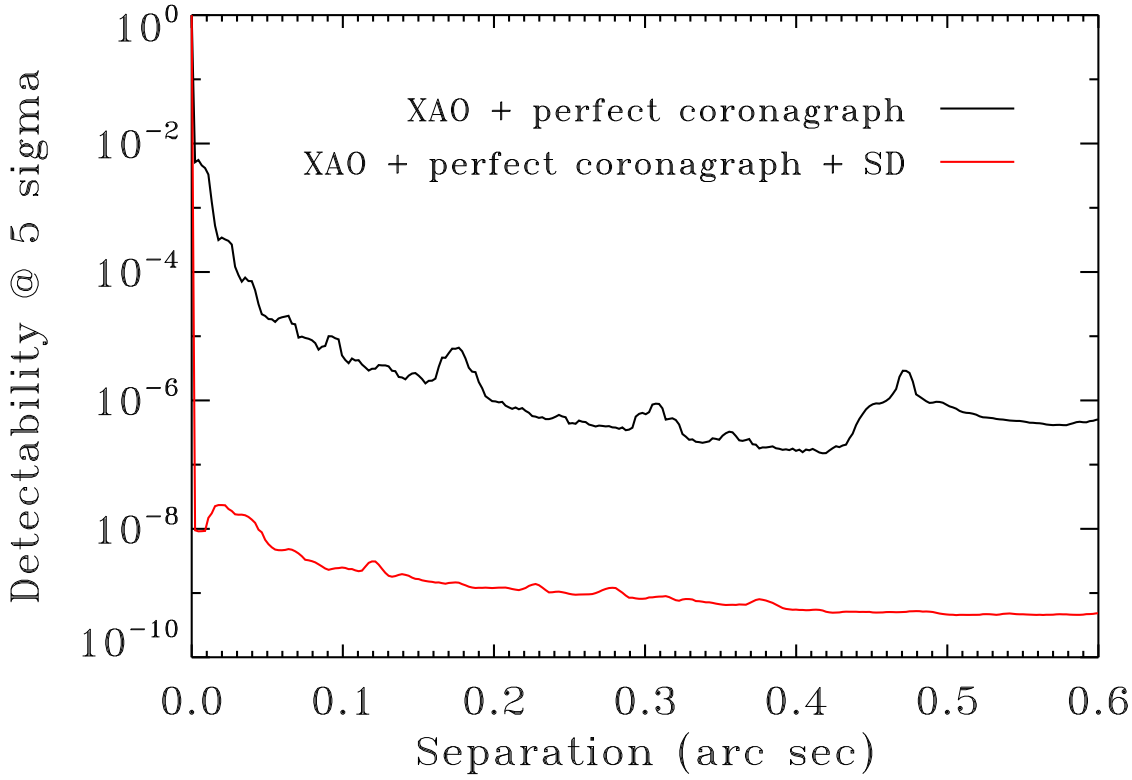

Figure 5.5: Contrast curves before, and after spectral deconvolution (SD), for the case introduced in Fig. 5.4 and scaled to 10 h.

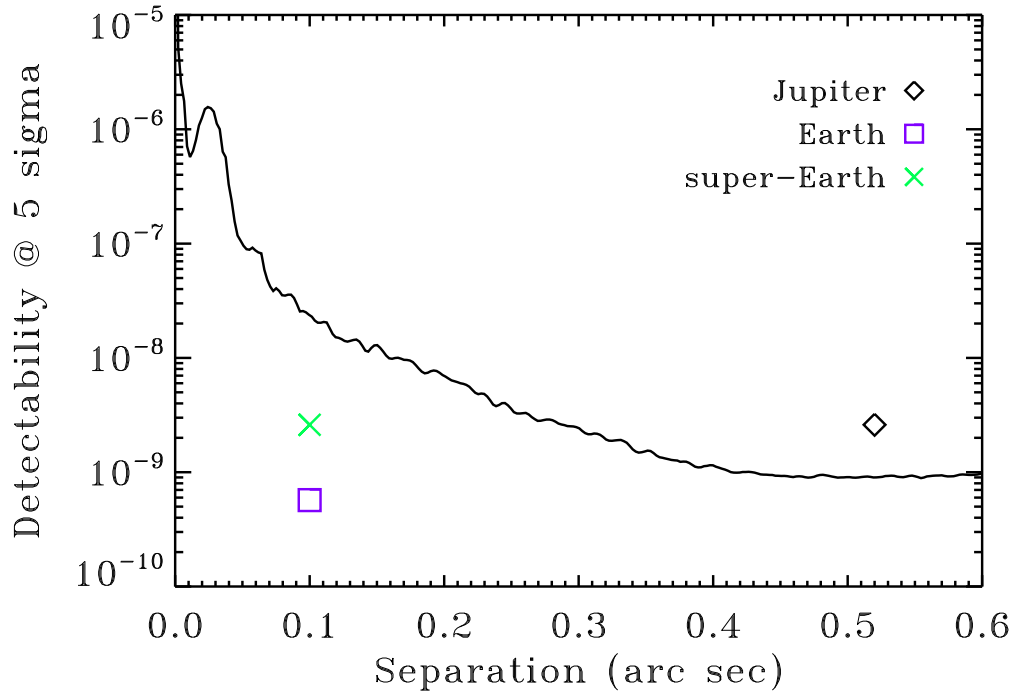

Figure 5.6: Contrast curve for a G-type star at a distance of 10 pc, APLC coronagraph, integration time = 10 h. Expected fluxes from three types of planets are over-plotted.

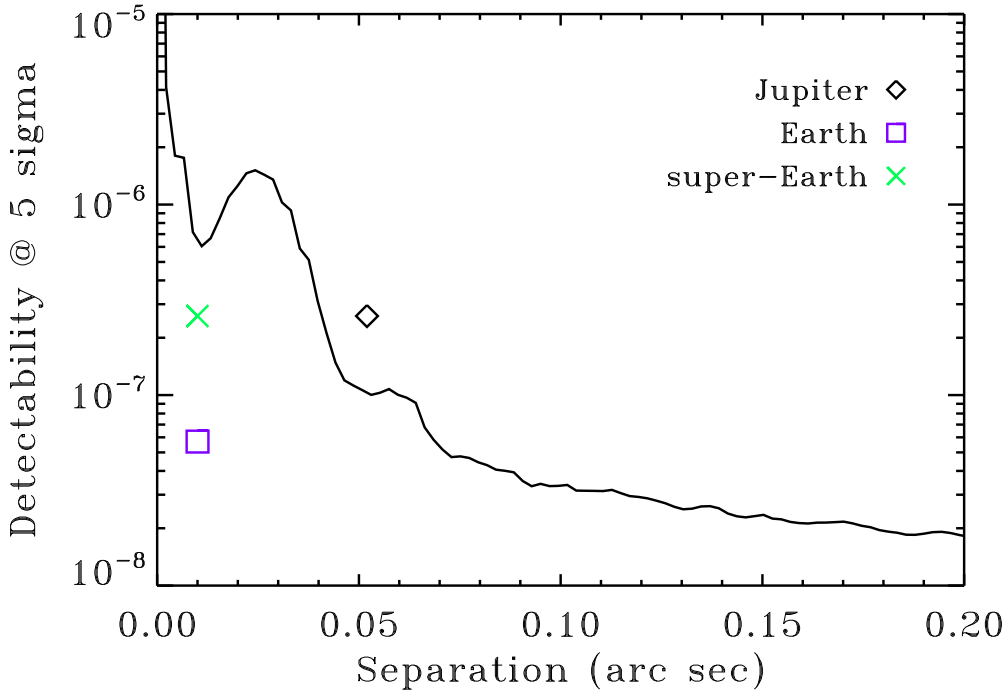

Figure 5.7: Contrast curve for an M-type star at a distance of 10 pc, APLC coronagraph, integration time = 10 h. Expected fluxes from three types of planets are over-plotted.

5.5.3 Compliance with figures of merit

Fig. 5.6 suggests that while it is likely that the E-ELT will make many discoveries of Jupiter-type, giant planets it seems that obtaining images of Earth-like planets will not be feasible with the XAO system, and especially with the APLC coronagraph used throughout this report.

The same analysis repeated for an M-type star (Fig. 5.7) shows the challenges associated with imaging of old and cold planets. It seems once again that direct imaging of rocky planets in the habitable zone would require a light-suppression system which is an order-of-magnitude more efficient than our combination of APLC and spectral deconvolution.

5.5.4 Sensitivity to input parameters

In the following sections we show how the change of a particular parameter impacts the final contrast curve. All contrast curves shown here correspond to an integration time of 10 h.

5.5.4.1 XAO vs. SCAO

Here we look at the gain afforded by increasing the actuator density. We ran one simulation with the SCAO system (cf. Table 5.1), and also two XAO simulations: one with the 200 $\times$ 200 actuator system, and one with XAO having 400 $\times$ 400 actuators. The other parameters of this 'super XAO' system were the same as for the 200 $\times$ 200 actuator system. The simulations were then repeated for the 'perfect' coronagraph.

Fig. 5.8 suggests that at small separations – less than 0.05″ – the APLC contrast curve seems to be independent of the level of AO correction. This is due to coronagraph chromaticity, which results in non-trivial spatial scaling of the images which cannot be removed by spectral deconvolution. Further away from the core the clearing of the halo by XAO brings an advantage of approximately one order of magnitude. This is important for giant planets which are expected to have star/planet contrasts of around 10^9.

SCAO is not sufficient if one wants to detect planets in reflected light. This can be seen in the curves for the 'perfect' coronagraph and the APLC. On the other hand the effect of increasing the actuator density above 200 $\times$ 200 is very small. This can be understood if one remembers that decreasing the WFS lenslet size gives less photons for the wavefront reconstruction operation. While the fitting error will decrease, the WFS noise error will increase: the XAO400 system's WFS consists of lenslets with size equal to 10 cm, and records stellar sub-images at a frame rate of 3 kHz. The Strehl ratios at the longest wavelength (1.7 μm) are 96 and 98 % for XAO200 and XAO400, respectively, and the difference in the final contrast

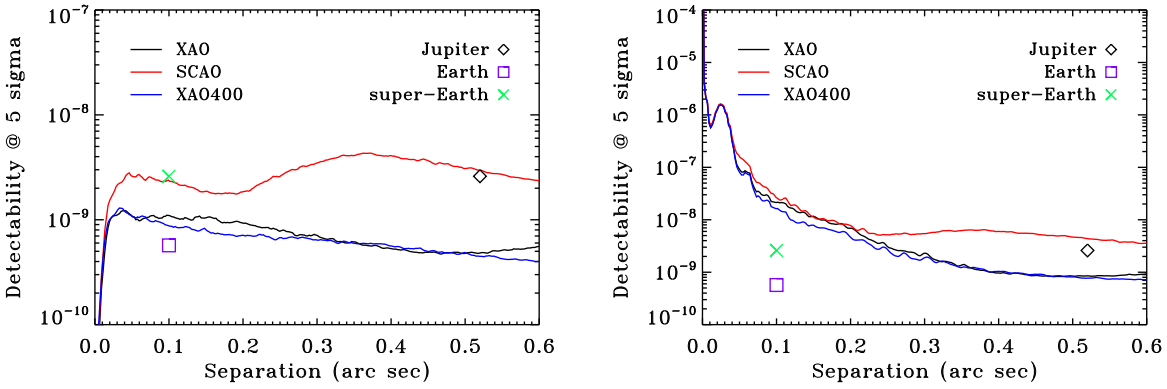

Figure 5.8: Left: Effect of varying the number of actuators on the contrast curve for a G-type star at a distance of 10 pc, 'perfect' coronagraph. Right: Same analysis for the APLC coronagraph.

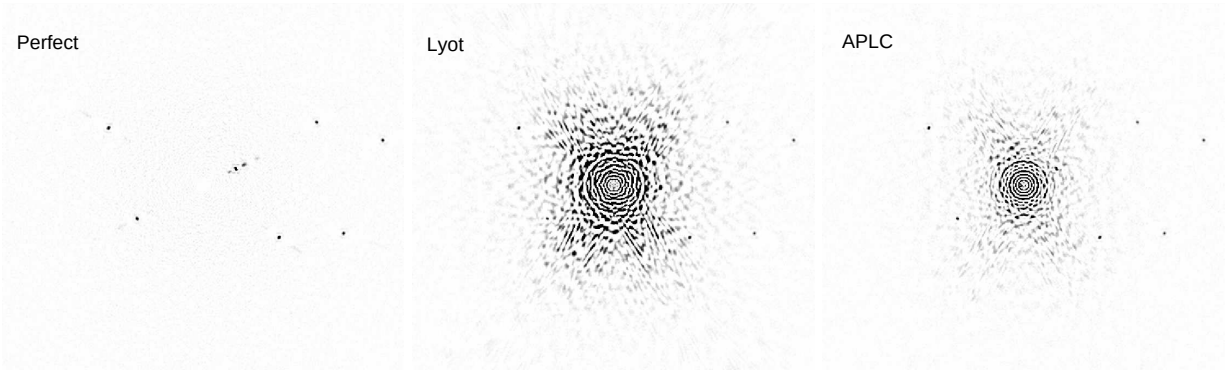

Figure 5.9: Results after spectral deconvolution for three different coronagraphs. Star G2 (m_v = 4.7 mag), distance 10 pc. Images are displayed on the same linear scale, with identical minimum and maximum values. Colours are inverted for better contrast.

is minimal. Going from a 200 $\times$ 200 actuator system to a 400 $\times$ 400 XAO will not be helpful as far as the detection of Earth-like planets is concerned.

5.5.4.2 Coronagraphs: 'Perfect' vs. APLC vs. Lyot

Spectral deconvolution can in theory remove post-coronagraphic residuals as long as they are achromatic. The resulting image should be free of speckle noise. This is not the case with real, chromatic coronagraphs as illustrated in Fig. 5.9. Chromaticity results in non-trivial spatial scaling of the images in the spectral data cube. This is visible primarily in the vicinity of the optical axis and this is where both the Lyot and APLC designs have significantly lower contrasts compared to the 'perfect', achromatic coronagraph.

Fig. 5.10 emphasises this point. Close to the optical axis chromaticity of the coronagraph is the 'bottleneck'. The spatially varying transmission of the coronagraph, i.e. the presence of the focal-plane mask and the resulting inner working angle, is insignificant. We compared Fig. 5.10 to an equivalent plot but without taking into account the blocking by the focal-plane mask in the computation of the contrast curve, and the differences between the figures were only visible at separations smaller than 0.01″. This means that at small separations the dominant factor is the chromaticity of the coronagraph. If spectral deconvolution is going to be the method of choice for high-contrast imagery, care has to be taken to use achromatic optics and to design 'Fresnel-friendly' instruments (Verinaud 2010). Fresnel effects, i.e. phase aberrations translating to amplitude changes, cannot be removed by spectral deconvolution.

It should be mentioned here that EPICS has an apodizer-only design, i.e. no focal-plane mask is present (Verinaud 2010). This is precisely because of the effects outlined above. Nevertheless, APLC is considered very suitable for high-contrast imaging. Fig. 5.10 shows that its performance, as quantified by the contrast curve, is almost ten times better than the classical Lyot coronagraph. This is due to the fact that an APLC is significantly less chromatic than the Lyot design. Its chromaticity can be further suppressed by using a Gaussian focal mask, instead of the hard-edge stop (Verinaud 2010).

5.5.4.3 Distance to the star

When stars appear brighter three effects are expected on the contrast curves. Firstly, the relative level of Poisson noise should decrease. This can be explained as follows: the contrast curve is related to the standard deviation of stellar noise. When we consider two stars, the first twice as bright as the second, Poisson noise in the halo of the former should increase by a factor of $\sqrt{2}$, while the peak of the PSF will increase by a factor of 2. Since the contrast curve is normalised by the peak of the un-obscured PSF it will scale as $\sqrt{2}/2$, i.e. it will be steeper by a factor of 0.7 for the brighter star.

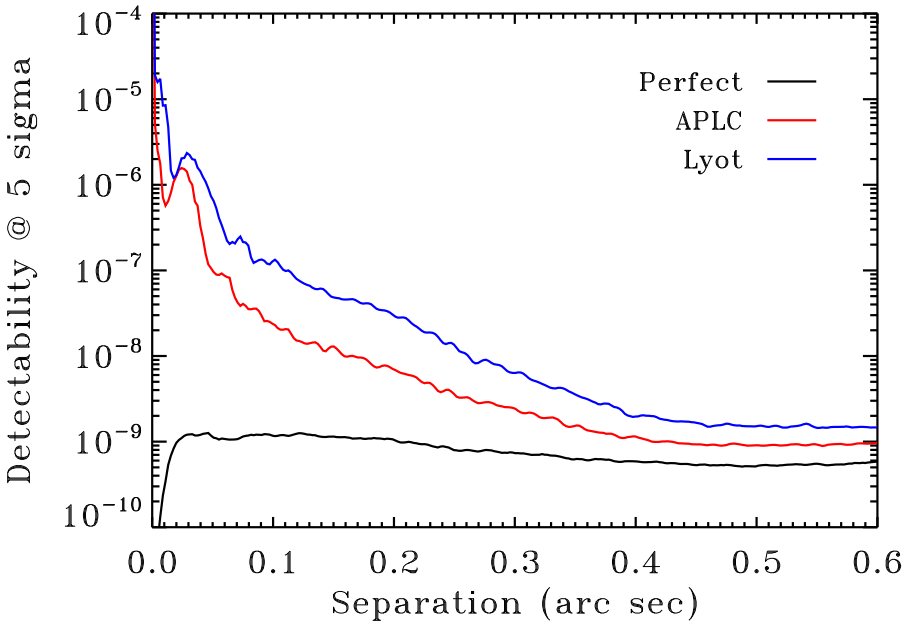

Figure 5.10: Contrast curves for the three coronagraphs considered in this project. For the 'perfect' coronagraph a throughput of 1 has been assumed. Star G2 (m_V = 4.7 mag), distance 10 pc.

Secondly, a brighter star will provide more photons for the WFS, and the AO correction will be better (assuming the AO error budget is dominated by WFS noise). The Strehl ratio will increase, i.e. the PSF will be sharper, and fewer photons will be scattered into the halo.

Thirdly, the increased Strehl ratio will concentrate more light on the coronagraphic mask. This coherent light will be removed, and the suppressed Airy residuals will generate less photon noise and fewer 'pinned' speckles – speckles which are modulated and amplified by the Airy pattern (Soummer et al. 2007).

In essence, all of the simulations corresponding to various distances (Fig. 5.11) and spectral types (Fig. 5.12) only generate the three effects mentioned above. The contrast curves will be steeper for brighter stars. To first order, when the recorded flux ratio between two stars is, say, two, the contrast curve will be steeper by a factor of 0.7.

Fig. 5.11 also shows that the contrast at small separations is again limited by the chromaticity of the coronagraph. Beyond 0.1″ the effect of Poisson noise starts to play a role and the contrast curve is shallower for fainter stars.

5.5.4.4 Star's spectral type

The contrast curves for various spectral types are shown in Fig. 5.12. As described in the previous section the contrast is better for brighter stars.

5.5.4.5 Influence of M1 aberrations

To see the effect of higher M1 aberrations on the final contrast curve we ran one simulation with the M1 phase divided by two, and another where it was multiplied by three. The resulting rms wavefront values on M1 were 0.15 μm and 0.9 μm before XAO, and 4 nm and 24 nm after correction, respectively. The wavefront phase after AO (rms = 55 nm) was dominated by post-AO aberrations, kept at the level of 30 nm, and atmospheric residuals. Fig. 5.13 shows that the effect of having very tight M1 specifications is negligible if instrumental errors remain. Both the 'perfect' coronagraph's and the APLC's contrast curves

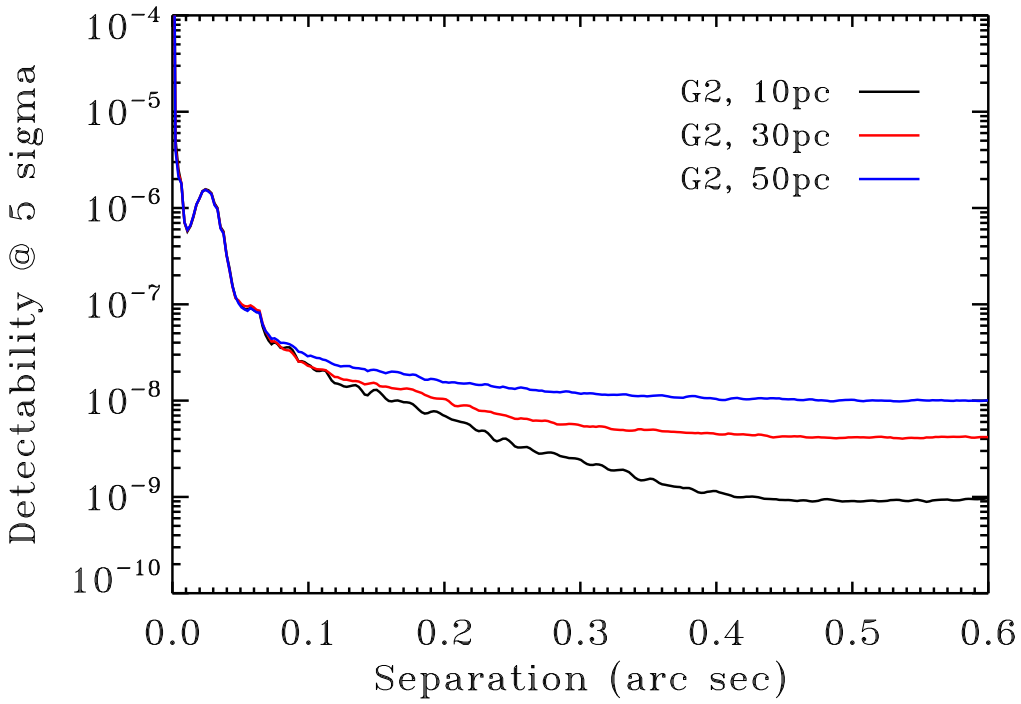

Figure 5.11: Contrast curves for stars at distances of 10, 30 and 50 pc. Stars are of type G2, APLC coronagraph was used.

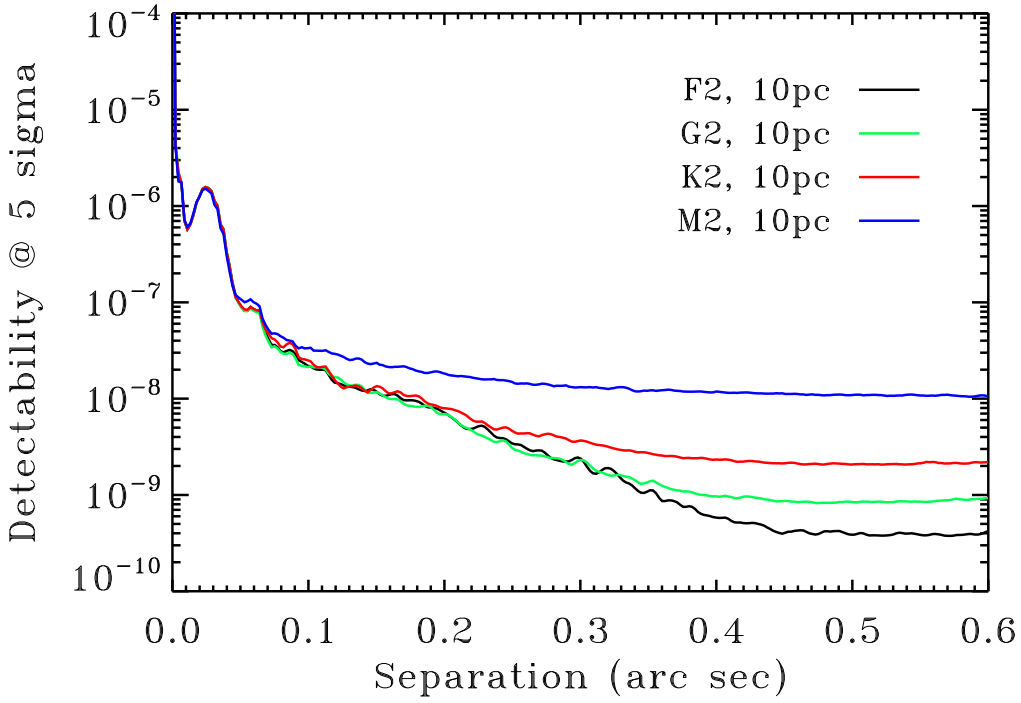

Figure 5.12: Contrast curves for F, G, K and M-type stars at a distance of 10 pc, APLC coronagraph.

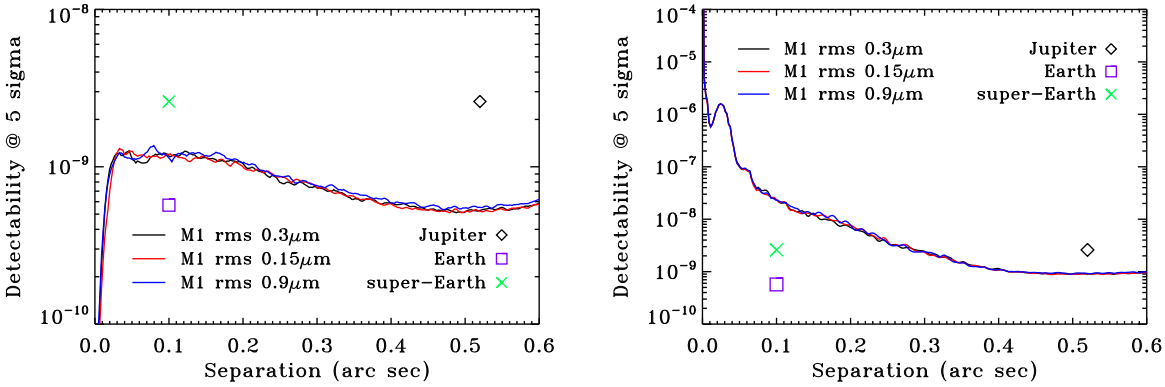

Figure 5.13: Left: Effect of varying low-spatial-frequency M1 aberrations on the contrast curve for a G-type star at a distance of 10 pc, 'perfect' coronagraph. Right: Same analysis for the APLC coronagraph.

are not susceptible to varying M1 aberrations by almost one order of magnitude. Earth-like planets are still not detectable but super-Earths could be imaged with a less chromatic coronagraph than the APLC.

5.5.4.6 Influence of post-AO aberrations

From Fig. 5.14 we can see that post-AO aberrations determine the contrast curve – even though their magnitude is small compared to pre-AO M1 errors. Clearly, it is essential to keep post-AO aberrations below 60 nm rms in order to have a chance of imaging a super-Earth. To detect an Earth-like planet these aberrations must be kept at the level of 10 nm rms.

5.6 Concluding remarks

In this report we have described the results of XAO simulations for the E-ELT. The simulations account for mirror segmentation, realistic segment aberrations, AO error budget, chromaticity of the coronagraph, and various sources of noise. The output of this extensive modelling suggests that with the combination of hardware and software described it will be possible to directly capture reflected light of super-Earth planets in the habitable zone of stars in the solar neighbourhood ($d < 10$ pc) but only with a close to perfect coronagraph. The major limitation is imposed by the chromaticity of the coronagraph chosen for realistic simulations (APLC), post-AO aberrations being the next most important factor determining the

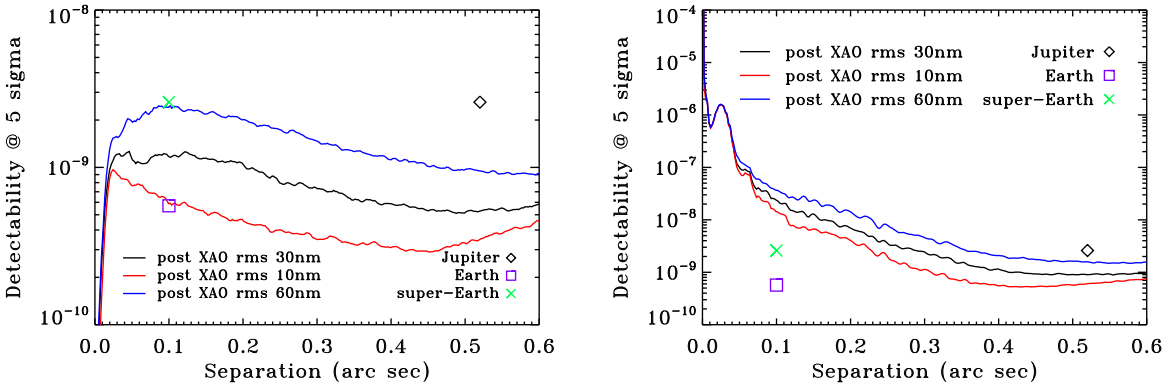

Figure 5.14: Left: Effect of varying the level of post-XAO aberrations on the contrast curve for a G-type star at a distance of 10 pc, 'perfect' coronagraph. Right: Same analysis for the APLC coronagraph.

contrast that can be achieved. This conclusion confirms the results of the simulations by the EPICS consortium (Verinaud 2010).

Imaging a true Earth twin within 10 pc from the Sun would be feasible but will require keeping the instrumental post-AO aberrations at the level of 10 nm rms, in addition to a perfect coronagraph (Fig. 5.14). The E-ELT should open up new discovery space when it comes to reflected light by old Jupiter-like planets and smaller giant gaseous planets. We have shown through our simulations that with a realistic coronagraph (APLC) and moderate integration times these planets could be detected up to distances of 30 pc if instrumental aberrations are kept at the level of around 30 nm.

Acknowledgements

The authors would like to thank Bruno Femenia Castella, Christophe Verinaud, Patrice Martinez, Arkadiusz Swat, Babak Sedghi, Laurent Jolissaint and Isabelle Surdej.

References

Antichi J., et al., 2009, ApJ, 695, 1042
Boccaletti A., Carbillet M., Fusco T., Mouillet D., et al., 2008, SPIE Proceedings, 7015, 70156E
Bonnet H., 2009, E-ELT report E-SPE-ESO-313-0202
Bonnet H., 2009, E-ELT report E-TRE-ESO-313-0387
Bonnet H., 2009, E-ELT report E-TRE-ESO-313-0637
Brandl B., et al., 2008, SPIE Proceedings, 7014, 70141N
Brown R., 2004, ApJ, 610, 1079
Cavarroc C., Boccaletti A., Baudoz P., Fusco T., Rouan D., 2006, A&A, 447, 397
Chanan G., 2009, Applied Optics, 48, 6074
Crossfield I., Troy M., 2007, Applied Optics, 46, 4533
Dimmler M., 2009, E-ELT report E-TRE-ESO-288-0211
Fusco T., et al., 2006, Optics Express, 14, 7515
Gonte F., 2009, E-ELT report ELT-TRE-ESO-04600-0099
Hinkley S., et al., 2007, ApJ, 654, 633
Jolissaint L., Veran J.-P., Conan R., 2006, Journal of the Optical Society of America A, 23, 382
Kasper M., et al., 2008, SPIE Proceedings, 7015, 70151S
Lyot B., 1939, MNRAS, 99, 580
Macintosh B., Poyneer L., Sivaramakrishnan A., Marois C., 2005, SPIE Proceedings, 5903, 170
Marois C., Lafreniere D., Doyon R., Macintosh B., Nadeau D., 2006, ApJ, 641, 556
Marois C., Macintosh B., Soummer R., Poyneer L., Bauman B., 2008, SPIE Proceedings, 7015, 70151T
Martinez P., Boccaletti A., Kasper M., Cavarroc C., Yaitskova N., Fusco T., et al., 2008, A&A, 492, 289
McGlamery B.L., 1976, in “Image Processing”, J.C. Urbach, ed., SPIE Proceedings, 74, 225
Müller M., 2008, E-ELT report E-TRE-ESO-303-0273
Piatrou P., Chanan G., 2010, Applied Optics, submitted
Poyneer L., Macintosh B., 2004, Journal of the Optical Society of America A, 21, 810
Poyneer L., Macintosh B., 2006, Optics Express, 14, 7499
Racine R., Walker G., Nadeau D., Doyon R., Marois C., 1999, PASP, 111, 587
Ragazzoni R., 1996, Journal of Modern Optics, 43, 289
Selsis F., Kasting J.F., Levrard B., Paillet J., Ribas I., Delfosse X., 2007, A&A, 476, 1373
Soummer R., 2005, ApJ, 618, L161
Soummer R., Ferrari A., Aime C., Jolissaint L., 2007, ApJ, 669, 642
Sparks W., Ford H., 2002, ApJ, 578, 543
Tecza M., Thatte N., Clarke F., Freeman D., 2009, Science with the VLT in the ELT Era, 267
Thatte N., Abuter R., Tecza M., Nielsen E.L., Clarke F.J., Close L.M., 2007, MNRAS, 378, 1229
Troy M., Crossfield I., Chanan G., Dumont P., Green J., et al., 2006, SPIE Proceedings, 6272, 62722C
Verinaud C., Le Louarn M., Korkiakoski V., Carbillet M., 2005, MNRAS, 357, L26
Verinaud C., 2010, EPICS report E-TRE-LAO-556-0002

6 S3-2: Earth twins in the habitable zone of solar-type stars

Authors: D. Naef, X. Dumusque, S. Udry, C. Lovis

6.1 The science case

Five years of sub-m s^{-1} radial-velocity measurements of quiet stars in the HARPS-GTO planet-search programme have unveiled the tip of a large population of Neptune and super-Earth-mass planets present around $\sim$30 % of G and K dwarfs of the solar neighbourhood, within 0.3 AU from the central star. These findings support recent results of synthetic planet-population models that also predict the existence of a large population of Earth-mass planets at all separations. We propose in this ambitious DRM case to directly probe the presence of Earth twins in the habitable zone in a sample of 20 close-by solar-type stars, through radial velocities obtained with a high-resolution optical spectrograph on the E-ELT. A high-resolution spectrograph with long-term stability coupled with the large collecting area of the E-ELT provide an unequalled facility for measuring stellar radial velocities at the few cm s^{-1} level. Simulations show that stellar noise (p-mode, granulation, activity) can be averaged down to this level for the quietest dwarf stars. The scientific goals of the proposal are (i) to detect Earth-mass planets in the habitable zone around solar-type stars and thus build up a list of suitable targets for future space missions aiming at characterising their atmosphere; (ii) to determine the frequency of Earth twins around neighbouring stars; (iii) to derive statistical properties of low-mass planets, priceless constraints for planet-formation models; and (iv) to characterise the multi-planet aspect of systems with Earth twins.

The observing strategy suggested in this DRM proposal is the following. First, a suitable target sample has to be selected: a set of 20 bright ($m_V \leq 10$ mag) inactive and slowly rotating stars. The following observing cadence has to be applied: 3 measurements of 15 minutes per night (separated by $\simeq$2 hours) for a given star repeated every night during a 10 night run. The same is then repeated every observing run and the run frequency is set to 1 run per month. All velocities acquired during a run are averaged in order to get one velocity average per month. Assuming that the target can be observed during 7 months every year, we end up with 7 velocity epochs per year and per star. Finally, everything has to repeated at least during one other observing season. The total telescope time required for such a strategy was estimated to be 1400 hours.

6.2 Goals of the DRM simulations

The radial-velocity semi-amplitude of the orbit of a 1 $M_\oplus$ planet with a period of 1 yr (in other words, an Earth twin) around a 1 $M_\odot$ star is $\simeq$ 9 cm s^{-1} (see Fig. 6.1 and equation 6.3). Assuming an appropriate spectrograph (see the CODEX concept study) we expect to reach a radial-velocity precision (instrumental and photon noise error) well below 4 cm s^{-1} (see Section 6.4.1.2). If we only had to deal with these two error sources, then the detection of an Earth twin would be quite easy and no detailed simulation would be required. But this ideal case is far from the reality. Other error sources are present. Stars themselves are not stable and they are able to generate their own radial-velocity signals in the absence of perturbing companions. There are at least three stellar intrinsic phenomena that are able to produce radial-velocity signals that can mimic or hide planetary signatures: stellar granulation, stellar oscillations (mostly pressure driven modes or p-modes in the case of Solar-type stars) and stellar activity. In this study, we have modelled the radial-velocity effects induced by these phenomena and we have generated realistic data sets containing all known error sources. We have also tested various observing strategies aiming at averaging out the perturbing signals and boosting the detection efficiency. The main goals of our simulation effort are:

- to realistically model the intrinsic stellar signals (granulation, oscillations and activity);
- to establish a robust planet detection criterion;
- to simulate the observing process in a realistic way and to determine detection limits;

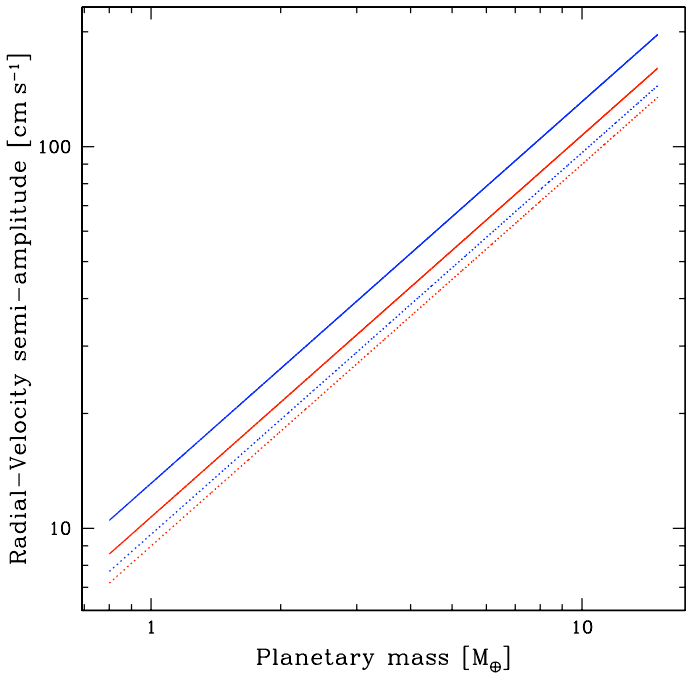

Figure 6.1: Signatures of planets in the habitable zone. This plot shows the radial-velocity semi-amplitude of a circular orbit as a function of the planetary mass. The blue curves are of a 1 $M_\odot$ (G2V) and the red ones for a 0.9 $M_\odot$ (K1V). The dashed curves are for an orbital period of one year. The blue solid curve is for a period of 216 d which corresponds to the inner edge of the habitable zone around a G2V star. The red solid curve is for the inner edge of the habitable zone around a K1V star (P = 145 d).

- to find an appropriate and efficient observing strategy;
- to estimate the smallest radial-velocity signal detectable;
- to check if the main scientific goals of the DRM proposal can be achieved and if not, how far can we expect to be from them.

In the DRM proposal, an observing strategy had been proposed (see the end of Section 6.1). This had been determined without including the effect of stellar activity. Results of simulations including activity have since shown that this strategy is no longer appropriate for active stars (even for the only marginally active ones). This original strategy remains valid only in the case of stars without spots or plages.

6.3 Metrics / figures of merit

We will build radial-velocity detection limits. These limits are curves in the planet-mass versus planet orbital period plane showing the locus where the planet detection rate reaches a selected value. These limits will take all known error sources into account, and they depend on various input parameters such as the primary star characteristics and the observing strategy. Showing with these detection limits that an Earth twin can be detected with a decent probability in the habitable zone will be our figure of merit. In order to avoid to include too many figures in this report, we will mostly focus on the limits obtained for the optimal observing strategy we have identified, and for typical targets of the considered stellar sample.

6.4 DRM simulations

6.4.1 Methodology

The significance of the detection of a radial-velocity orbit can be estimated using the following signal-to-noise ratio (SNR) relation:

$$\frac{\sigma_{planet}}{\sigma_{noise}} \geq T \tag{6.1}$$

where $\sigma_{\rm planet}$ is the velocity rms of the signal induced by the presence of the planetary companion, $\sigma_{\rm noise}$ is the rms of the noise present in the radial-velocity data and T is a threshold (the threshold we adopted is presented in Section 6.4.1.5). If the planetary orbit is not too eccentric and if it is well sampled by the data, then $\sigma_{\rm planet}$ can be related to the radial-velocity semi-amplitude of the orbit K: $K \simeq \sqrt{2}\,\sigma_{\rm planet}$. The SNR relation becomes:

$$\frac{K}{\sigma_{noise}} \geq \sqrt{2}\,T \tag{6.2}$$

The radial-velocity semi-amplitude of an orbit can be directly computed using Kepler's law. The following relation gives the radial-velocity semi-amplitude K (in m s^{-1}) of the signal induced on a primary star of mass M_1 by a planet with a minimum mass $m_2 \sin i$ (in M$_\oplus$) on a circular orbit of period P (in years):

$$K \approx 0.09\ M_1^{-2/3}\ \ m_2 \sin i\ \ P^{-1/3} \tag{6.3}$$

The noise term in equation (6.1), $\sigma_{\rm noise}$, contains various contributions which we will describe in Section 6.4.1.1. Some of these contributions are easy to quantify or safe assumptions can be made about them. Other contributions are related to complex phenomena at the surface of the star and are quite difficult to estimate.

In order to generate the selected figures of merit (see Section 6.3), we have adopted the following methodology:

- We generate a noise model containing all known error sources. This noise model depends on the characteristics of the primary star (mass or spectral type, activity level, apparent magnitude).
- Monte-Carlo simulations:
  - Selection of the primary star characteristics (spectral type, activity level).
  - Definition of a grid of planetary characteristics: a grid in the mass versus orbital period plane (see Dumusque et al. 2009, for details about the grid).
  - Selection of the observing strategy (epochs of observations, number of integrations per epoch, time gap between the integrations, integration time). The output of this step is a set of j Julian dates where j equals the number of epochs times the number of integrations per epoch.
  - Generating the planetary signal: for each grid point, we generate N radial-velocity data sets. These sets consist of j velocities (one per Julian date). The planetary signal is of course based on the period and mass of the considered grid point and on a small set of Monte-Carlo variables that are drawn randomly from statistical distributions (we will list all these variables in Section 6.4.1.6). Note that the orbital inclination i was set to 90° (i.e. edge-on orbits) in all simulations.
  - Generating the noise signal: as for the planetary signal, we generate N velocity noise data sets each of them containing j velocities. All of these velocities are generated from the noise model corresponding to the selected stellar characteristics. The noise signal also depends on a Monte-Carlo variable, a phase term drawn from a uniform distribution (see Section 6.4.1.6). Note that the noise signal is integrated over the selected DIT as it would happen for a real integration.
- Selection of a data binning scheme. The scheme may depend on the orbital period but the binning length (in days) needs to be always much shorter than the orbital period in order not to average out or distort the planetary signal.

- Binning: the planetary and noise data sets are binned following the selected scheme. Note that different binning schemes can be tested on the outputs of the same Monte-Carlo simulation.
- Statistics: for each of the N binned planetary and noise data sets, we compute the two terms of the SNR equation: $\sigma_{\rm planet}$ and $\sigma_{\rm noise}$.
- Determining the detection limit: this curve is obtained using the adopted detection criterion. We will give more details about this step in Section 6.4.1.5.

We can use these simulations to predict the radial-velocity noise rms as a function of the target characteristics, observing strategy, integration time and binning scheme. An example of the detection limits obtained in the way we just described is shown in Fig. 6.14.

6.4.1.1 Noise model

6.4.1.2 Instrumental error and photon noise

In our study, we have set the instrumental radial-velocity long-term precision to 2 cm s^{-1}. This value corresponds to the CODEX instrument specification and is assumed to include all instrumental error sources one can currently think of (uncertainties on the wavelength calibration, photon-noise of the reference spectrum, guiding errors, etc.). We stress that we assume this to be the long-term precision. In our case, long term means 4–5 years. As it observed for many other radial-velocity instruments, the precision over shorter timescales (e.g. over one night or a few days) may well be better than this value but we have assumed in this study that the 2 cm s^{-1} value was the same at all timescales.

The SNR of a spectrum can be estimated with the E-ELT Spectroscopic Exposure Time Calculator (ETC, see [RD1]). This estimate can be used to derive the photon-noise radial-velocity uncertainty $\epsilon_{\rm RV}$. It can be shown (either theoretically or with numerical simulations) that the radial-velocity photon-noise error

Table 6.1: Limiting magnitudes for obtaining radial-velocity photon-noise errors of 1, 2 and 5 m s^{-1} in 15 min for two different spectral types as a function of telescope diameter, site and mirror coating. Al and Ag/Al stand for the aluminium and silver/aluminium coatings. PA and H&D stand for the Paranal-like and High & Dry sites. These limits were estimated using the E-ELT Spectroscopic ETC (version 2.14). We have set the resolving power to 100 000, the seeing to 0.8″ and the airmass to 1.15. The last three lines show the magnitudes at which the detector linearity limit is reached in 15 min.

Spectral type	G2V				K1V			
Coating	Al		Ag/Al		Al		Ag/Al	
Site	PA	H&D	PA	H&D	PA	H&D	PA	H&D
Diameter [m]	Photon noise error of 1 cm s^{-1} in 900 s							
30	6.53	6.58	6.82	6.87	8.09	8.14	8.38	8.43
42	7.26	7.31	7.55	7.60	8.82	8.87	9.11	9.16
60	8.03	8.08	8.32	8.37	9.59	9.64	9.88	9.93
	Photon noise error of 2 cm s^{-1} in 900 s							
30	8.04	8.09	8.32	8.37	9.59	9.64	9.88	9.93
42	8.77	8.82	9.05	9.10	10.32	10.37	10.61	10.67
60	9.54	9.59	9.82	9.87	11.09	11.14	11.38	11.43
	Photon noise error of 5 cm s^{-1} in 900 s							
30	10.03	10.08	10.31	10.36	11.58	11.63	11.87	11.92
42	10.76	10.81	11.04	11.09	12.31	12.36	12.60	12.65
60	11.53	11.58	11.81	11.86	13.08	13.13	13.37	13.42
	Linearity limit in 900 s							
30	8.58	8.63	8.86	8.91	8.63	8.68	8.92	8.97
42	9.31	9.36	9.59	9.64	9.36	9.41	9.65	9.70
60	10.08	10.13	10.36	10.41	10.13	10.18	10.42	10.47

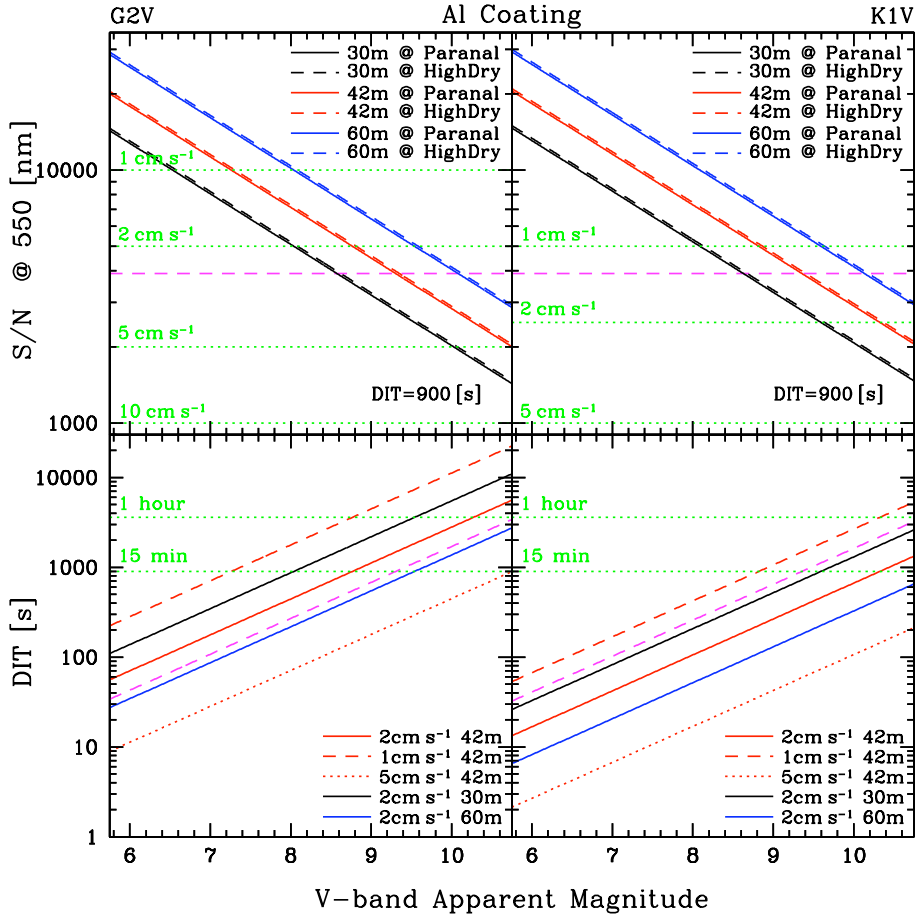

Figure 6.2: Top panels: SNR at 550 nm as a function of the apparent visual magnitude for a G2V star (left) and a K1V star (right) for the case of the aluminium coating. Various primary mirror sizes and two sites have been considered: 30 m (black), 42 m (red) and 60 m (blue); Paranal-like (solid curves) and High & Dry (dashed). These curves have been obtained using the E-ELT Spectroscopic ETC (version 2.14) and the following input parameters: 0.8″ seeing, 1.15 airmass, 100 000 resolution, 900 s DIT. The horizontal green dotted lines show corresponding radial-velocity photon-noise errors. Bottom panels: Radial-velocity photon-noise curves in the DIT versus apparent visual magnitude plane for a G2V star (left) and K1V star (right) for the case of the aluminium coating and a Paranal-like site. In all panels, the magenta dashed lines show the expected detector linearity limit (multiple integrations are required above this limit).

$\epsilon_{\rm RV}$ mostly depends on the SNR and on the mean contrast C of the spectral line used for the velocity computation. The relationship between these quantities is:

$$\epsilon_{RV} = \frac{\alpha}{C \times SNR} \tag{6.4}$$

where α is a constant that depends on the instrument spectral format and that needs to be calibrated. The SNR is per resolution element. Note that $\epsilon_{\rm RV}$ also depends on the mean width of the spectral lines (the error increases when the width increases) but we have neglected this term as all the surveyed targets will be very slow rotators. The line width effect will thus be nearly constant and it is included in the constant α. For α, we will assume that the calibration obtained for HARPS will also be valid for a high-resolution spectrograph at the E-ELT: a photon-noise error of 1 m s^{-1} is obtained for Sun-like star (i.e. a G2V star with Solar metallicity) with SNR = 100 (i.e. 1 cm s^{-1} is obtained with SNR = 10 000). This calibration was made for the SNR at 550 nm. For K dwarfs, we have assumed that the mean line contrast was about twice bigger than for G dwarfs (i.e. for a given SNR, the photon-noise error on a K dwarf is half that of a G dwarf).

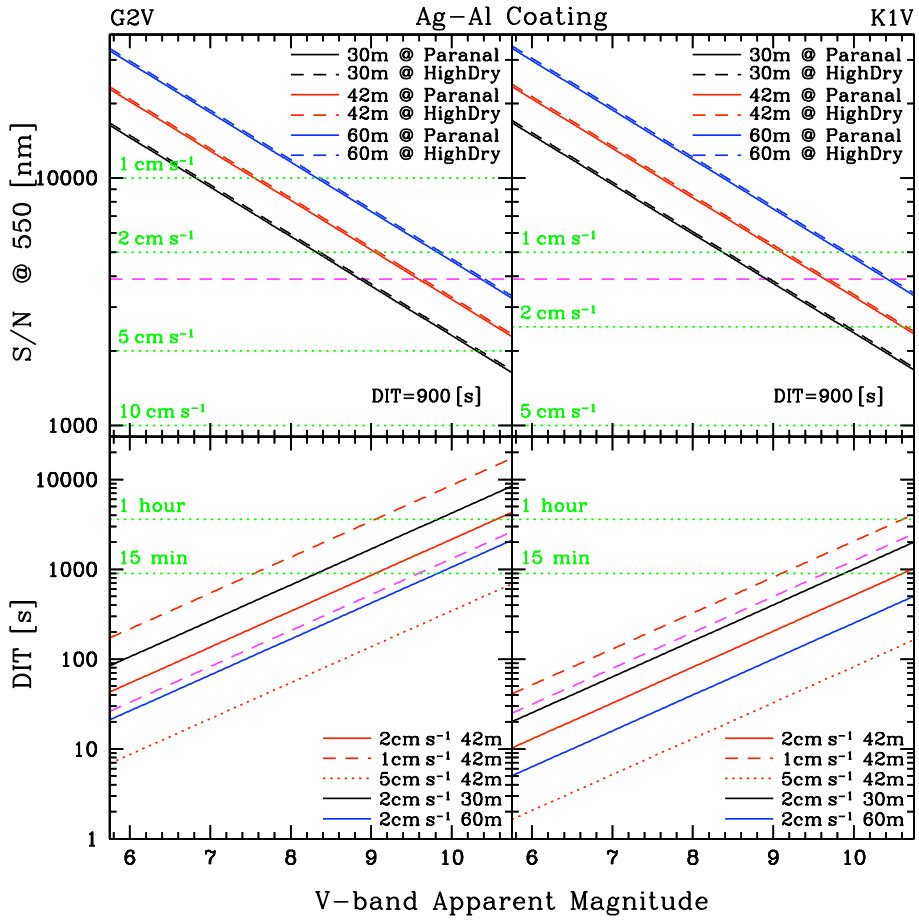

Figure 6.3: The same as Fig. 6.2 but for the protected silver/aluminium coating.

We have used version 2.14 of the ETC for obtaining the SNR at 550 nm for two reference targets (a G2V star and a K1V star). The following assumptions and input parameters were used: apparent magnitude $m_V = 10$ mag, DIT = 15 min, point source, seeing-limited case, seeing = 0.8″, airmass = 1.15, fibre diameter = 0.82″ (hence SNR was computed over an area of radius 0.4″), a spectral resolution of 100 000 and a sampling of 4 pixels per resolution element (which implies a total number of spectra of 280 on the detector). We have also varied the telescope main mirror diameter (30, 42 and 60 m), the mirror coating (aluminium and protected silver/aluminium) and the observing site (Paranal-like and High & Dry).

We show the results in Fig. 6.2 for the aluminium coating and in Fig. 6.3 for the silver/aluminium coating. For the two cases, we show SNR versus target magnitude curves for a fixed 15 min exposure time (upper panels). We also plot the exposure times required for achieving a fixed photon-noise error as a function of the target magnitude (lower panels). In Table 6.1, we list limiting magnitudes for various photon-noise errors.

Our conclusions on the photon-noise are:

- the impact of the site is very small (0.05 mag effect);
- the mirror coating only has a minor impact (0.28 mag effect);
- a photon-noise error better than 5 cm s^{-1} can be achieved for a G2V star brighter than 10.8 mag (12.3 mag for K5V star) in less than 15 min, assuming a 42 m telescope with the aluminium coating on a Paranal-like site;

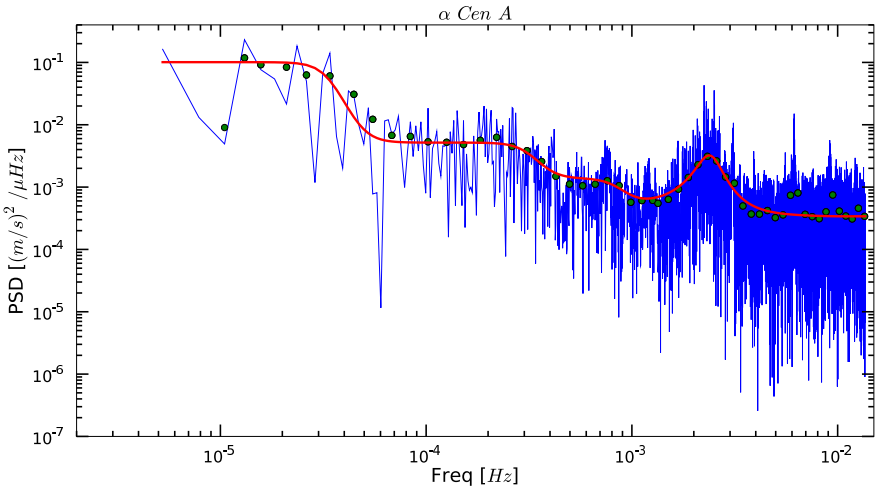

Figure 6.4: Radial-velocity power spectrum density (PSD) for the G2V star α Cen A (HD 128620). The blue solid line shows the PSD computed from a HARPS asteroseismologic data set. The green dots correspond to the binned spectrum and the red solid line is the fitted model (see Section 6.4.1.3). From Dumusque et al. (2009).

- a photon-noise error better than 2 cm s^{-1} can be achieved for a G2V star brighter than 8.8 mag (10.3 mag for K5V star) in less than 15 min, assuming a 42 m telescope with the aluminium coating on a Paranal-like site.

As we will see later, the impact of photon-noise on the velocity error budget is almost negligible in the cases we have considered: target brighter than $m_V = 10$ mag and exposure times larger than 5 min (see Figs. 6.11 and 6.12).

6.4.1.3 Stellar oscillations and granulation induced noise

In this section, we present our effort aimed at modelling the radial-velocity signals induced by physical phenomena taking place in stars: stellar oscillations and granulation. These two effects originate in the interiors of stars but both induce effects at the stellar surface. These surface effects produce the radial-velocity signatures we will model.

Stellar oscillations result from the propagation of pressure waves ("p-modes") from the interior of the star to its surface inducing dilatations and contractions of its outer layers. The typical timescale of these non-radial oscillations is a few minutes (5 minutes for the Sun, see Schrijver & Zwaan 2000). Depending on the star type and its evolutionary stage, the amplitude of these modes varies from 10 to 400 m s^{-1} (Schrijver & Zwaan 2000). The current observational knowledge shows that the amplitude and the pulsation period increase with the stellar mass on the main sequence (Mayor et al. 2003; Arentoft et al. 2008). From the theoretical point of view, it is predicted that the frequencies of the p-modes increase with the square root of the mean stellar density and that the amplitudes are proportional to the ratio of the luminosity over the mass (Christensen-Dalsgaard 2004). Unevolved low-mass stars thus seem to be the most suitable targets for a very low-mass planet search.

The various types of granulation originate in the convective energy transport present in cool stars. Three types of granulation are considered: granulation, mesogranulation and supergranulation. These types differ in their typical timescales: from several minutes to several hours (up to 33 hours; Del Moro et al. 2004). The typical amplitudes of the granulation signatures are similar to those produced by p-modes (see e.g. Schrijver & Zwaan 2000; Kjeldsen et al. 2005).

The method we have used for modelling the noise induced by these phenomena is described in detail by Dumusque et al. (2009). We briefly summarise the main steps here. We started with real radial-velocity

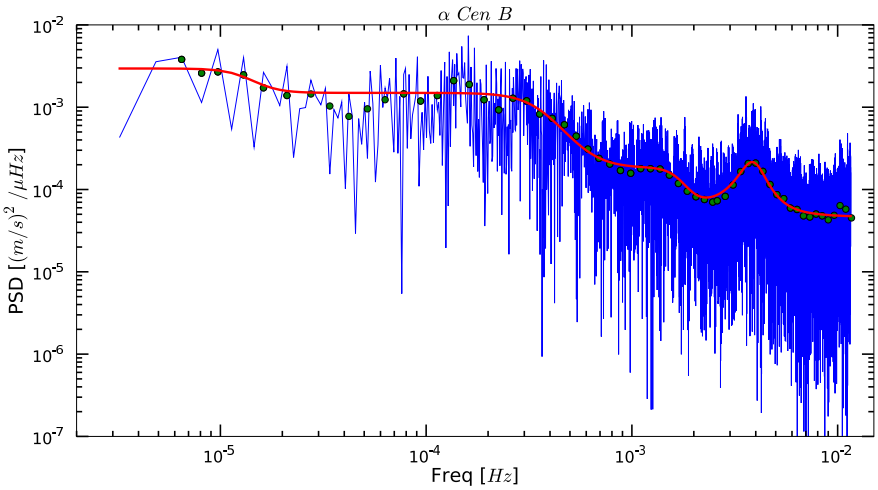

Figure 6.5: The same as Fig. 6.4 but for the K1V star α Cen B (HD 128621). From Dumusque et al. (2009).

data sets obtained with HARPS for our two reference targets: α Cen A and B (HD 128620 and HD 128621, see Section 6.4.3.1). The targets' spectral types are G2V (A) and K1V (B) and both stars have very low levels of activity. The two data sets consist of long, uninterrupted series of short integration time (the two stars are very bright) observations over several consecutive nights. Their sampling are thus appropriate for characterising the noise signatures over the typical timescales covered by both granulation and p-modes. We have computed the radial-velocity power spectrum densities (PSD) for the two targets. The observations are not evenly spaced so standard Fast Fourier Transforms algorithms could not be used here. Instead, we have computed weighted Lomb-Scargle periodograms (Lomb 1976; Scargle 1982; Zechmeister & Kürster 2009). PSDs are then obtained by normalising the periodograms by the inverse of the area under the spectral window (Kjeldsen et al. 2005). With such a normalisation, PSDs no longer depend on the observing window. The obtained PSDs are shown in Fig. 6.4 for α Cen A and in Fig. 6.5 for α Cen B (blue solid curves). Three main regimes can be distinguished on the two plots. Below 1 mHz, the power decreases with increasing frequency. This low-frequency part is essentially due to meso and supergranulation. Then, between 0.001 and 0.01 Hz, a bump is present. This bump is essentially due to p-modes. In this frequency range, (corresponding to 100–1000 s) granulation also contributes to the power spectrum. At higher frequencies, both PSDs become more or less flat (i.e. white noise). This latter part of the spectra at high frequency results from other noise sources: instrumental error, photon-noise, guiding errors, etc. Note that the white noise part can be subtracted if one wants to study stellar intrinsic noise alone. One can also modify the level of this flat part in order, e.g., to artificially modify the contribution of photon-noise.

The next step is to fit more physical models to the obtained PSDs. Before the fit, the PSDs have been smoothed. The region of the p-mode bump is convolved with a Gaussian function with a full width half maximum of $4\Delta\nu$ where $\Delta\nu$ is the p-mode large frequency separation (Kjeldsen et al. 2005). The $\Delta\nu$ values we have used as well as the paper they come from are listed in Section 6.4.3.1. In the other frequency regions, the PSDs have been binned. The results of theses smoothing processes are illustrated by the green dots in Figs. 6.4 and 6.5.

The model used for fitting the granulation part was first proposed by Harvey (1984) and later reviewed by Andersen et al. (1994). The empirical models they proposed are exponentially decaying time functions (one per granulation type). In frequency space, these "Harvey" functions take the following form:

$$P(\nu) = \frac{A}{1 + (B\nu)^C} \tag{6.5}$$

where $P(\nu)$ is the velocity PSD, A the power density of the corresponding convection motion, B its characteristic time scale and C the slope of the power law. The power remains approximately constant on timescales larger than B and drops off for shorter timescales. The total velocity PSD due to granulation

phenomena is the sum of the PSD for each phenomenon (G = granulation, MG = mesogranulation and SG = supergranulaion):

$$P(\nu) = P_{\rm G}(\nu) + P_{\rm MG}(\nu) + P_{\rm SG}(\nu) \quad (6.6)$$

Finally, we fitted the p-mode bump with a Lorentzian function (for details, see e.g. Lefebvre et al. 2008) and the high frequency part with a constant. The final fits are also shown in Figs. 6.4 and 6.5 with red solid curves.

Synthetic noise data sets can be obtained from the fitted PSD. Such data sets will reproduce the noise induced by both granulation and oscillations in a realistic way. In order to get a noise value at the desired epoch t_i, we have used the following equation which is justified by the fact that the Lomb-Scargle periodogram is equivalent to a weighted sine wave fit (Scargle 1982; Zechmeister & Kürster 2009):

$$RV(t_i) = \sum_{\nu} \sqrt{PSD(\nu)}\ \sin(2\pi\nu t_i + \phi(\nu)). \quad (6.7)$$

In this equation, the frequency ν spans from $2/T$ (T is total time span of the data set used for computing the PSD) and the Nyquist frequency with a step of $1/T$. The phase $\phi(\nu)$ is drawn from a uniform distribution between 0 and 2π. We do so assuming that all noises due to stellar phenomena are independent. $\phi(\nu)$ is one of the Monte-Carlo variables (see Section 6.4.1.6).

In order to test the validity of this approach, we have generated synthetic data sets with the same observing dates as the one of the input HARPS radial-velocity data. We have then computed the rms of these synthetic data and compared it to the rms of the real data. For both stars, the discrepancy between the two rms values was less than 5%.

6.4.1.4 Stellar activity induced noise

The presence of cool and dark spots on the surface of a rotating star generates spectral lines distortions by altering the symmetry of the stellar disk illumination. These distortions change with time as rotation goes on. In addition, spots appear and disappear. The number of spots or groups of spots, their position (longitude, latitude) as well as their aspects (e.g. the filling factor) depend on the activity level.

In this section, we briefly describe the method we have used for simulating the radial-velocity signals induced by the presence of spots. The detailed description of the method can be found in Dumusque et al. (2010). The basic idea is to generate realistic spot configurations. These configurations mostly depend on the activity level which is characterised by the index $\log R^{'}_{\rm HK}$ (see e.g. Noyes et al. 1984). The time evolution of the initial spot configuration is then simulated (using time steps of 2 h) again in as realistic a way as possible (for example by taking differential rotation into account). The output of this step (positions, numbers and sizes of spot groups at every time step) are then plugged into the programme SOAP (Huélamo et al. 2008, or Bonfils & Santos 2010, in prep.). This programme is used for computing the radial-velocity signal produced by a given spot configuration. For every activity level we considered, we generated 100 simulations (see details about the Monte-Carlo variables in Sect 6.4.1.6).

In these simulations, we have used many observational constraints obtained for the Sun. For example, we have constraints on the latitudes of spot groups and the number of spots as a function of the activity level, on the statistics of spot lifetimes, on the relation between the maximal spot group filling factor (i.e. the ratio between the spot group area and the area of the visible hemisphere) and on the statistical law governing the probability of appearance of a spot group. For more details about these constraints, we refer the reader to Dumusque et al. (2010) and references therein.

The level of activity considered in our simulations are $\log R^{'}_{\rm HK}$ = −5.0, −4.9 and −4.75, corresponding to the Sun with no, average and maximum activity, respectively. Note that an activity index of −5.0 refers to a spotless star. The lack of observational constraints (e.g. direct spot maps and statistics) did not allow us to simulate stars with activity levels higher than −4.75. This is not really a problem as the search for very low-mass planets will be focused on stars as inactive as possible. Also, we lack observational constraints on stars with spectral type different from that of the Sun. Results on the activity effects on cooler stars (e.g. the K1V star α Cen B) have been obtained under the strong assumption that the spot group configurations and evolution for a given level of activity were similar for G2V and K1V stars.

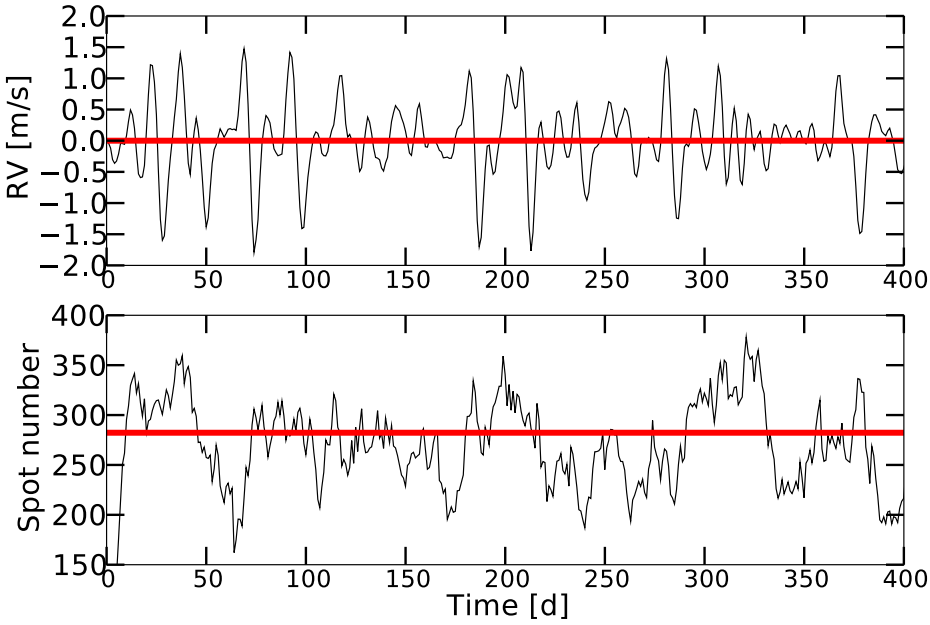

Figure 6.6: Simulations of stellar activity. Top: Stellar dark spots induced radial-velocity effect as a function of time. Bottom: Total number of spots present on the stellar surface as a function of time. These results have been obtained for the case of a star with a $\log R'_{\rm HK}$ activity index of -4.75 and exhibiting 6 active longitudes (3 per hemisphere). From Dumusque et al. (2010).

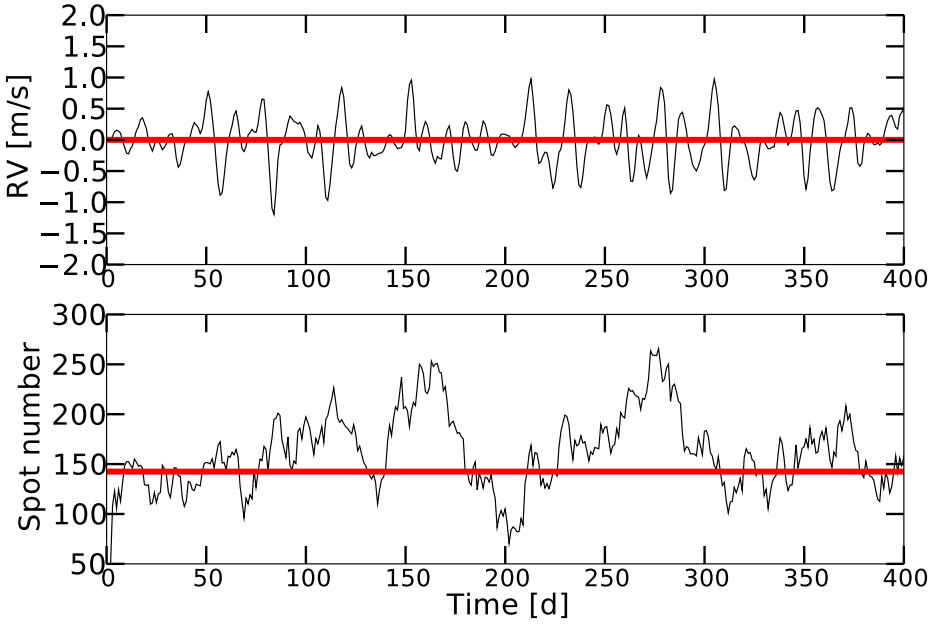

Figure 6.7: Same as for Fig. 6.6 but for a very low-activity star (activity index $\log R'_{\rm HK} = -4.9$). From Dumusque et al. (2010).

We show examples of results obtained with this simulation scheme for two different activity levels in Figs. 6.6 ($\log R^{'}_{\rm HK} = -4.75$, i.e. the Sun at its maximum activity) and 6.7 ($\log R^{'}_{\rm HK} = -4.9$, i.e. the Sun at average activity). In these two figures, we show the activity induced radial-velocity effect as a function of time (upper panels) and the total number of spots as a function of time (lower panels).

As for the case of pulsation and granulation, we have compared the results of our activity simulations with real Solar radial-velocity data presented in Lagrange et al. (2010). In particular, we have compared their measured velocity rms, 59 cm s^{-1}, with the one obtained from our simulations: 63 cm s^{-1}. The agreement between the two quantities is very good.

6.4.1.5 Detection criterion

In Section 6.4.1 we already introduced the SNR relation we will use (see equation 6.1). In this section, we present the threshold T we have selected and we show that this is a robust detection criterion. A planetary signal will be considered as detected if:

$$\frac{\sigma_{planet}}{\sigma_{noise}} \geq 1.8 \tag{6.8}$$

As mentioned in Section 6.4.1, we can relate $\sigma_{\rm planet}$ to the radial-velocity semi-amplitude K under the assumptions that the orbit is not too eccentric and the data sampling is sufficient (if the sampling is too poor, then the velocity extrema of the orbit can be missed resulting in an under-estimation of the velocity rms). The detection criterion becomes:

$$\frac{K}{\sigma_{noise}} \geq 2.5 \tag{6.9}$$

Fig. 6.8 shows the F-test probability as a function of $K/\sigma_{\rm noise}$ for simulated data sets. Each of these data sets contained a real planetary signal and noise (with an rms of $\sigma_{\rm noise}$) generated using the noise model presented above. For each set, we fitted a Keplerian orbit which gave us the semi-amplitude K. Then, using the pre-fit and post-fit velocity dispersion, we computed the fit F-test probabilities for each set. The F-test probability allows the comparison between the two different models: the constant model (i.e. 1 free parameter: the velocity mean) and the Keplerian model (6 free parameters). A high probability corresponds to a high significance of the 6-parameter Keplerian model. In other words, a high F-test probability corresponds to a low false-alarm probability and to a high significance of the fit. In Fig. 6.8, we see that $K/\sigma_{\rm noise} \geq 2.5$ corresponds to an F-test probability of $\leq$ 0.3% (i.e. 3σ).

For the determination of the detection limits from the simulation scheme presented in Section 6.4.1, we have used the detection threshold in the following way. For every point in the mass versus period grid, we have in hand N simulated data sets. For each of them, we compute the $\sigma_{\rm planet}/\sigma_{\rm noise}$ ratio and we count the number of sets for which this ratio exceeds the selected threshold (i.e. 1.8). For every period, we then search the minimal mass for which the fraction of ratios above 1.8 is above 80%. This minimal mass is then set as the detection limit for that period. The detection limits we obtain in this way are thus 80% detection limits.

6.4.1.6 Monte-Carlo variables

In this section we list the Monte-Carlo variables we have used in our simulations together with the statistical distributions we used for generating random values for these variables.

First, here are the variables used when generating the planetary signals:

T_0 = instant of periastron passage of a planet, uniform distribution between 0 and the orbital period.

ω = longitude of periastron of a planet, uniform distribution between 0 and 2π. Note that ω is set to 0 if the orbit is circular. In this case T_0 corresponds to the instant of the radial-velocity maximum.

e = orbital eccentricity. We have used the observed eccentricity distribution of exoplanets. This distribution is a function of the orbital period (e.g. short-period orbits are assumed to be circularised). For the definition of this distribution, see Dumusque et al. (2009) or Udry & Santos (2007).

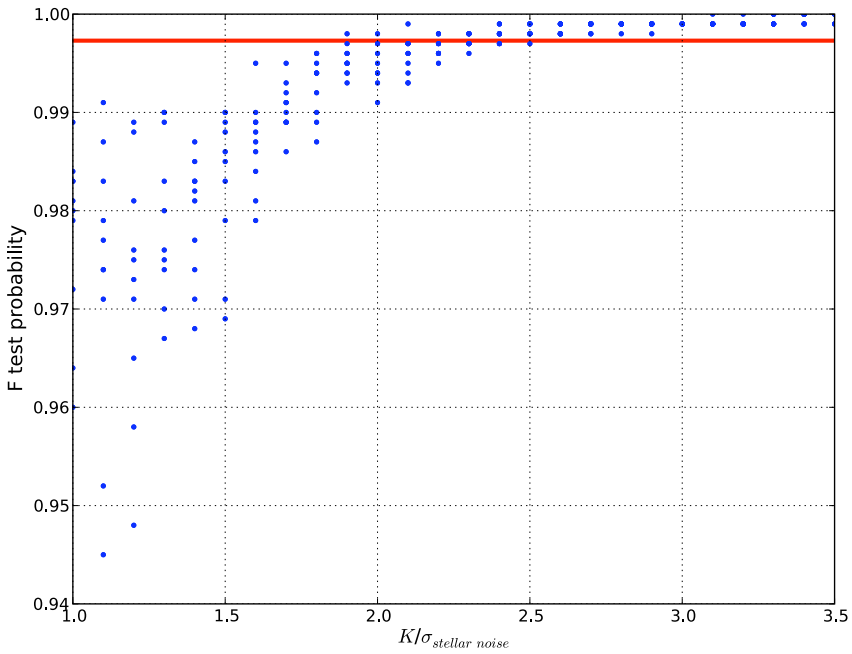

Figure 6.8: Testing the adopted detection criterion. F-test probabilities (i.e. 1−false-alarm-probability) of fitted orbits to simulated data versus K/σ (radial-velocity semi-amplitude of the orbit over rms of stellar noise). We see that for the adopted criterion ($K/\sigma \geq 2.5$), the false-alarm-probability is well below 1 %. The red horizontal line corresponds to a 3σ detection (99.7 %)

The only Monte-Carlo variable used for generating the noise signal induced by oscillations and granulation is the phase of the noise signal $\phi(\nu)$. This term is drawn from a uniform distribution between 0 and 2π.

Finally, we list the variables for generating the activity-induced noise signal. Note that these are irrelevant for simulations of stars with $\log R^{'}_{\rm HK} = -5.0$ (no activity) because these are spotless.

N_L = number of active longitudes. We used 4, 5 or 6 longitudes, with equal probability for the three cases. Note that the angle between active longitudes in a given hemisphere is not random. In the case of 2 active longitudes, the angle between them is 180°. If a third longitude is present, then it is shifted by 90 degrees with respect to the two others. Also, it seems that the positions of the active longitudes in one hemisphere are not independent from the positions of the active longitudes in the other hemisphere. It seems that a 90° shift is a likely configuration.

L = longitude of a spot group. Once the number of longitudes is selected, spot groups are placed on them. The exact longitude is obtained via a Gaussian distribution centred on the active longitude and with a 15° dispersion (matching what is observed in the case of the Sun).

T = spot group lifetime. 50 % have a lifetime between 1 hour and 2 days, 40 % between 2 and 11 days, 10 % between 11 days and 2 months.

SA = spot appearance. The probability of spot appearance is governed by a Poisson process:

$$P[(N(t+\tau) - N(t)) = k] = \frac{(\lambda\tau)^k e^{-\lambda\tau}}{k!}, \quad k = 0, 1, \ldots \tag{6.10}$$

where P is the probability, N the number of spot groups, t the time, τ the time step and λ the average rate of appearance of spots per unit of time. This value is manually adjusted so that the total number of spots matches the Solar value for the considered activity level. The appearance of k spots is tested at every time step with this equation.

Note that the spot group latitude is not a random variable. Solar observations show that the latitude of spot groups is a function of the activity level: $\pm 22.5^\circ$ for $\log R'_{\rm HK} = -4.9$ and $\pm 15^\circ$ for $\log R'_{\rm HK} = -4.75$, i.e. spot groups migrate towards the equator when activity increases. Note that the minimal distance between two spot groups was set to 1°. When drawing the longitude from the Gaussian distribution (see above), spot groups too close to others are rejected and another longitude is drawn.

6.4.2 Pipeline

As explained in Section 6.4.1, the main simulation effort of this study focuses on the generation of a realistic noise model. Once the model is ready, we select the observing strategy we want to test and we generate the corresponding radial-velocity data set containing all the signals (planetary signal if present and realistic stellar noise signal). This data set is then directly used for the analysis without the need of special processing or additional reduction. The analysis consists of some basic statistical operations: data binning (i.e. averaging) and rms computation. These operations are standard and can thus be done with any mathematical software. We have not used any specific (or generic) pipeline.

6.4.3 Inputs

6.4.3.1 Scientific data

For this study, we have used several HARPS radial-velocity data sets. Two of them have been used for building up the velocity PSDs of our two reference targets: α Cen A and B (HD 128620 and HD 128621). The α Cen A HARPS data set was retrieved from the ESO archive (from programme 075D-0800A, see Bazot et al. 2007). The one for α Cen B is not published yet and it was kindly provided to us by the PI of the programme (F. Pepe, priv. comm.). The last data set we have used is the one for the star HD 69830. In fact, we have only used the dates of observation of these latter data in order to obtain a realistic and dense enough observing sequence. These data are published in Lovis et al. (2006).

In order to model the power spectrum densities of our two reference targets, the p-mode oscillation large frequency separation $\Delta\nu$ is required (see Section 6.4.1). We have used the following values: $\Delta\nu = 106\,\mu$Hz for α Cen A (Kjeldsen et al. 2005) and $\Delta\nu = 162\,\mu$Hz for α Cen B (Kjeldsen et al. 2005; Bazot et al. 2007).

Finally, we have used the the definition of the habitable zone (in particular, the location of its lower edge) published in Selsis et al. (2007).

6.4.3.2 Technical data

The main explicit technical assumption made in our study is the long-term (up to 4–5 years for this science case) radial-velocity precision that will be achieved with a high-resolution spectrograph. We have assumed that the CODEX instrument specification will be achieved: $2\,{\rm cm\,s^{-1}}$.

There are of course some implicit technical assumptions. In particular, we have used the E-ELT Spectroscopic ETC (version 2.14, [RD1]) for estimating the photon-noise radial-velocity errors that would affect our measurements. Table 6.1 and Figs. 6.2 and 6.3 have been obtained using the ETC for the case of a seeing-limited spectrograph (i.e. no AO). In the table and plots, various technical aspects are considered (mirror size, coating, site). In the simulations, only the following case was considered: 42 m telescope, aluminium coating, Paranal-like site. However, we note that the targets considered here are bright ($m_V \leq 10$ mag), so that photon-noise errors are always negligible when compared to stellar intrinsic error sources. This is why we decided not to test all the mirror size + coating + observing site combinations.

6.4.4 Outputs

The output of the simulations simply consists of radial-velocity rms values and detection limits in the planet-mass versus orbital period plane as a function of the target star characteristics, observing strategy and binning scheme.

6.5 Results of simulations

6.5.1 Simulation runs

Several simulation runs were performed for this study. Below we list the values or range of values that were used for the various input parameters of the simulations:

Target spectral type. In this study, we have focused our efforts on two spectral types: G2V and K1V. Results on other spectral types (including one sub-giant star) can be found in Dumusque et al. (2009)

Target activity level. Three different levels have been tested in our study: $\log R^{'}_{\rm HK}$ =−5.0 (spotless Sun), −4.9 (average Solar activity) and −4.75 (maximal Solar activity).

Number of measurements per observation epoch. A measurement can in fact be a set of several consecutive integrations if the target is bright. Three different values have been tested: 1 meas./night (i.e. the current HARPS rate), 2 and 3 meas./night. The time gap between measurement was set to 5 h in the 2 meas./night case and to 2 h in the 3 meas./night case.

Number of epochs per month and time gap between epochs. Two main strategies have been tested, both with 10 epochs per month. The first strategy is the one selected in the DRM proposal, i.e. 1 epoch per night during 10 consecutive nights. The second one is the one we would recommend after this study: 1 epoch every 3 nights, hence 10 epochs per months. We have also tested an ultimate strategy: 1 epoch per night every night. Such a strategy gives excellent results but its cost in terms of telescope time is very high.

Integration time of individual measurements. We have tested various values in the range 1–20 minutes.

Binning length. The binning of the synthetic data is not a part of the simulation. This operation is a post-simulation process. All possible binning lengths can be tested on the result of a given simulation run. The binning length values we have tested range from 1 day to 10 days.

6.5.2 Analysis

The first point we want to discuss here is the effect of the data binning. Several figures presented in this report clearly demonstrate the positive impact of this data manipulation. For example, we see in Fig. 6.9 that the rms of the noise induced by oscillations and granulation decreases when the binning length increases. We also see in this plot that the improvement seems to somehow saturate and the gain above a binning length of 10 days becomes marginal. Fig. 6.9 shows the results obtained for a K1V star but the same behaviour is observed for all the target types tested so far (Dumusque et al. 2009). The positive impact of the binning remains true when stellar activity induced noise is added. This can easily be seen in Fig. 6.10 where decreasing jumps are clearly visible each time the binning length is increased. Again, the effect does not depend on the observing strategy as it can again be seen in Fig. 6.10 (e.g. both the red dashed and the red solid curves exhibit these jumps).

The second point is related to the DIT (i.e. the integration time of individual exposures). This parameter plays only a minor role if and only if more than one exposure is taken every observing night. In this case, integrating for more than 5 min marginally improves the detection efficiency. This can be clearly seen in Figs. 6.11 (for a G2V star) and 6.12 (for a K1V star). In these plots we see that the various black curves become more vertical for DITs above 5 min. This result is valid for all spectral types. This result is also valid for all levels of activity which is not a big surprise since the typical timescale of the activity signal, namely the star rotation period, is much longer (typically weeks) than any considered DIT (up to 20 min). The same effect is also observed when a different strategy is used (see Fig. 6.13). Again, we stress the fact that reducing the integration time down to 5 min is only fine when more than one exposure is taken each observing night. Of course, this is no longer a valid approach for faint targets as photon noise needs to be kept low (see the coloured curves in Figs. 6.11 and 6.12).

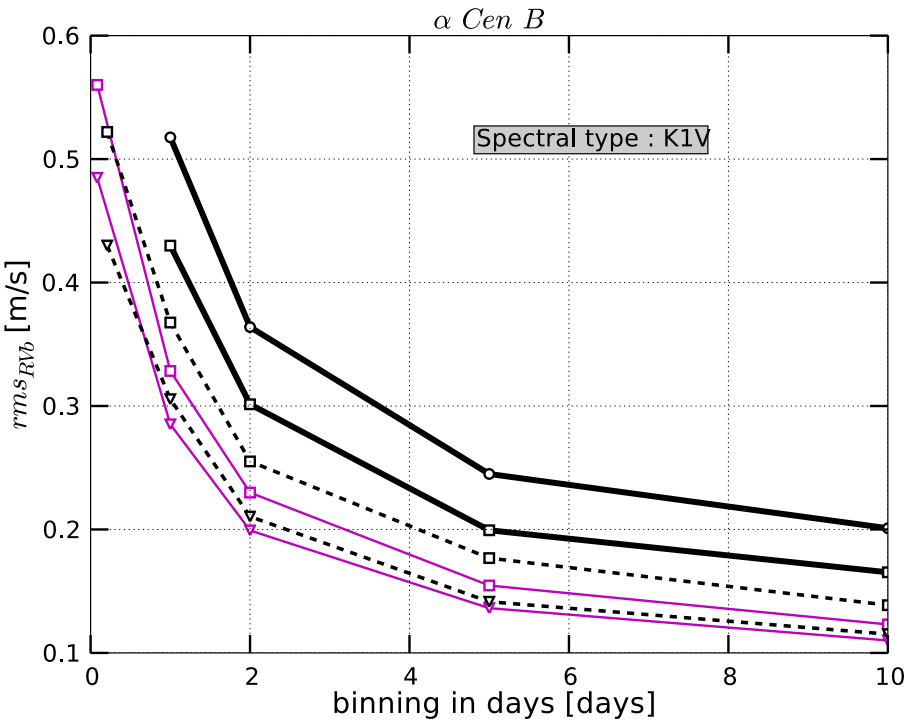

Figure 6.9: Radial-velocity dispersion as a function of the binning for different observing strategies (i.e. number of radial-velocity measurement per night, exposure time used for the individual measurements and the time gap between the observations taken during a given night). The velocities were simulated using the noise model generated for α Cen B (HD 128621). The dates of observation were selected using the calendar of the HD 69830 data set (Lovis et al. 2006). This calendar is typical of a HARPS GTO monthly 10 night run. Stellar activity induced effects are not included in the noise model used here. The observing strategies displayed here are: 1 meas./night (black solid curves); 2 meas./night with a 5 h time gap (black dashed curves) and 3 meas./night with a 2 h gap (solid magenta curves). The symbols represent the total integration time for a given night: 15 min (open circles, 1×15 min); 30 min (open squares, 1×30 min or 2×15 min or 3×10 min) and 1 h (open triangles, 2×30 min or 3×20 min). From Dumusque et al. (2009).

The third point to be discussed is the observing strategy. First, it is mandatory to take more than one observation per night because this is the only way for averaging out oscillation and granulation signals. As explained before, if this condition is matched and the targets are bright enough (this is true for the targets considered here), then the DIT can be reduced down to 5 min. We see in Fig. 6.9 that strategies involving 3 meas./night seem more efficient than those involving only 2 meas./night. The efficiency difference between the two strategies increases when the DIT decreases. So with a DIT of 5 min, a strategy with 3 meas./night has to be preferred. The next parameter of the observing strategy is the sampling of the observing nights. The suggested strategy in the DRM proposal was: 3 meas./night every night for a 10 night observing run). This was based on noise simulations in which stellar activity was not included. With stellar activity, such a strategy is no longer valid. The primary targets of the proposed planet search are low-activity Solar-type stars. The typical rotation period of this kind of target is 1 month. In other words, the initially proposed strategy only allows a poor sampling of the activity cycle resulting in a non optimal averaging out of activity noise. We would thus recommend another strategy: 3 meas./night every 3 nights. This way, the total number of observations remains the same and the sampling is much better. The last parameter related to the observing strategy is the total number of observing nights per season. Instead of taking 3 meas./night every 3 nights, one could consider to take 3 meas./night *every* night. Here, the conclusion is easy. The detection efficiency always increases when the the number of observing epochs increases. Fig. 6.13 shows the results obtained in this 1 epoch/day case. This has to be compared to

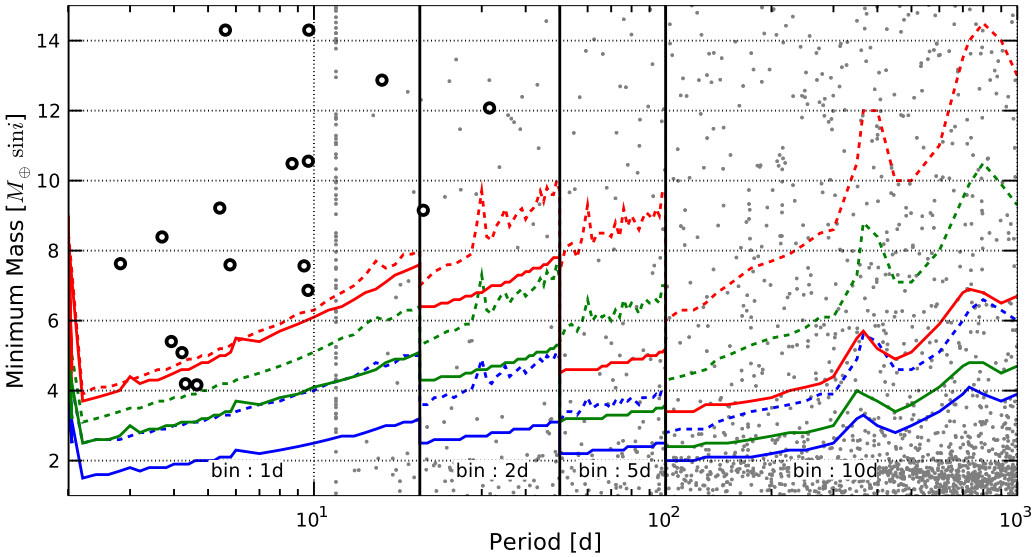

Figure 6.10: Effects of the binning, of stellar activity and of the observing strategy on the detection limits for a K1V bright star. The instrumental and photon noise contributions, although present, are negligible. The activity level is colour-coded: $\log R^{'}_{\rm HK} = -5.0$ (blue), -4.9 (green) and -4.75 (red). Results of a HARPS-like observing strategy (one 15 min measurement every night during a 10 night run, one observing run per month) are displayed with dashed curves. Results of the "optimal" observing strategy (three 10 min measurements every 3 nights during the entire observing season) are shown with solid curves. The binning length is adapted to the period P of the planetary companion: 1 d-binning for $10 \leq P \leq 20$ d; 2 d for $20 < P \leq 50$ d; 5 d for $50 < P \leq 100$ d and 10 d for $P > 100$ d. The small grey dots are synthetic planets from the Bern group formation models (Mordasini et al. 2009) and the big open dots represent discovered and confirmed small planets orbiting K dwarfs (from K0V to K9V). Adapted from Dumusque et al. (2010).

Fig. 6.12. The difference is obvious as this "ultimate" strategy could allow the detection of planets lighter than 1 $M_\oplus$ in the habitable zone of inactive K dwarfs. The only reason not to apply such a strategy is of course its prohibitive cost in terms of telescope time.

Regarding the target characteristics, the conclusions are easy to draw. Later spectral types (namely K dwarfs) have several advantages. Their lower masses allow the detection of lower mass planets as the signal induced by a planet of a given mass increases when the primary mass decreases. Low-mass planets are easier to detect in the habitable zone around them because it is located closer. Noise contributions from stellar oscillations and granulation have smaller amplitudes and shorter periods and they are thus easier to average out. The result of all these effects can easily be seen in Fig. 6.14 where the blue curves are well below the red ones. Stellar activity also plays an important role as expected. It induces signals that can be reduced down to some level but the careful selection of targets as inactive as possible is of critical importance when one wants to detect very-low amplitude Keplerian signals.

The next point we want to mention is related to the instrument/telescope combination. Several times throughout this report, we have mentioned the fact that the instrument and photon noise contributions to the error budget were negligible. From the results of the various simulations we have performed, we have reached the conclusion that obtaining noise rms better than a 10 cm s^{-1} was a very difficult task. In all the cases we have considered, the photon-noise term was lower than 5 cm s^{-1}. The quadratic sum of this value and the assumed instrumental error remains below 5.5 cm s^{-1} which is much smaller than these "ultimate" 10 cm s^{-1} and we can neglect these contributions. This result critically depends on the noise model used. This model, or at least its oscillations and granulation part, critically depends on the HARPS input data used for building up the PSDs. It may well be that the HARPS precision is ultimately

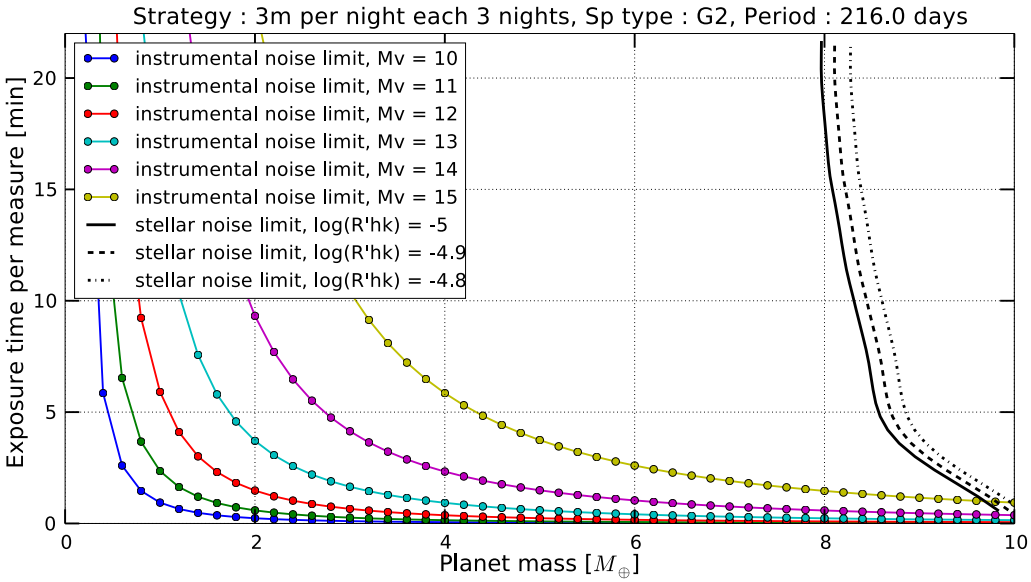

Figure 6.11: Effect of the DIT on the radial-velocity detection limits for a planets at the inner edge ($P = 216$ d) of the habitable zone around a G2V star. The black curves show the detection limits for 3 different activity indeces: $\log R^{'}_{\rm HK} = -4.75$ (dotted), $\log R^{'}_{\rm HK} = -4.9$ (dashed) and $\log R^{'}_{\rm HK} = -5.0$ (solid, spotless and plageless case). The black curves only account for the effect of stellar intrinsic noise sources (activity, pulsation and granulation) and thus correspond to the case of bright stars. The coloured curves show the instrument noise limits, i.e. the limits one would obtain in the absence of stellar noise sources for several magnitudes. The blue curve shows that the effects related to the instrument can be neglected for targets brighter than $m_V < 10$ mag. Increasing the DIT only has a small effect on the detection limits (small leftward shift).

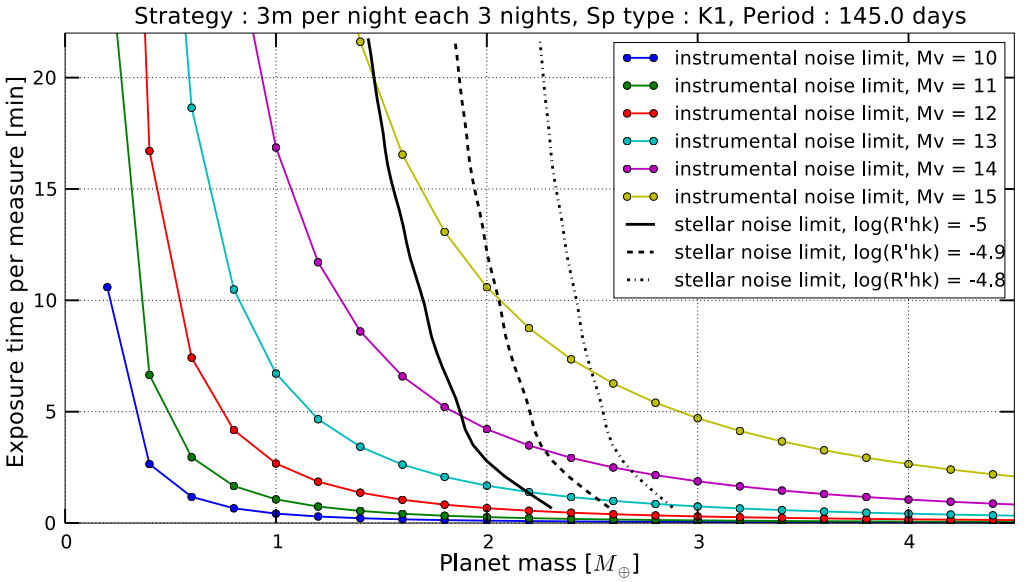

Figure 6.12: Same as for Fig. 6.11 but for a K1V star where the inner edge of the habitable zone corresponds to a period of 145 d. As for the G2V case, increasing the DIT above 5 min only marginally improves the detection sensitivity and instrument related effects can be neglected for bright targets.

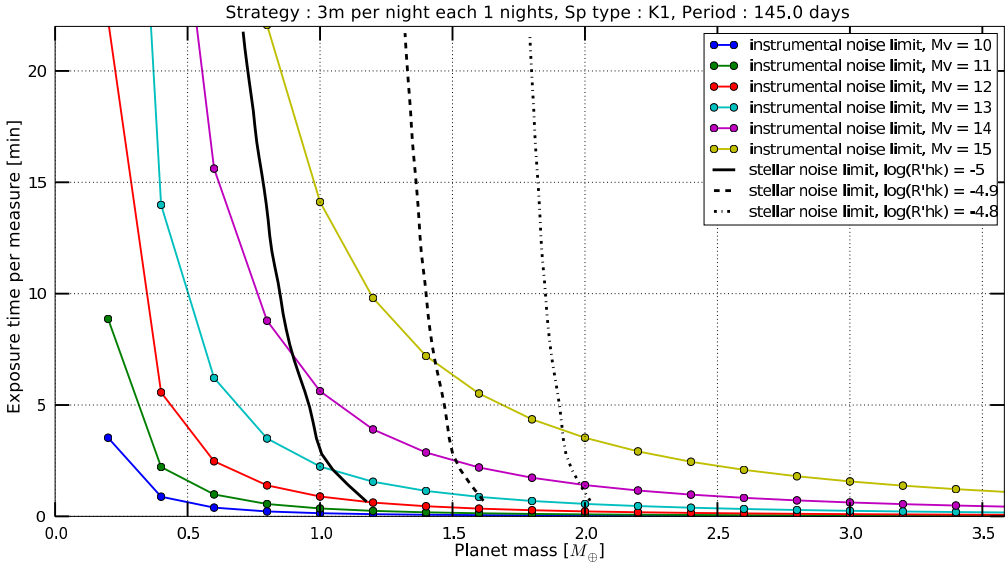

Figure 6.13: Same as for Fig. 6.12 but with a different observing strategy: 3 measurements per night every night instead of just every 3 nights. With this ultimate strategy, we see that the detection of a 1 $M_\oplus$ planet at the inner edge of an inactive K1V star is possible. Of course, the telescope time cost of such an observing strategy would be extremely high.

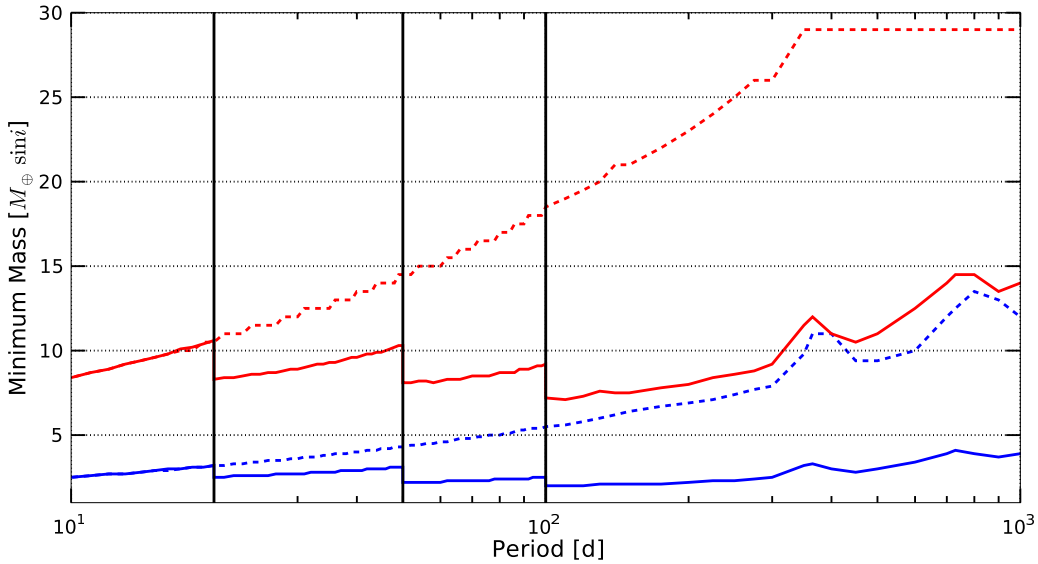

Figure 6.14: Effects of the binning and of the spectral type on the radial-velocity detection limits. The red curves show the case of the star α Cen A (HD 128620, G2V star) and the blue ones the case of α Cen B (HD 128621, K1V star). Three error sources have been neglected here: instrumental error, photon-noise and activity jitter. These detection limits thus correspond to the ideal cases of bright ($m_V \ll 10$ mag) spotless and plageless stars only affected by pulsation and granulation noise. The binning length is adapted to the period P of the planetary companions: 1 d-binning for $10 \leq P \leq 20$ d; 2 d for $20 < P \leq 50$ d; 5 d for $50 < P \leq 100$ d and 10 d for $P > 100$ d. The dotted curves show the detection limits obtained with a 1 d-binning for $P > 10$ d. Adapted from Dumusque et al. (2010)

not sufficient when dealing with such small signals. An instrument like ESPRESSO will probably play a decisive role in addressing this issue.

Finally, using Fig. 6.14, we estimate the mass of the smallest planet in the habitable zone detectable with the E-ELT under all the assumptions made in this study: 2 $M_{\oplus}$ if the primary star is an inactive K1V and 8 $M_{\oplus}$ if the primary star is an inactive G2V. In the case of a low-activity ($\log R^{'}_{\rm HK} = -4.9$) K1V star, the minimum mass increases to 2.5 $M_{\oplus}$ and to 4 $M_{\oplus}$ when $\log R^{'}_{\rm HK} = -4.75$.

6.5.3 Compliance with figures of merit

Depending on the definition used for an Earth-twin, the scientific goal of detecting such a planet in the habitable zone can or cannot be achieved. If an Earth-twin is defined as a 1 $M_{\oplus}$ planet, then the goal will probably not be achieved at least with a reasonable strategy. If a 2 $M_{\oplus}$ planet can still be considered as an Earth-twin (such a planet would only be 25% bigger in radius than the Earth if the densities are the same), then the goal can be achieved, at least for the case of an inactive K1V primary star.

6.5.4 Sensitivity to input parameters

Most of the input parameters related to the observing strategy play a significant role in the detection efficiency. Among these parameters, the number of observations per night (which has to be bigger than one) and the sampling of the observing epochs (1 epoch every 3 nights) are clearly the most important. However, the most sensitive input parameters are the ones related to the stellar characteristics (spectral type and activity level). Because of this fact, the target selection based on data obtained with the best radial-velocity instruments (HARPS and ESPRESSO in the future) will be of crucial importance.

6.5.5 Calibration requirements

Throughout this study we have made the assumption that the contribution to the radial-velocity error budget from the instrument is at the level of 2 cm s^{-1} over the timescale of several years. Hence we require that the instrument can be calibrated to this level. This requires in particular an extremely precise and ultra-stable wavelength calibration.

6.5.6 Limitations

We see three main limitations in our study. The first two are related to stellar activity. Our results, together with theoretical predictions (Christensen-Dalsgaard 2004), show that K dwarfs are the most promising targets for such a scientific programme. There is little doubt that both stellar oscillations and granulation are less problematic for K dwarfs than for G dwarfs. Now, our simulations of the effect of stellar activity were built on the Solar case as the Sun is the only star for which detailed observational constraints on cool spots are available. We have assumed that these constraints are also valid for cooler stars. This may not be the case.

The second limitation is related to the low frequency part of the PSDs we have obtained in Section 6.4.1.3. We have assumed that this part of the spectrum was generated by granulation. In fact, part of the power at low frequencies may also result from stellar activity. The data sets used for computing the PSDs have time spans of several days (i.e. a significant fraction of the star rotation period) and we cannot exclude the presence of a small activity signature in these PSDs. If this were the case, then the effect of stellar activity included in our simulations would be slightly over-estimated.

Finally, the last limitation we want to mention is related to the input data used for constructing our noise model. Although they were obtained with the instrument that is currently by far the best radial-velocity machine, HARPS, their intrinsic precision may be insufficient for correctly describing what is happening at or below the 10 cm s^{-1} level.

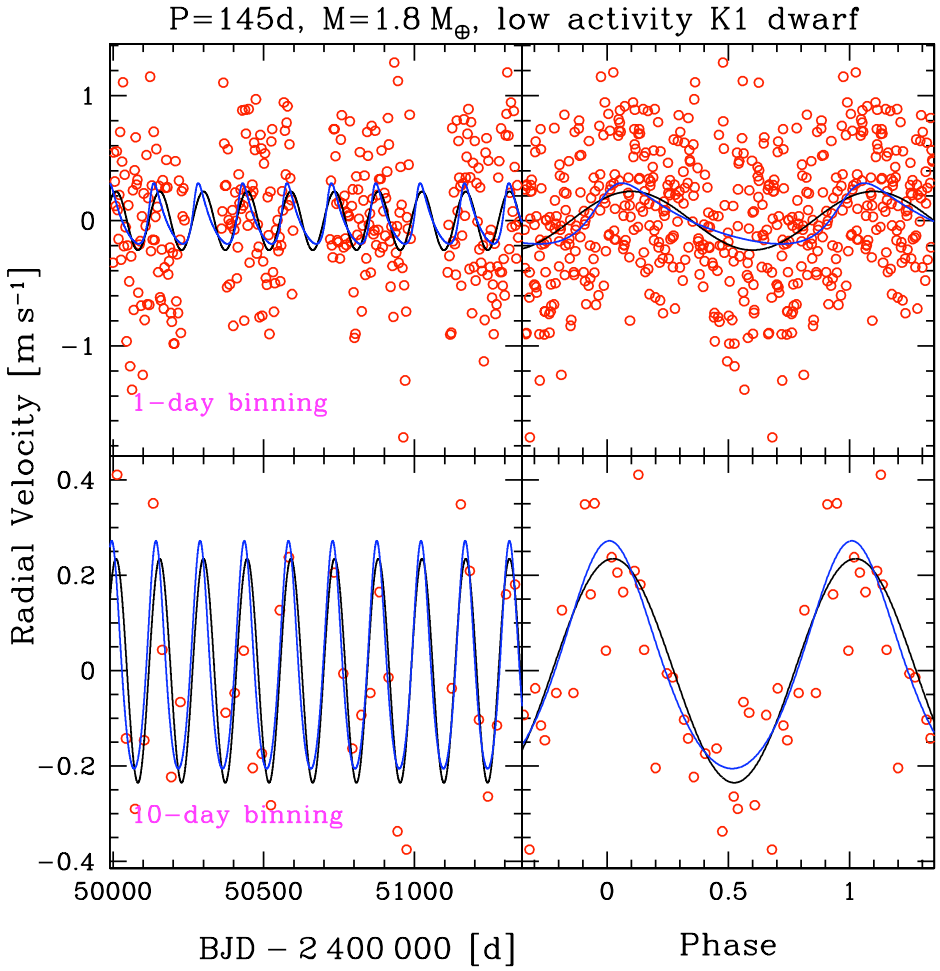

Figure 6.15: Simulated data set. Radial-velocity data (red dots) corresponding to the case of a 1.8 $M_\oplus$ planet at the inner edge of the habitable zone around a bright K1V low-activity star ($\log R^{'}_{\rm HK} = -4.9$).The observing strategy was the following: 3 meas./night every 3 nights obtained during four 8-month observing seasons. This case has a $K/\sigma \simeq$ ratio of 2.2 i.e. slightly below our selected detection threshold (2.5). We show in Figs. 6.16 and 6.17 that such a limit case is nevertheless a highly significant detection. Top panels: data obtained with 1-day binning. Bottom panels: data obtained with 10-day binning. Left panels: temporal data. Right panels: phase-folded data. In all panels, the black curves correspond to the input orbit and the blue curves to the orbits fitted to the data.

6.6 Concluding remarks

Finally, we would like to soften a little bit our slightly pessimistic conclusions of the previous sections. It is clear that the various sources of stellar intrinsic noise form an obstacle making the detection of an Earth-twin in the habitable zone a very difficult task. However, we have used a very robust detection threshold. It is already certain that planets below this threshold will be detected. To illustrate this fact, we have generated a radial-velocity data set for a 1.8 $M_\oplus$ planet at the inner edge of the habitable zone ($P = 145$ d) around a low-activity ($\log R^{'}_{\rm HK} = -4.9$) bright K1V star. This data set was obtained using our noise model and the observing strategy is the one we consider as the optimal one (i.e. efficient and not too expensive): three 10 min meas./night every 3 nights. Fig. 6.15 shows this data set with two different binnings (1 and 10 d). We also display the orbits fitted to the data (independent fits were performed for the two binnings). We also show the input orbit. The results of the fits are presented in Table 6.2. We see that the fitted parameters are very close to the input ones for both fits. This shows that the binning process is not critical for obtaining the correct orbital parameters. It only helps in making the detection more visible or more obvious.

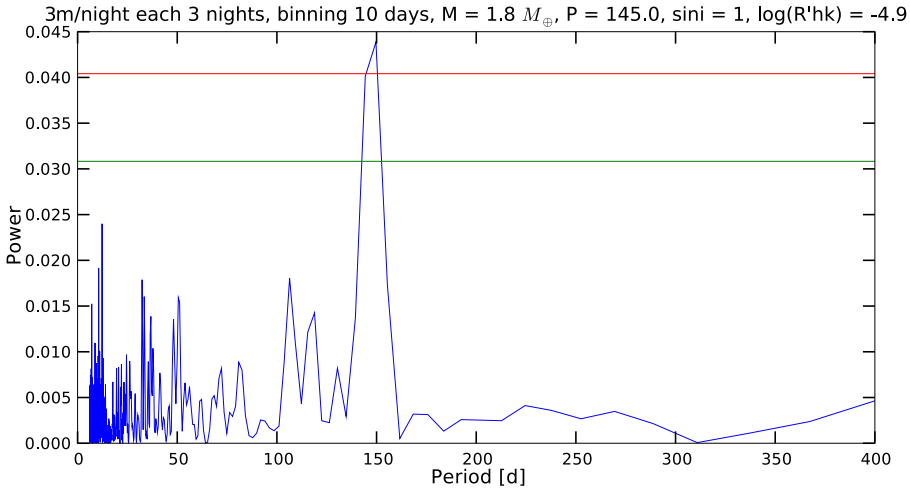

Figure 6.16: Periodogram of the simulated data set presented in Fig. 6.15. The green horizontal line represent the power of a 2σ detection and the red one a 3σ detection. The input period, 145 d is detected with a low false-alarm-probability.

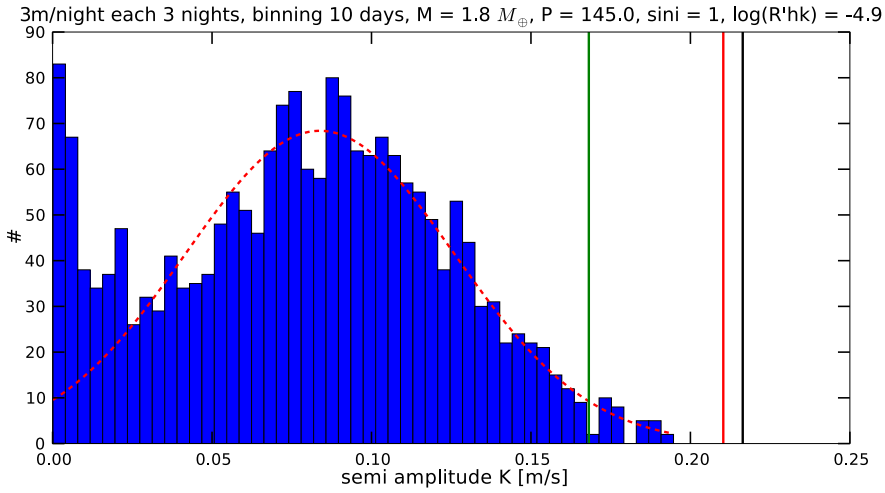

Figure 6.17: Bootstrap test. Simulated data have been randomly re-arranged. Orbital fits have been performed to 2000 re-arranged data sets. The blue histogram shows the distribution of the fitted radial-velocity semi-amplitudes. The red dotted curve shows the result of a Gaussian fit to the distribution. From this fit, we computed the 2σ (green line) and 3σ (red line) semi-amplitudes. The black line corresponds to the semi-amplitude fitted to the original data set. Clearly, the detected signal is very significant.

Table 6.2: Simulated data set. Comparison between the input orbital parameters and the orbital parameters fitted to the simulated data set presented in Fig. 6.15. We show the results for two different binnings (1 and 10 days).

Parameter	Unit	Input	1-day binning	10-day binning
P	[d]	145.0	147.18 ± 1.17	146.55 ± 1.00
e		0.0	0.34 ± 0.16	0.14 ± 0.13
K_1	[m s^{-1}]	0.235	0.24 ± 0.08	0.24 ± 0.04
$M_{\rm planet}$	[$M_\oplus$]	1.8	1.89 ± 0.54	1.82 ± 0.51

Figs. 6.16 and 6.17 further demonstrates that this detection is significant and that there would not be any doubt about it. The rms of the data with a binning of 1 d is 52 cm s^{-1} which is more than twice as big than the radial-velocity semi-amplitude. The rms of the residuals around the orbit fitted to the 10-day binning data is 11 cm s^{-1}. This latter value corresponds to the $\sigma_{\rm noise}$ term of the SNR relation (equation 6.1). With a fitted semi-amplitude of 24 cm s^{-1}, we compute a SNR of 2.18 which is clearly below the 2.5 threshold adopted in our study. This example demonstrates that the minimal masses we presented at the end of Section 6.5.2 are somewhat conservative and that slightly lower-mass planets are within reach.

References

Andersen B.N., Leifsen T.E., Toutain T., 1994, Sol. Phys., 152, 247
Arentoft T., Kjeldsen H., Bedding T.R., Bazot M., Christensen-Dalsgaard J., et al., 2008, ApJ, 687, 1180
Bazot M., Bouchy F., Kjeldsen H., Charpinet S., Laymand M., Vauclair S., 2007, A&A, 470, 295
Christensen-Dalsgaard J., 2004, Sol. Phys., 220, 137
Del Moro D., Berrilli F., Duvall Jr. T.L., Kosovichev A.G., 2004, Sol. Phys., 221, 23
Dumusque X., Santos N., Udry S., 2010, A&A, submitted
Dumusque X., Udry S., Lovis C., Santos N., 2009, A&A, submitted
Harvey J.W., 1984, in "Probing the Depths of a Star: the Study of Solar Oscillations from Space", R.W. Noyes & E.J. Rhodes Jr., eds., Pasadena, JPL/NASA
Huélamo N., Figueira P., Bonfils X., Santos N.C., Pepe F., Gillon M., et al., 2008, A&A, 489, L9
Kjeldsen H., Bedding T. R., Butler R.P., Christensen-Dalsgaard J., Kiss L.L., et al., 2005, ApJ, 635, 1281
Lagrange A., Desort M., Meunier N., 2010, A&A, 512, 38
Lefebvre S., García R.A., Jiménez-Reyes S.J., Turck-Chièze S., Mathur S., 2008, A&A, 490, 1143
Lomb N.R., 1976, Ap&SS, 39, 447
Lovis C., Mayor M., Pepe F., Alibert Y., Benz W., Bouchy F., Correia A.C.M., et al., 2006, Nature, 441, 305
Mayor M., Pepe F., Queloz D., Bouchy F., Rupprecht G., et al., 2003, The Messenger, 114, 20
Mordasini C., Alibert Y., Benz W., Naef D., 2009, A&A, 501, 1161
Noyes R.W., Hartmann L.W., Baliunas S.L., Duncan D.K., Vaughan A.H., 1984, ApJ, 279, 763
Scargle J.D., 1982, ApJ, 263, 835
Schrijver C.J., Zwaan C., 2000, Solar and Stellar Magnetic Activity, Cambridge University Press
Selsis F., Kasting J.F., Levrard B., Paillet J., Ribas I., Delfosse X., 2007, A&A, 476, 1373
Udry S., Santos N.C., 2007, ARA&A, 45, 397
Zechmeister M., Kürster M., 2009, A&A, 496, 577

7 S9: Imaging the planet-forming regions of circumstellar disks

Authors: D. Villegas, M. McCaughrean

7.1 The science case

In the paradigm of the formation of low-mass stars and their planetary systems, circumstellar disks are a critical component, as they provide a conduit to channel material onto the forming star (Hartmann et al. 1997) and supply a reservoir of material from which planets may form (Bodenheimer & Pollack 1986). They are also directly involved in the formation and collimation of jets and outflows, which are generally thought to be implicated in the key process of angular momentum removal (e.g. Coffey et al. 2004).

After centuries of speculation that our own solar system must have formed from a rotating disk of gas and dust, the existence of circumstellar disks around other young stellar objects is now firmly established (Beckwith et al. 1990; Strom et al. 1993; Beckwith & Sargent 1996). Disks are found around ∼30–50 % of all low-mass (0.3–3 $M_{\odot}$) young (0.3–10 Myr) stars, with radii ∼10–1000 AU, and masses ∼0.01–0.1 $M_{\odot}$. Most of these results are based on indirect measurements (e.g. spectral energy distributions, asymmetric wind profiles, polarisation mapping) but in the past decade well-resolved direct images of circumstellar disks have become available and can provide us with important information on the structure of the disks and how they are affected by their ambient environment (e.g. Burrows et al. 1996; McCaughrean et al. 1996; Watson et al. 2007).

Over 10–100 Myr, these optically-thick young disks lose their original gas and dust through accretion onto the central star, photo-evaporation, or agglomeration and accretion to form planets and planetesimals. The reservoir of small dust grains is then replenished by collisions between rocky planetesimals, leading to the so-called dusty “debris disks” seen around older nearby stars such as Vega, β Pictoris, Fomalhaut, ε Eridani, AU Microscopii and HD 107146. There is considerable evidence that non-axisymmetric structures in these debris disks can be used to infer the presence of planetary systems (e.g. Kalas et al. 2005). The transition between young, optically-thick, gas-dominated disks and older, residual dust debris disks is crucial to our understanding of the planet formation process (e.g. Meyer et al. 2007).

In particular, high spatial resolution direct imaging of disks from 1–100 Myr in age has a crucial role to play, as protoplanets and mature planets can leave tell-tale signatures in the disks in terms of radially axisymmetric structures such as gaps and belts, and non-axisymmetric structures such as resonances and spiral density waves (Kalas et al. 2006; Wyatt et al. 2006). Fig. 7.1 shows some examples of known circumstellar disks in nearby star forming regions.

However, such direct imaging studies are very challenging, particularly in the crucial regions where analogues of gas giants like Jupiter and terrestrial-mass planets like the Earth are expected to form. This is simply a matter of the available spatial resolution: the 5 AU radius orbit of Jupiter would subtend only 33 mas at 150 pc, the distance to the nearest low-mass star-forming regions, while a search for signs of Earth-like planets forming in a potentially habitable zone would require just 7 mas resolution at the same distance. Matters are somewhat improved for the older, nearby debris disks, which have been probed in optical/near-IR scattered light with HST, VLT, Gemini and Keck, and in direct thermal emission from dust grains by Spitzer and ground-based sub-mm/mm telescopes. The same fiducial distances of 5 AU and 1 AU subtend 170 mas and 35 mas at 20 pc: such resolutions are achievable at optical/near-IR wavelengths with HST and ground-based AO, but not by Spitzer or single dish sub-mm telescopes. A fully diffraction-limited E-ELT should yield angular resolutions of 13 mas at 2 μm, 60 mas at 10 μm and 120 mas at 20 μm, enabling very substantial gains in disk imaging, both in terms of scattered light at the shorter wavelengths and direct thermal emission from dust at the longer wavelengths. Images obtained of scattered light and warm (100–300 K) thermal emission from disks at near to mid-IR wavelengths, respectively, with the E-ELT would be directly complementary to those obtained at sub-mm/mm wavelengths with ALMA of the colder dust and gas at larger distances from the parent star, with a spatial resolution of ∼0.1″ at 1 mm, equivalent to 15 AU at 150 pc. Such images would be further complemented by the E-ELT spectroscopic capability to study the dynamics of gas and the chemical and physical evolution of gas and dust in the inner regions of circumstellar disks.

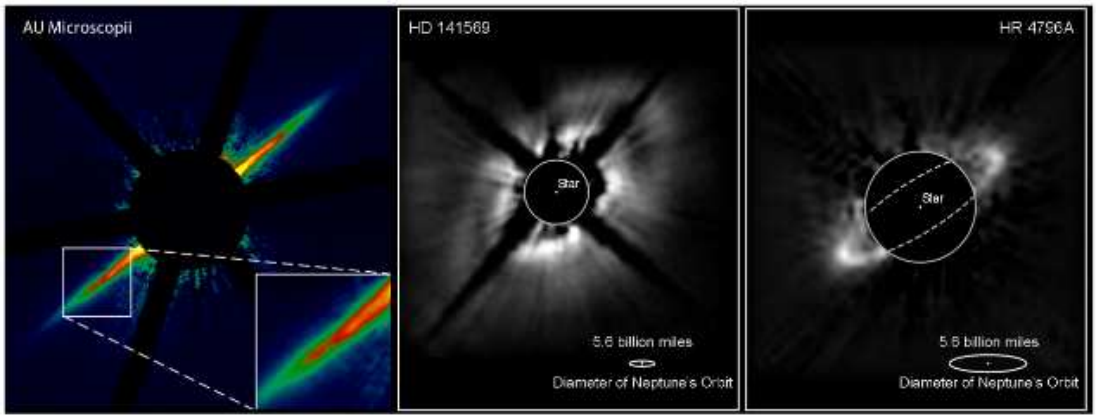

Figure 7.1: Near-IR scattered light coronagraphic images of representative circumstellar disks. From left to right: AU Microscopii (Keck image, resolution $\sim$0.1″, 1 AU at 10 pc; Kalas et al. 2004); HD 141569 (NICMOS image, resolution $\sim$0.2″, 20 AU at 100 pc; Weinberger et al. 1999); HR 4796A (NICMOS image, resolution $\sim$0.2″, 15 AU at 70 pc; Schneider et al. 1999).

This project aims to investigate the spatial structure of circumstellar disks via direct near and mid-IR imaging with the E-ELT. The goal is to test the detectability of gaps and non-axisymmetric structure such as spirals and hotspots in the young and more mature disks indicative of ongoing or completed planet formation. In particular, we want to reach as far as possible into the zone where putative terrestrial-mass habitable planets may be forming or have already formed.

7.2 Goals of the DRM simulations

The DRM simulations will produce near and mid-IR images of young, low-mass circumstellar disks with evidence of gaps formed as a result of an ongoing planet formation process. Given the large parameter space involved in this kind of system we decided to study the feasibility of detection of planetary formation features in a “model case” as the distance to the observer varies.

We have defined the following goals, ordered by increasing complexity in the simulations:

- Simple detection of a young, low-mass circumstellar disk. We aim to trace the detectability of a continuous disk as a function of distance to the observer. How far away can we observe and still detect emission coming from the disk within a reasonable amount of observing time?
- To which extent would planetary formation features be detected in a model case of a T Tauri star surrounded by a young, low-mass circumstellar disk with a 2 AU-wide gap at an orbital radius of 5/10 AU (which roughly corresponds to the signature of a 1 Jupiter-mass planet). This case will be tested as a function of the distance to the observer, site location, etc.
- Simulations of coronagraphic imaging will be required at short wavelengths in order to suppress the central star from a disk-star composite model. This will allow us to test the level of suppression against the occulting disk and the minimum distance from the star that can be probed given the telescope capabilities.

7.3 Metrics / figures of merit

As a general metric we will use the distance at which a “standard” disk and its particular planet formation features can be detected within a maximum observing time of one night.

The criteria to define a “detection” will be derived from a comparison between the azimuthally averaged radial profile of the star+disk observation and the noise expected during an equivalent observation of a star that does not have any circumstellar material.

7.4 DRM simulations

7.4.1 Methodology

Given the relatively large parameter space of this science case we decided to constrain the simulations to a model case of a T Tauri star surrounded by a face-on circumstellar disk with the physical properties described in Section 7.4.3.1. This basic disk model will then be studied as a continuous disk and also in a disk+gap configuration.

The continuous disk case was designed to trace the detectability of the disk component as the distance to the observer increases. The disk+gap model will be tested in two configurations of a 2 AU-wide gap that differ only in the physical location of the gap. This gap morphology is meant to roughly represent the planet-formation feature of a 1 $M_{\rm Jup}$ planet and will provide us with a means to determine the level of detectability of these features. The gap will be modelled with two different particle distributions: a Gaussian shape that represents the short-lived period during which the planet is cleaning the area around its orbital radius, and a top-hat distribution without any particles in the gap area representing a fully cleaned orbital path.

This set of reference cases will be used to assess the influence on the results of the simulations of:

- the different disk physical parameters given a certain observational configuration;
- different observational parameters given a fixed set of physical properties of the disk.

7.4.2 Pipeline

Circumstellar disk images are generated using the 3D continuum radiative transfer code MCFOST (Pinte et al. 2006, see Section 7.4.3.1 for details). The model images are produced in three different configurations:

1. A T Tauri star surrounded by a continuous circumstellar disk.
2. A T Tauri star surrounded by a circumstellar disk with a 2 AU-wide gap (Gaussian and top-hat) at a distance of 5 AU from the star.
3. A T Tauri star surrounded by a circumstellar disk with a 2 AU-wide gap (Gaussian and top-hat) at a distance of 10 AU from the star.

These images are generated in the I, H, K, L, N and Q bands and are subsequently convolved with a PSF of the same wavelength (see Section 7.4.3.2) using the task FCONVOLVE in IRAF. The pixel scales of the original images are the same as those used in the PSF images, which are determined by the size of the diffraction limited core of the PSF at a given wavelength.

In parallel an image containing only a star (i.e. a delta function) is generated using IRAF's MKOBJECTS task. This image is then used to create a so-called "true PSF image" by convolving it with the same PSF used above to convolve the star+disk image with, as well as a simulated "PSF calibration image" by convolving it with a variety of different PSFs.

An E-ELT image simulator is then used to add all the remaining observational effects to the PSF-convolved frames. This package is based on the Large Binocular Camera Image Simulator (Grazian et al. 2004) and allows the creation of realistic 2D images of the sky as it will be observed by the E-ELT. The simulator accounts for the main instrumental characteristics such as the telescope diameter and the efficiencies of the various components (mirrors, instrument and CCD). Considering these parameters it generates an output image including noise, flat-field inaccuracies and sky background, and simulates the observational strategy: the detector integration time, the number of exposures and adopted pixel size at a given wavelength.

A schematic view of the simulation pipeline is shown in Fig. 7.2.

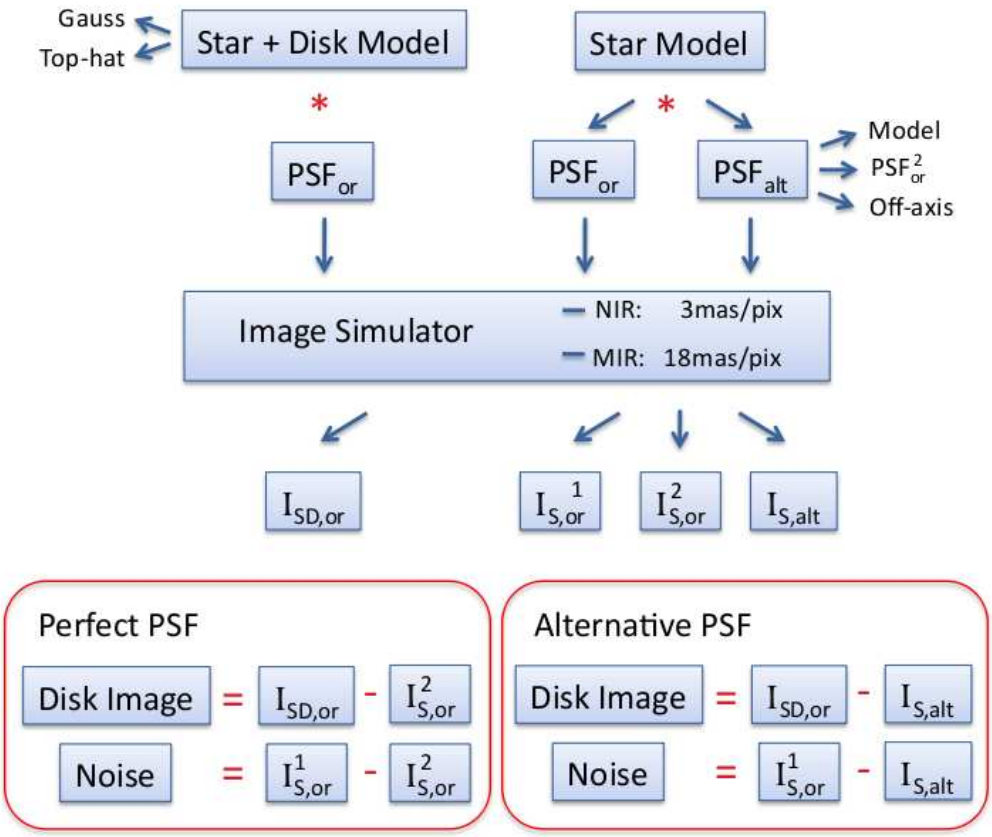

Figure 7.2: Schematic illustration of the method followed during the simulations.

7.4.3 Inputs

Scientific and technical data were provided as inputs to the simulations as follows.

7.4.3.1 Scientific data

The simulations were performed assuming a characteristic circumstellar disk model based on the observational properties of IM Lupi, an M0 T Tauri star located within the Lupus star forming clouds ($d \sim 140$ pc). Despite its relatively low emission line activity the presence of an extended disk surrounding this star has been detected in scattered light.

Synthetic images of circumstellar disks were computed using MCFOST, a 3D continuum radiative transfer code based on the Monte-Carlo method (Pinte et al. 2006). The model includes multiple scattering with a complete treatment of polarisation, dust heating and continuum thermal re-emission. Dust properties are modelled to reproduce vertical dust settling such as the one observed in our model case IM Lupi. Calculations are performed using a long-characteristic Monte Carlo method similar to the one presented in Hogerheijde & van der Tak (2000).

The images are produced assuming an axisymmetric flared density structure of the dust disk with a Gaussian vertical profile. In cylindrical coordinates we have:

$$\rho(r, z) = \rho_0(r) \exp(-z^2/2\, h(r)^2) \tag{7.1}$$

which is valid for a vertically isothermal, hydrostatic, non-self-gravitating disk. We use power-law distributions for the surface density $\Sigma(r) = \Sigma_0(r/r_0)^\alpha$ and the scale height $h(r) = h_0(r/r_0)^\beta$, where $h_0 = 10$ AU

Table 7.1: Sky position of some characteristic nearby star forming regions. Columns 5 and 6 show the best possible airmass reachable from the latitude of Cerro Paranal in Chile and the Canary Island of La Palma, respectively.

Region	RA	Dec	Distance	Paranal	La Palma
	J2000		[pc]	sec(z)	sec(z)
Chamaeleon	11 53	$-$79 06	150	1.7	–
Lupus	16 03	$-$38 06	140	1.0	2.6
Ophiucus	16 28	$-$24 32	150	1.0	1.7
Orion	05 35	$-$05 23	400	1.1	1.2
Taurus	04 41	$+$25 52	150	1.6	1.0

is the scale height at radius $r_0 = 100$ AU. The values adopted for the simulations are $\alpha = -1.5$ and $\beta = 1.25$. The total dust mass of the disk is $10^{-3}\ M_\odot$. We consider homogeneous spherical grains and calculate optical properties with Mie theory. The grain sizes are distributed according to the power-law $dn(a) \propto a^{3.5}\ \mathrm{d}a$, with minimum and maximum grain sizes of 0.03 μm and 1 mm, respectively.

The final goal of these simulation is to test the detectability of planetary formation features in the circumstellar disks. These features have been implemented in the model in a simple way by generating a gap in the continuity of the particles in the disk at a given radius. The simulations have been implemented with two different gap morphologies. In the first approach we use an inverted Gaussian profile with zero-counts at its peak value and a full width at half maximum of 2 AU. This is meant to emulate the short-lived planet-formation signature of a $\sim$1 M_{Jup} planet that is "cleaning" its orbit as part of its own assembly process. A perhaps more realistic scenario was also modelled by generating a top-hat-shaped gap without any particles over the 2 AU extent of the gap, representing a planet that has completely cleaned its orbital path. This may be a better approximation of the kind of feature we should expect to observe because this phase will last much longer than the cleaning phase. Three test cases have been implemented: (i) a continuous disk (i.e. no gap), (ii) a disk with a Gaussian/top-hat gap at 5 AU and (iii) a disk with a Gaussian/top-hat gap at 10 AU. Fig. 7.3 shows azimuthally averaged radial profiles of all the 29 (the I-band model with a Gaussian gap at 10 AU is not available) star+disk models used in these simulations.

Table 7.1 lists the sky positions of a set of nearby star forming regions where evidence of circumstellar material surrounding young stars has been detected. We list the coordinates and the distances to these regions, as well as the best possible airmass reachable from a site located at the latitude of Cerro Paranal in Chile and La Palma in the Canary Islands. According to these values good targets are available in both hemispheres. We have performed the simulations assuming an average airmass of 1.15, but this parameter does not play an important role in shaping the results.

7.4.3.2 Technical data

The technical data used for the simulations were taken mainly from [RD1].

Site and background: We follow [RD1]. We use the background emission and atmospheric transmission models for both the Paranal-like and High & Dry sites. We assume a zenith distance of 30° (i.e. an airmass of 1.15) for all of the simulation runs.

Telescope: As described in [RD1]. We use both the bare aluminium and the protected silver/aluminium coatings.

Instrument: We assume an imager operating in the near and mid-IR wavelength regimes, with 3 mas pixels in the near-IR and 18 mas pixels in the mid-IR. With these pixel sizes the images are significantly under-sampled at the shortest wavelengths in each regime (i.e. in the I and L bands). Due to the smallness of our targets (only a few arcsec even for the nearest objects) the field of view of the instrument is practically irrelevant. We use the read-out noise, dark current and detector QE values for each regime as specified in [RD1]. Apart from the detector QE no further instrument efficiency is applied. We assume a detector saturation level of 120 000 counts and limit the integration times of individual exposures

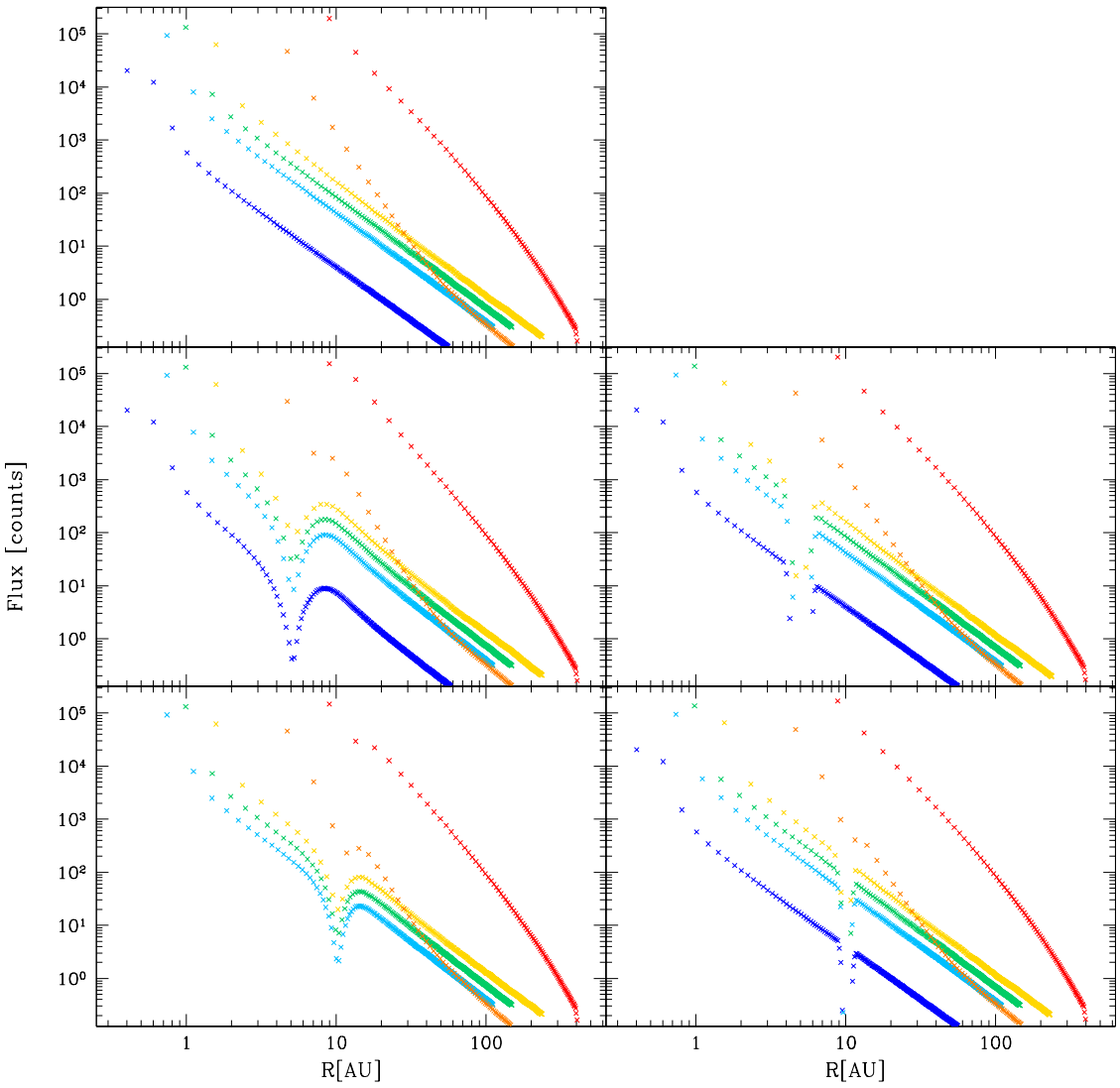

Figure 7.3: Azimuthally averaged radial intensity profiles of the star+disk models in I (blue), H (light blue), K (green), L (yellow), N (orange) and Q-band (red). From top to bottom the figures show a continuum disk, a disk with a 2 AU-wide gap at 5 AU and a disk with a 2 AU-wide gap at 10 AU, respectively. Both the Gaussian (left panels) and the top-hat (right panels) gap morphologies are shown here.

to a maximum of 0.1 and 0.01 s in the near-IR and mid-IR, respectively, in order to limit the effects of saturation.

AO mode and PSFs: Given the need of a high quality on-axis AO correction in the small field of view subtended by the target star+disk system (only a few arcsec for an object at $\sim$140 pc), the simulations have been performed adopting Laser Tomography Adaptive Optic (LTAO) PSFs as described in [RD1]. We only used on-axis PSFs to convolve the star+disk images with, and we used the actual images rather than the analytical fits. The speckle noise in the PSF images is actually a welcome feature in this DRM case because we are dealing with very short exposure times here.

The PSF calibration image is generated using four different PSFs, in order to explore the effects of imperfect PSF knowledge on the final results of the simulations. The PSF models used are:

1. The same PSF image as the one that was used to convolve the "science" image with. This is the best case scenario where the PSF uncertainties are only due to photon and read-out noise. We use this as a reference case but we point out that it is quite unrealistic.
2. A PSF image generated from the analytic fit to the radial intensity profile of original PSF image. This model is intended to be representative of a PSF calibration obtained at the telescope by averaging several exposures of PSF calibration stars.
3. A different realisation of the AO simulation of the same PSF used for the "science" image convolution. In this case the "original" and the "alternative" PSFs are only different because of the random nature of the simulated atmosphere and because of the noise associated with the guide star images. These simulations are only available for the H and K bands.
4. A PSF image generated for a different position within the field of view, namely 10 arcsec off-axis. These simulations are also only available for the H and K bands.

Although we assumed a zenith distance of 30° when calculating the background emission, PSFs for this airmass were not available for all bands considered here, and so we only used PSFs that were simulated for airmass 1.

7.4.4 Outputs

For each simulation run, the following set of FITS files was generated in each of the bands studied:

- One "science" image containing the luminosity distribution of a star surrounded by a circumstellar disk modified to include all the observational features added by the image simulator, including the background.
- One "true PSF" image containing only a star but otherwise generated in exactly the same way as the "science" image above (including the use of the exact same PSF).
- Two–four "PSF calibration" images (depending on the band), each also containing only a star but using one of the four different PSFs discussed in the previous section in order to simulate the effect of imperfect PSF knowledge. We refer to these as the "perfect", "analytical", "different realisation" and "off-axis" PSF calibration images, respectively. Note that the "perfect" PSF calibration image is just another realisation of the "true PSF" image above.

7.5 Results of simulations

7.5.1 Simulation runs

The different simulation runs performed and their input parameters are listed in Table 7.2. In addition to these, all of the listed runs were also repeated with different technical input parameters to test for the effect of changes in observation parameters such as site of the observatory, mirror coating and pixel size.

We have only run the full simulation for the 4 shorter wavelengths (I, H, K and L) because, as already seen in Fig. 7.3, the chances of detecting the gap at longer wavelengths is negligible. Fig. 7.4 shows a comparison between the intensity profile of the (not-convolved) disk model with a gap at 5 AU and the expected PSF normalised to the flux of a T Tauri star at a given wavelength. We can clearly see that the most favourable contrast between the disk and the extended wings of the PSF is obtained in the L and Q bands. However, the gap is of course best resolved at the shortest wavelength.

Table 7.2: List of the simulation runs performed and their input parameters. The columns indicate: (1) type of disk model used, (2) passband, (3) PSF model used to subtract from the final simulated image, (4) assumed pixel size, (5) maximum time for a single exposure and total exposure time per simulation.

Model (1)	Band (2)	PSF (3)	Pixel size (4)	Exposure time (5)
No gap	I	original	3 mas	0.1 s / 10 s
		analytical model		
	H	original	3 mas	0.1 s / 10 s
		analytical model		
		different realisation		
		off-axis		
	K	original	3 mas	0.1 s / 10 s
		analytical model		
		different realisation		
		off-axis		
	L	original	18 mas	0.01 s / 10 s
		analytical model		
Gaussian gap 5 AU	I	original	3 mas	0.1 s / 10 s
		analytical model		
	H	original	3 mas	0.1 s / 10 s
		analytical model		
		different realisation		
		off-axis		
	K	original	3 mas	0.1 s / 10 s
		analytical model		
		different realisation		
		off-axis		
	L	original	18 mas	0.01 s / 10 s
		analytical model		
Top-hat gap 5 AU	I	original	3 mas	0.1 s / 10 s
		analytical model		
	H	original	3 mas	0.1 s / 10 s
		analytical model		
		different realisation		
		off-axis		
	K	original	3 mas	0.1 s / 10 s
		analytical model		
		different realisation		
		off-axis		
	L	original	18 mas	0.01 s / 10 s
		analytical model		
Gaussian gap 10 AU	H	original	3 mas	0.1 s / 100 s
		analytical model		
		different realisation		
		off-axis		
	K	original	3 mas	0.1 s / 100 s
		analytical model		
		different realisation		
		off-axis		
	L	original	18 mas	0.01 s / 100 s
		analytical model		
Top-hat gap 10 AU	I	original	3 mas	0.1 s / 100 s
		analytical model		
	H	original	3 mas	0.1 s / 100 s
		analytical model		
		different realisation		
		off-axis		
	K	original	3 mas	0.1 s / 100 s
		analytical model		
		different realisation		
		off-axis		
	L	original	18 mas	0.01 s / 100 s
		analytical model		

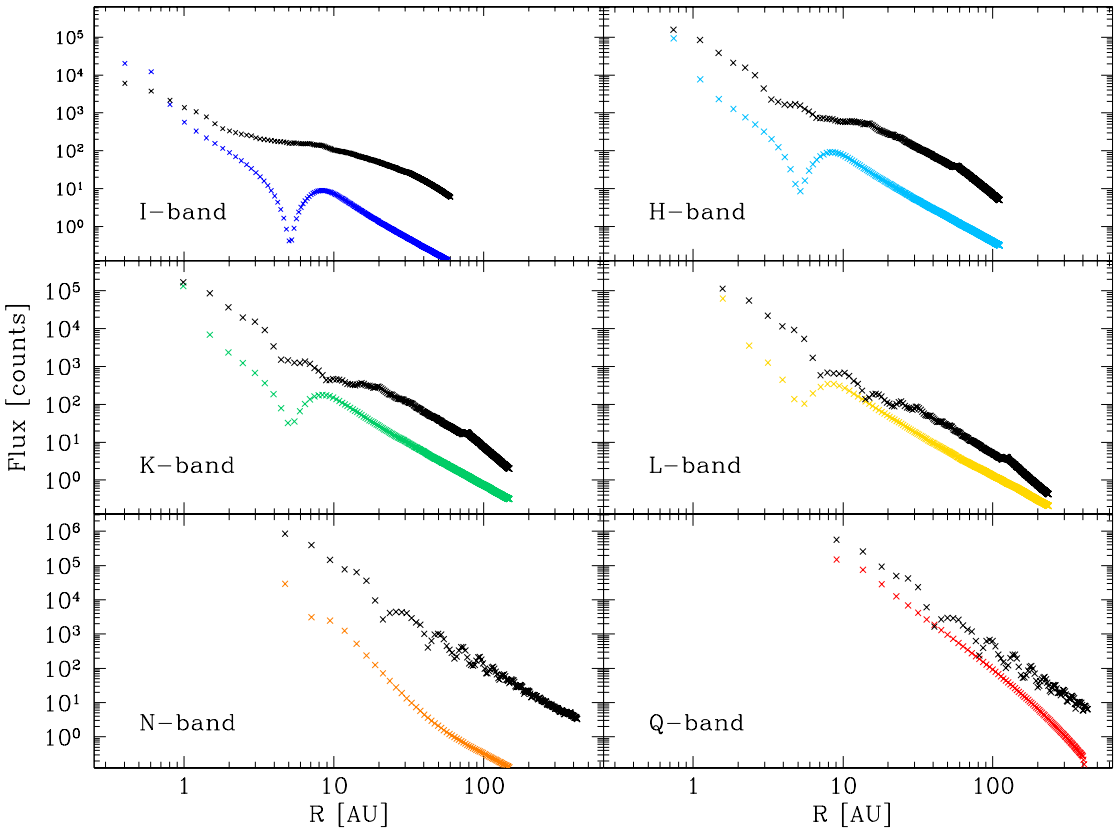

Figure 7.4: Azimuthally averaged radial intensity profiles of the disk model with a gap at 5 AU (not convolved with a PSF) compared to the profile of the PSF normalised to the flux of the T Tauri star in the corresponding wavelengths.

7.5.2 Analysis

We aim to detect traces of the light coming from the circumstellar disk in enough detail to be able to "see" the gap produced by a planet in formation. The first step then consists of isolating the light of the disk. To do this we subtract one of the PSF calibration images from the star+disk image. This procedure allows us to remove the star light as well as to account for the background, but (depending on which PSF calibration image is used) introduces more or less severe errors due to PSF uncertainties. We also generate a separate frame in which we subtract the PSF calibration image from the "true PSF" image. This delivers an estimation of the flux we will expect to obtain when performing the same procedure on the image of a star that does not host circumstellar material.

Our analysis consists then of a comparison between the flux level detected in the "disk-only" image (star+disk image − PSF calibration image) and the "PSF residuals" image ("true PSF" image − PSF calibration image). For the PSF calibration image we have considered different cases. In the ideal case ("perfect" PSF calibration image) we assume perfect knowledge of the PSF, meaning that the PSF calibration image is a star-only image generated with the same PSF as the one used for the star+disk image. In other words, the PSF calibration image is just another realisation of the "true PSF" image. In this case the "PSF residuals" image simply gives us an estimation of the photon and read-out noise in the simulation. In the other cases the PSF calibration image is generated using a different PSF from the one that was used to convolve the star+disk image with. Here we use an analytical model of the original PSF, a second realisation (from the AO simulations) of the same PSF, or a slightly off-axis PSF. In these cases the "PSF residuals" image will not only contain the photon and read-out noise but also any differences between the PSFs used for the star+disk image and the calibration image.

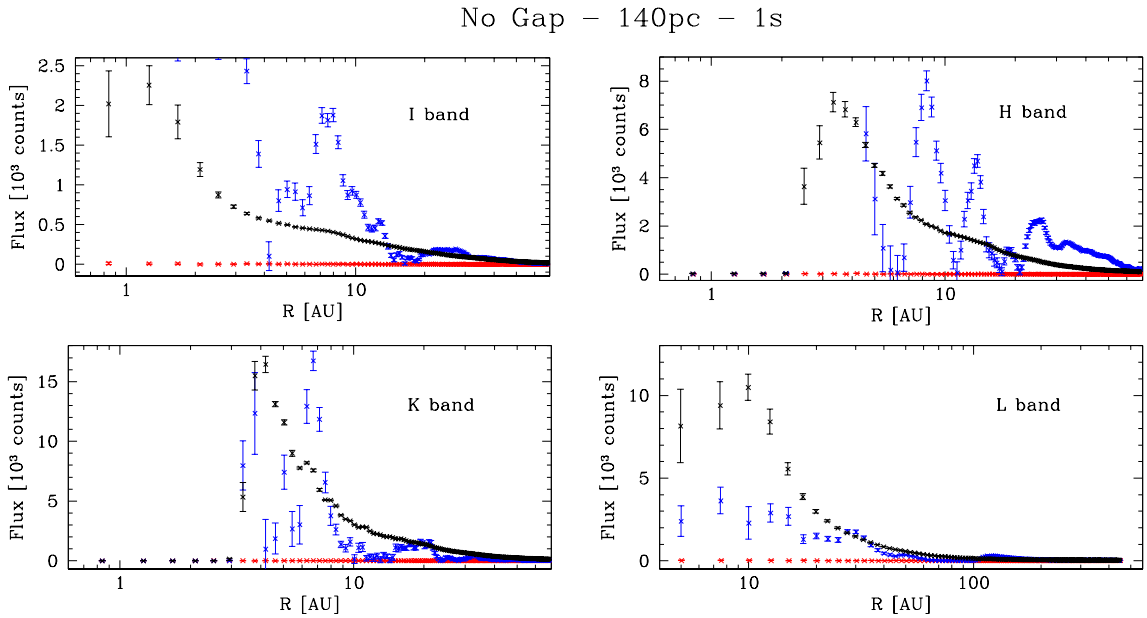

Figure 7.5: Azimuthally averaged radial intensity profiles of the final output of the simulations for a continuous disk at a distance of 140 pc in the indicated bands. In black we show the disk profile (star+disk image – "perfect" PSF calibration image). The red symbols show the residuals from subtracting the "perfect" PSF calibration image from the "true PSF" image (i.e. photon and read-out noise). The blue symbols correspond to the "PSF residuals" obtained from subtracting the "analytical" PSF calibration image from the "true PSF" image.

To compare the "disk-only" image with the "PSF residuals" image we generate azimuthally averaged light profiles of the images by running the IRAF task ELLIPSE that measures the image using an iterative method described by Jedrzejewski (1987). The flux level at a given radius is measured after fixing the centre of the object at the centre of the image and with a fixed ellipticity and inclination angle. The step size between successive ellipses is set to be 1 pixel. The uncertainty in the flux determination is derived as the square-root of the variance within a given annulus, divided by the square-root of the number of pixels in it.

7.5.3 Compliance with figures of merit

7.5.3.1 Detectability of the disk

As a first step in the simulation we want to understand to what extent we are able to detect the existence of circumstellar material surrounding the star. In order to quantify this we repeat our experiment of simulating the observations of a T Tauri star surrounded by a continuous low-mass disk for distances of 140, 400 and 1000 pc. These results are summarised in Figs. 7.5, 7.6 and 7.7.

Distance = 140 pc: This corresponds to the scale at which the circumstellar disk model was originally generated, so no scaling in terms of flux or pixel size (apart from the one required to include the instrumental effect) was applied in this case.

Fig. 7.5 summarises the results obtained for the I, H, K and L bands for a total integration time of 1 s divided into individual exposures of 0.1 s for the near-IR wavelengths and 0.01 s for the mid-IR. The N and Q bands were also simulated but the combination of high stellar flux, large pixel size and the very high contribution of the background at these wavelengths resulted in a fully saturated image not suitable for further analysis. The same output at N and Q bands was also observed in all the other realisations of the simulations. In Fig. 7.5, as well as in Figs. 7.6, and 7.7, we look to compare the disk-only flux to the residuals expected from two different kinds of PSF subtraction by showing three azimuthally averaged

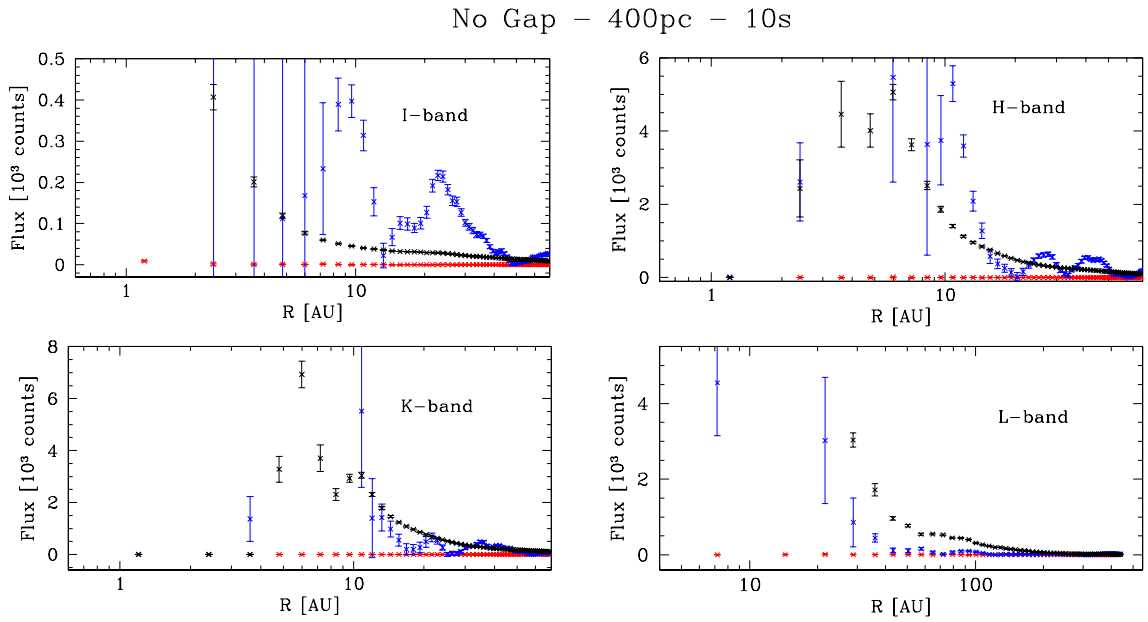

Figure 7.6: Azimuthally averaged radial intensity profiles of the final output of the simulations for a continuous disk at a distance of 400 pc in the indicated bands. The colour coding is the same used in Fig. 7.5.

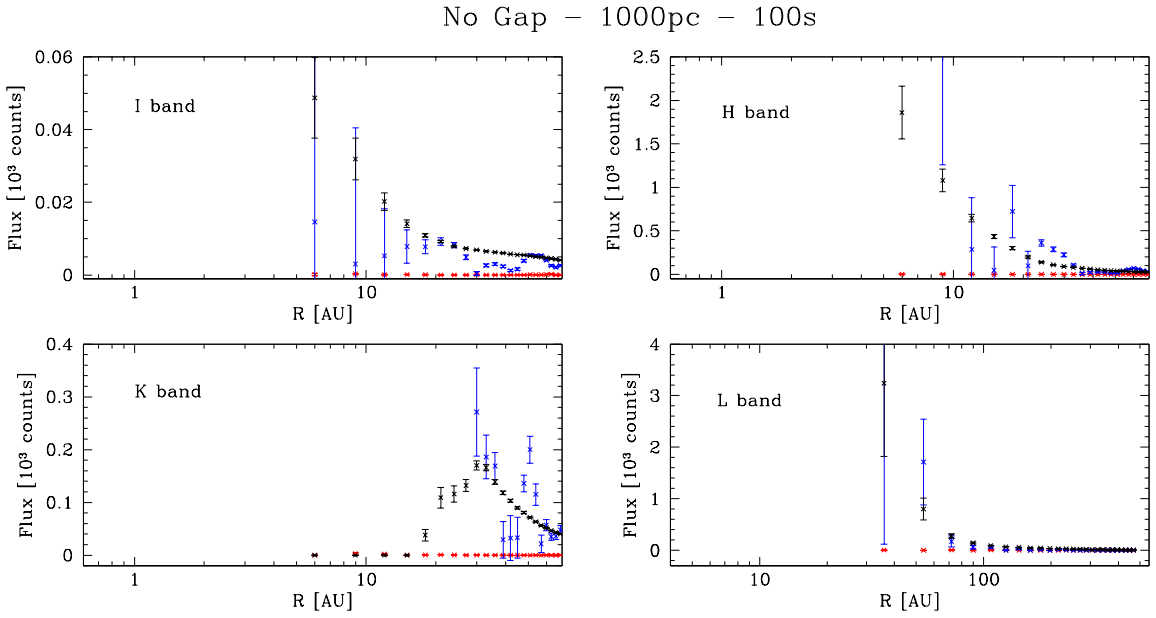

Figure 7.7: Azimuthally averaged radial intensity profiles of the final output of the simulations for a continuous disk at a distance of 1000 pc in the indicated bands. The colour coding is the same used in Fig. 7.5.

radial intensity profiles. The black symbols correspond to the output of the simulations for the star+disk model subtracted with a PSF calibration image generated with the same "original" PSF used to convolve the model in the first place. This profile is therefore showing the intensity pattern of the circumstellar disk. The red symbols represent the subtraction of the "true PSF" image and the "perfect" PSF calibration image, i.e. two realisations of the same thing, and is therefore just accounting for the photon and read-out noise. This give the detection limit we would expect from our near-perfect knowledge of the PSF. Finally, the blue symbols correspond to the subtraction of the "true PSF" image and the "analytical" PSF calibration image.

From Fig. 7.5 we can see how critically dependent the results are on an accurate knowledge of the PSF. The disk intensity profile (black) can be clearly distinguished from the photon and read-out noise (red) if we have a perfect knowledge of the PSF. When the procedure is repeated using a more realistic PSF description (blue symbols), the expected counts level coming from the disk is comparable to the uncertainty generated from the deficient PSF subtraction. The black and blue profiles are consistent within errors in the I and and H bands. It is worth noticing that the most promising results are observed in the K and L bands where the disk contribution to the total flux of the image is higher compared to the flux of the star. The better result in the L-band is also a result of the larger pixel size assumed at this wavelength, which tends to smooth-out the features in the PSF model.

In all the figures the rather odd behaviour of the profile at very small radius is explained by saturated pixels in these regions of very high flux due to the contribution of the stellar component.

Distance = 400 pc: At this distance as well as for the simulations at a distance of 1000 pc the images where scaled in flux to account for the effect of the increase in distance (lowering the total observed flux as $(140\ \mathrm{pc}/d)^2$) and the smaller apparent size of the objects in the sky (corresponding to an increase pc/pixel ratio of $d/140$ pc). The combinations of both of these effects results in only a relatively small change in the flux per pixels, so it is necessary to keep the minimum integration time per individual exposure in order to avoid saturation well beyond the central region. The total integration time was nonetheless increased to 10 s.

The results shown in Fig. 7.6 account for a disk that is better distinguished from the residuals of subtraction with the "analytic" PSF calibration image when compared to their counterparts at 140 pc. This is the result of looking at the more internal region of the PSFs and the smoothing effect in the profiles derived as a result of the distance scaling process. In all 4 studied cases we register a robust detection of the circumstellar emission when working with a close to perfect PSF knowledge, but a more uncertain detection when dealing with the "analytical" PSF calibration images.

Distance = 1000 pc: At a distance of 1000 pc and as shown in Fig. 7.7 the residuals from the PSF subtraction performed with the "analytical" calibration images are smoother than in the previous cases as a result of looking at the very inner region of the PSFs where the analytical models trace more closely the behaviour of the PSF. In this case we could expect that the circumstellar disk could be detected without the need of such a detailed knowledge of the PSF as in the previous cases. These simulations kept the single-exposure integration time used in the previously described realisations but increased the total exposure time to 100 s. The results show a much irregular shape in the detected circumstellar emission and given the smaller size of the target in the field of view we are only able to trace the more external regions of the disks.

Overall, the results presented in this section show that under the conditions assumed by the simulations the existence of a circumstellar disk with the physical properties described in Section 7.4.3.1 can be traced out to a distance of at least 1000 pc provided that a relatively good knowledge of the shape of the PSF can be obtained.

7.5.4 Gap detection

As a first step to prove to which extent a gap with the described physical properties could be detected given the observing conditions, we need to understand whether we would be able to resolve this or other similar size features given a certain set of observational parameters. Table 7.3 shows the characteristic

Table 7.3: Characteristic pixel sizes used in the simulations. (1) Passband, (2) original pixel size of the science and PSF images, (3) pixel size of the instrument assumed in the simulations, (4) width in pixels of a 2 AU-wide gap at a distance of 140 pc (first line), 400 pc (second line) and 1000 pc (third line), (5) equivalent in pixels of a 5 AU radius, (6) same as (5) but for a radius of 10 AU.

Band	PSF pixel size [mas/pix]	Inst. pixel size [mas/pix]	Gap width [pix]	Gap position 5 AU [pix]	Gap position 10 AU [pix]
(1)	(2)	(3)	(4)	(5)	(6)
			4.76	11.90	23.80
I	1.44	3	1.66	4.16	8.33
			0.66	1.66	3.33
			4.76	11.90	23.80
H	2.65	3	1.66	4.16	8.33
			0.66	1.66	3.33
			4.76	11.90	23.80
K	3.54	3	1.66	4.16	8.33
			0.66	1.66	3.33
			0.79	1.98	3.96
L	5.63	18	0.27	0.69	1.38
			0.11	0.27	0.55
			0.79	1.98	3.96
N	16.91	18	0.27	0.69	1.38
			0.11	0.27	0.55
			0.79	1.98	3.96
Q	32.37	18	0.27	0.69	1.38
			0.11	0.27	0.55

pixel size at the various steps of the simulations. In columns 4, 5 an 6 we have listed the equivalent in pixels of 2 AU, 5 AU and 10 AU for different pixel sizes and distances to the observer.

From the values shown in Table 7.3 we can immediately see that a 2 AU-wide gap could only be resolved at the minimum studied distance of 140 pc and, given the smaller pixel size, this would only be feasible in the near-IR wavelengths. We have in any case performed the whole set of simulations for the other distances/wavelengths hoping to detect a hint of the discontinuity in the circumstellar material, particularly in the near-IR bands at 400 pc, but no detection can rightfully be claimed from these results which are essentially equivalent to the corresponding profiles already shown in Figs. 7.5, 7.6 and 7.7.

Figs. 7.8 and 7.9 show the results of the simulations for a disk with a gap at 5 AU and 10 AU respectively. In both figures we show the results obtained for a gap with a Gaussian and a top-hat distribution of particles. The total exposure time is 1 s, divided in individual exposure of 0.1 s and 0.01 s for near-IR and mid-IR respectively. We have again performed one instance of the simulations assuming a perfect knowledge of the PSF and, in order to have a more accurate impression of how “real” these observing results are, we have also considered the profiles subtracted with the “analytical” PSF calibration image.

As already suggested by the continuous disk results, the measurements are highly dependent on the accuracy of our knowledge of the PSF of the system. From the disk profiles in the near-IR panels of Figs. 7.8 and 7.9 it is indeed possible to infer the existence of a gap in the circumstellar disk, especially from the simulations in the K-band. The gap detection is also more strong when the top-hat particle distribution is used. In the I and H bands, the contrast between the disk and the star is much less favourable and therefore any special features will be much more easily lost in what remains of the stellar light after the PSF subtraction. At the L-band the much larger pixel size makes it impossible to detect the planet formation features, as they appear located too close to the star (only 4 pixel away for the gap located at 10 AU), in the region of saturated pixels.

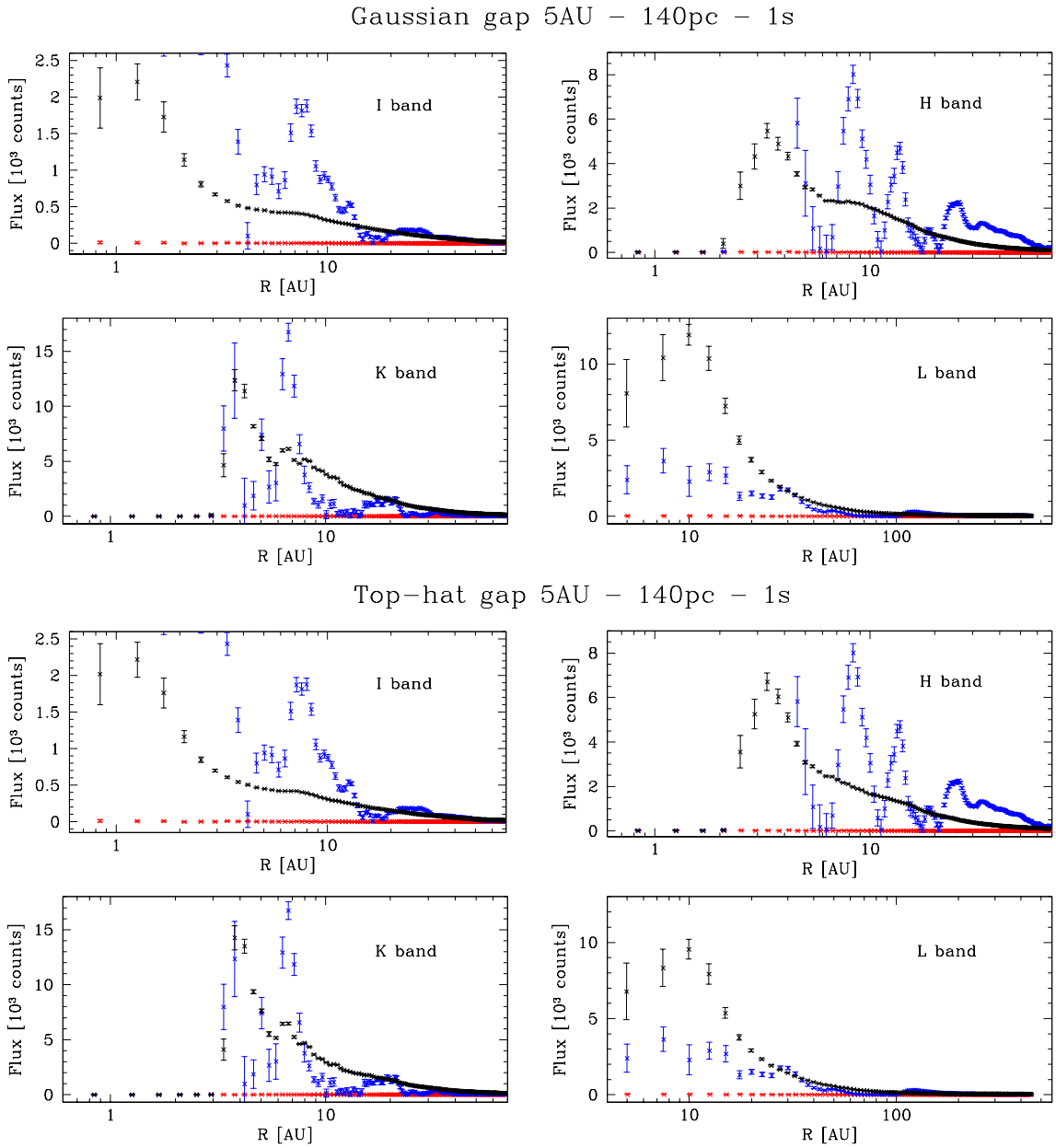

Figure 7.8: Azimuthally averaged radial intensity profiles of the final output of the simulations for a disk having a 2 AU-wide gap at a radius of 5 AU. The assumed distance to the object is 140 pc and the particle distribution in the gap is Gaussian (top) and top-hat (bottom). The colour coding is the same used in Fig. 7.5.

Figs. 7.8 and 7.9 also give us a hint of how a poor PSF subtraction could affect the results. In this case, the “analytical” PSF calibration image differs quite substantially from the “true PSF” image and would completely erase the features in the profile indicating the existence of a gap, both at 5 AU and 10 AU. On a real observation any trace of the planetary formation features that could remain after subtracting the stellar light will be attributed to artifacts introduced by a deficient PSF subtraction.

Overall the results presented here suggest again that the disk itself could be detected, but to secure a “gap detection” we will need to be able to provide an extremely good description of the PSF of the system. As shown in Figs. 7.8 and 7.9 even the introduction of a small uncertainty, would decrease the

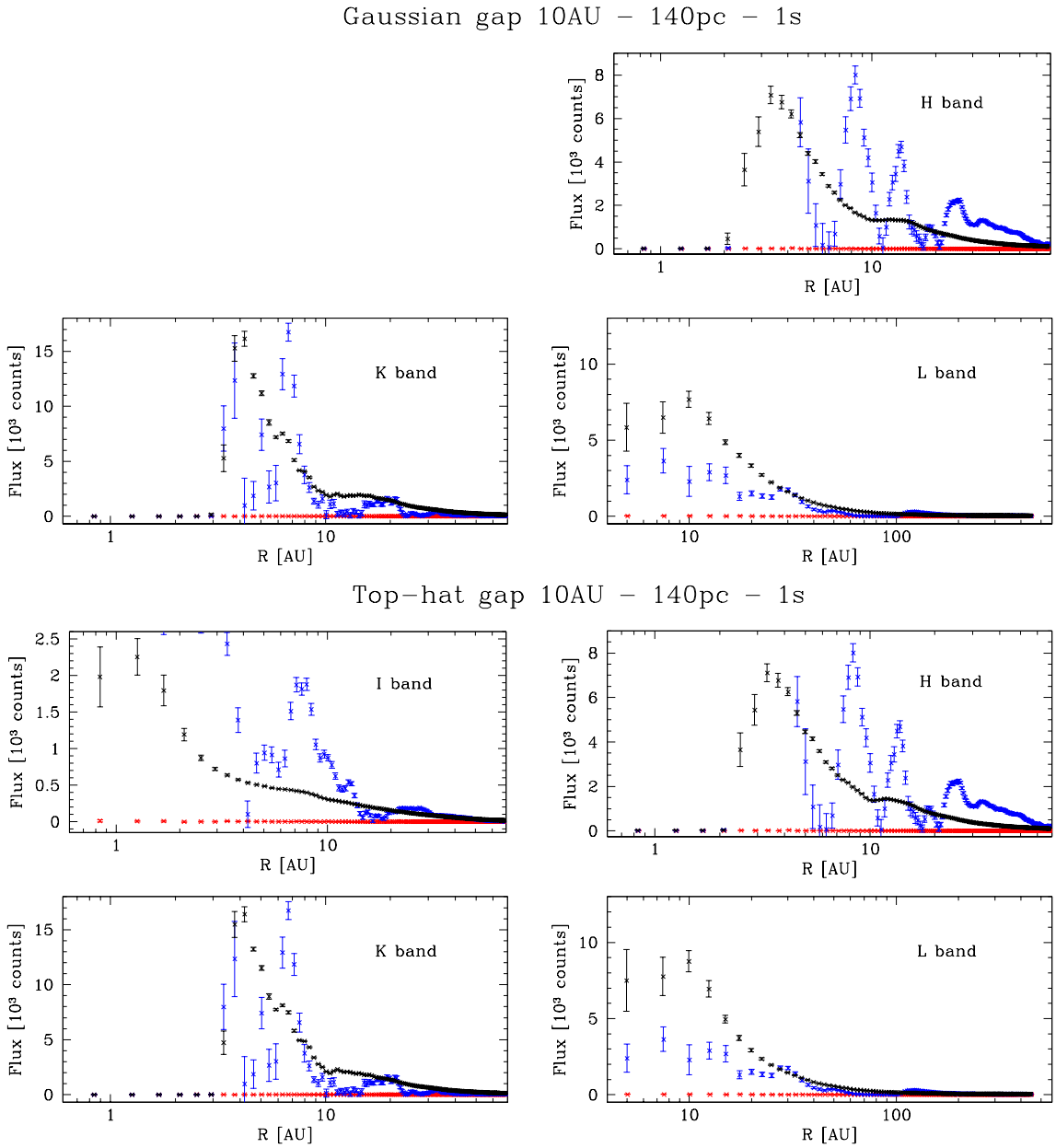

Figure 7.9: Azimuthally averaged radial intensity profiles of the final output of the simulations for a disk having a 2 AU-wide gap at a radius of 10 AU. The assumed distance to the object is 140 pc and the particle distribution in the gap is Gaussian (top) and top-hat (bottom). The colour coding is the same used in Fig. 7.5.

smoothness of the profile to such an extent that the gap could not be differentiated from other artifacts introduced by the faulty PSF subtraction.

As described in Section 7.4.3.2, we have also performed the simulations using PSF calibration images generated from a different AO realisation of the PSF and with an off-axis PSF. Given the limited availability of PSF images these simulations are only possible in the H and K bands.

The results of the simulations performed using a different AO realisation of the PSF for the calibration image are presented in Figs. 7.10 and 7.11. In these figures the black symbols correspond to the azimuthally averaged radial intensity profile obtained for the star+disk image from which the “different realisation” PSF calibration image has been subtracted. Equivalently, the red symbols show the residuals from the sub-

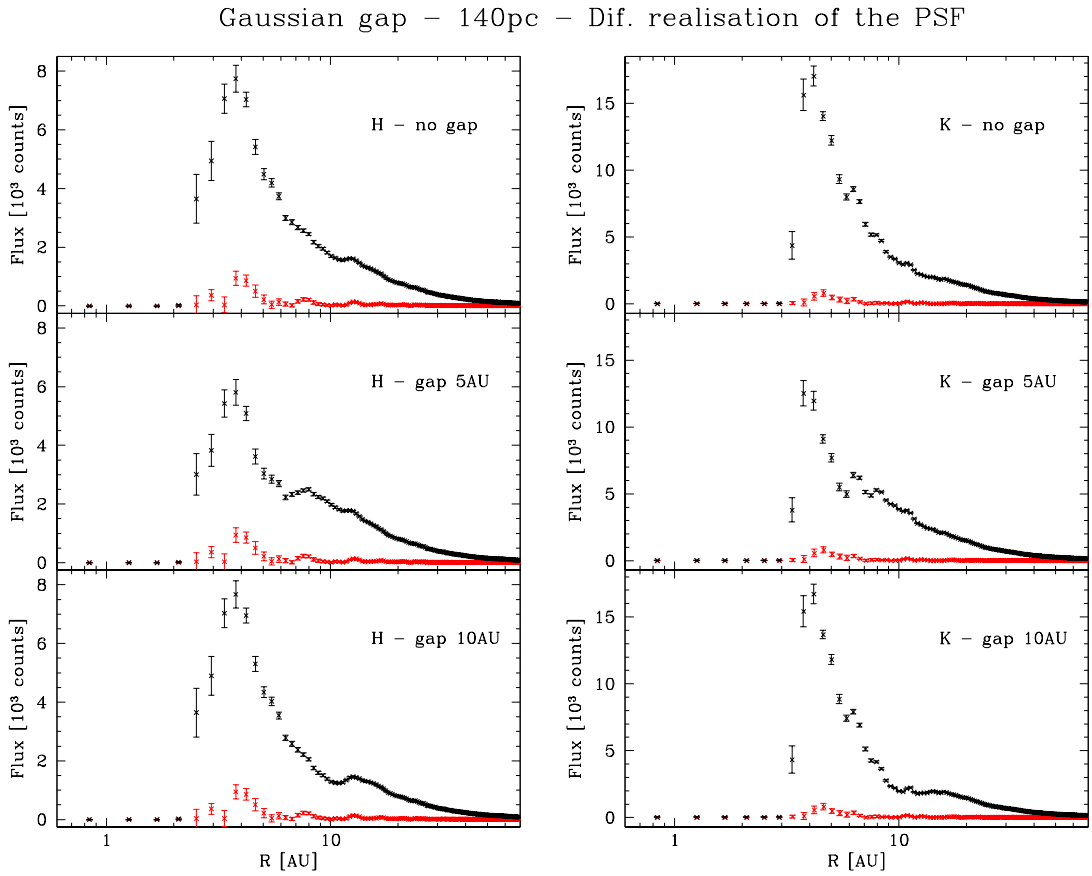

Figure 7.10: Azimuthally averaged radial intensity profiles of the final output of the simulations for a continuous disk (top), a disk with a Gaussian gap at 5 AU (middle), and a disk with a Gaussian gap at 10 AU (bottom), in the H (left) and K (right) bands. The black symbols correspond to the residuals obtained from subtracting the "different realisation" PSF calibration image from the star+disk model. The red symbols corresponds to the subtraction of the "different realisation" PSF calibration image from the "true PSF" image.

traction between the "true PSF" image and the "different realisation" PSF calibration image. In this case we are performing the PSF subtraction using a calibration PSF that is very similar in its shape to the one used for the star+disk image, but was simulated for a different realisation of the atmospheric conditions and has an additional random component due to the noise associated with the guide star images.

From Figs. 7.10 and 7.11 we can see that the PSF subtraction using this PSF calibration image is very successful, and that it allows to detect without any doubt the presence of circumstellar material in the images as well as the planetary formation features in it. The results are almost equally good for both particles distributions, and in both passbands. These results show a good example of the level of accuracy needed in order to detect features in the circumstellar disk, which as already suggested by Fig. 7.4 needs to be within a 10 % of precision in the K band.

Figs. 7.12 and 7.13 present the results of the subtraction performed with a 10″ off-axis PSF. Here the black symbols correspond to the star+disk image from which the "off-axis" PSF calibration image has been subtracted, while the red symbols show the residuals of the subtraction between the "true PSF" image and the "off-axis" PSF calibration image. In this case we can observe a much less precise PSF subtraction, because of the intrinsic differences between the two PSF images, in particular the differences in ellipticity. This level of accuracy on the PSF subtraction does not allow for a reliable gap detection, but delivers nonetheless reasonable indications of the presence of circumstellar material, particularly in the K-band.

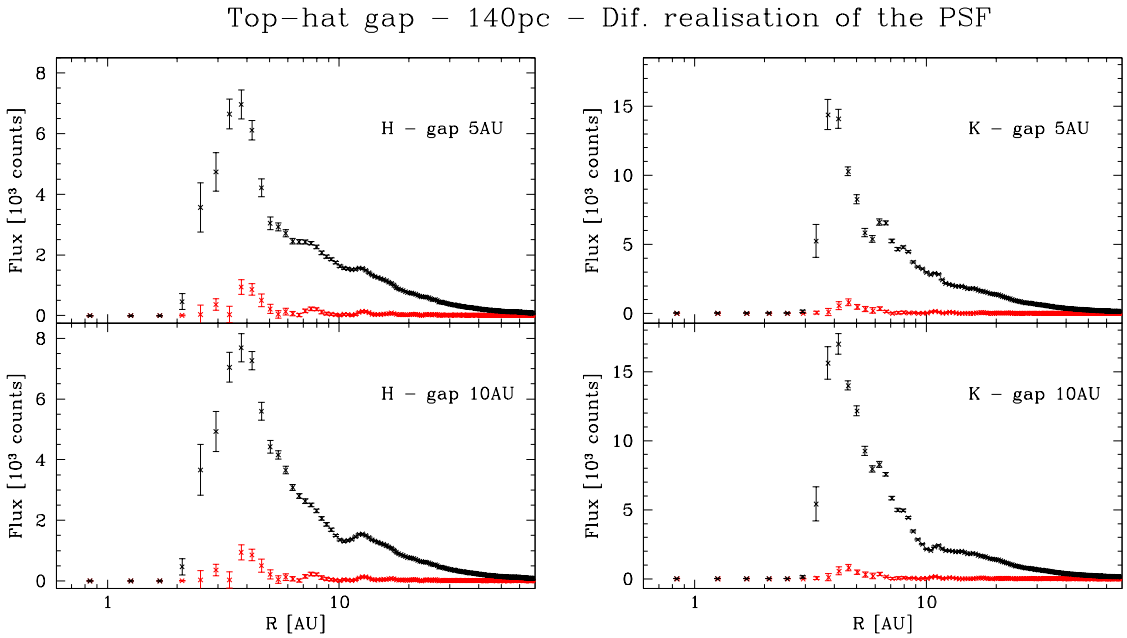

Figure 7.11: As Fig. 7.10 for a top-hat gap.

7.5.5 Sensitivity to input parameters

Besides the already mentioned critical dependency of these simulations on the modelling of the PSF calibration, there are no major variations in the results as a function of the input parameters. Below we present a few comments on the results obtained when varying some of them.

Site: As discussed before the only effect the site has on the simulations is reflected in the level of the background. This difference mainly affects the mid-IR wavelengths and therefore only the L-band in our case, with a decrease of $\sim$1 mag in a High & Dry site compared to a Paranal-like site. These simulations are mostly concerned with the study of the very central region of the star+disk system, where compared to the contribution of the stellar light the background is negligible. A lower background could nonetheless play an important role in the study of the external regions of the disks. The lower background achievable for the N and Q bands at a High & Dry site is also not enough to avoid extended saturation on these images.

Mirror coating: All the results presented above were computed using the transmission curves of a protected silver/aluminium coating. The results of the simulations were also tested using the slightly inferior bare aluminium coating, which produced a relatively minor decrease in the total flux observed at all wavelengths. As expected the major decrease in flux is observed in the I-band. These flux changes have nonetheless no effect on the final conclusions of this study. Note that the bare aluminium coating also increases the background in the K-band and at longer wavelengths, and so the comments made above also apply here.

Pixel size: For most of the wavelengths studied the assumed instrumental pixel size was relatively close to the original pixel size of the images, which is determined by the size of the diffraction limited core of the PSF. We expect then no major information lost due to this assumption.

The only exception to this rule would be the L-band, that goes from 5.63 mas/pix to a more than 3 times larger 18 mas/pix. The availability of a smaller pixel size at this wavelength will not improve much the situation in terms of the detection of the gap (2 AU corresponds to only 2.5 pixel at a pixel scale of 5.63 mas/pix), but could certainly make a contribution to the study of the inner region of the disk, which becomes automatically out of reach with a pixel scale of 18 mas/pix.

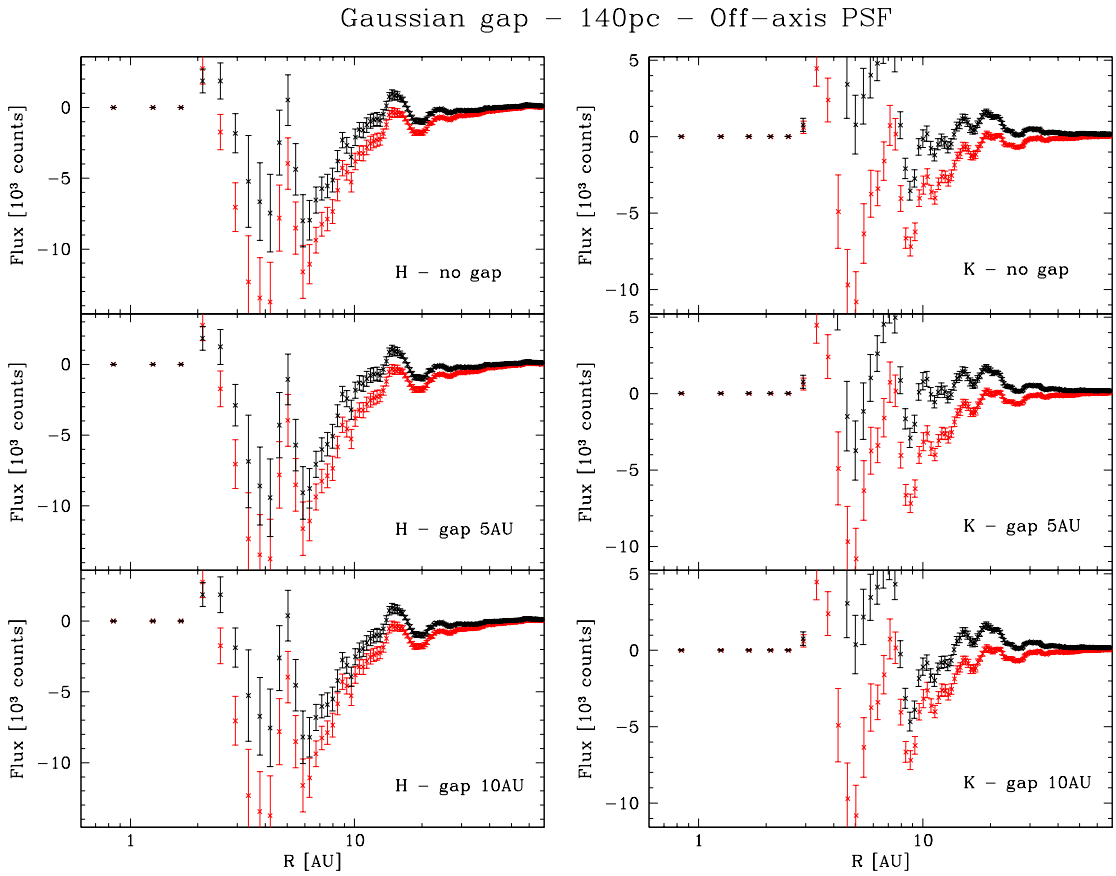

Figure 7.12: Azimuthally averaged radial intensity profiles of the final output of the simulations for a continuous disk (top), a disk with a Gaussian gap at 5 AU (middle), and a disk with a Gaussian gap at 10 AU (bottom), in the H (left) and K (right) bands. The colour coding is the same used in Fig. 7.10, but in this case we are subtracting the "off-axis" PSF calibration image.

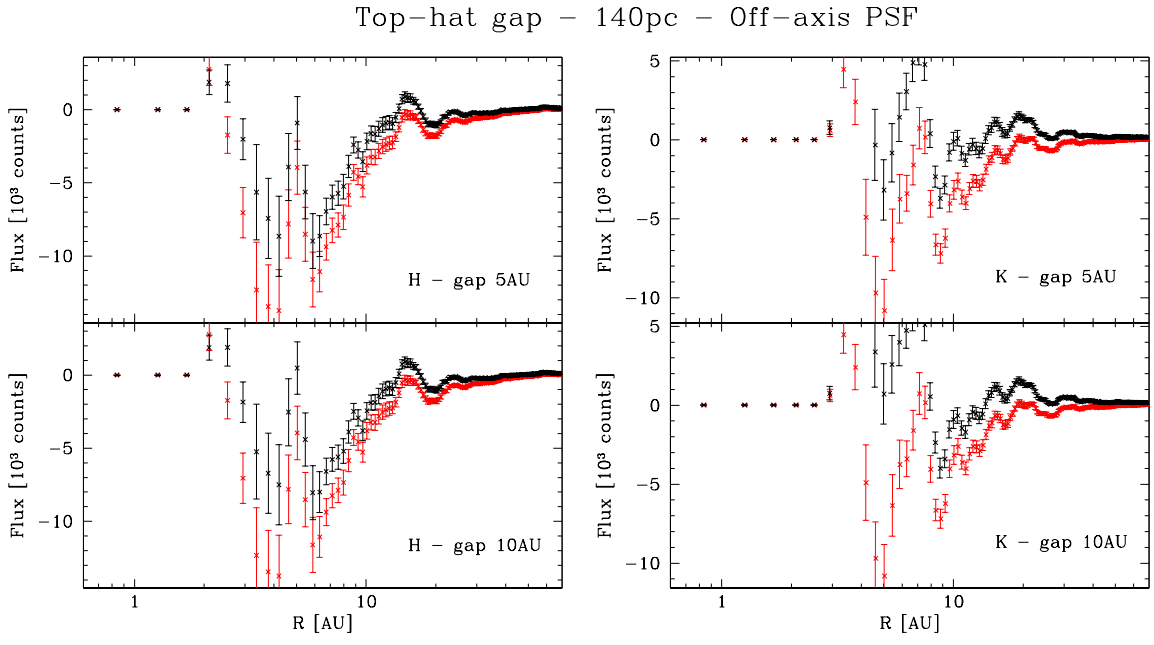

Figure 7.13: As Fig. 7.12 for a top-hat gap.

7.5.6 Calibration requirements

As discussed before the main calibration issue related to this science case is the accurate determination of the PSF of the system. Independently of the strategy for the analysis of the data, the PSF must be precisely measured, either to permit the subtraction of the central source as we did here and/or to allow for the deconvolution of the data.

In this case, PSF-reference stars close to the target will need to be selected and periodic PSF measurements planned during the observation. Ideally, the target and the PSF reference star magnitudes should be similar to ensure consistent AO performance.

7.5.7 Limitations

There are several limitations and possible improvements required for the simulation procedure adopted here. Among them the following should be considered:

- Our simulations make use of a theoretical model to reproduce the main observational properties of a circumstellar disk and the potential planet formation features in it. Even though we believe our model is a fairly accurate approximation of what would be observed in reality, there are always uncertainties associated with this kind of procedure. In particular, the model for the distribution of particles in the region of the gap has not been subject of much testing and might indeed not be very realistic. It is expected that as a planet is formed and interacts with the surrounding circumstellar region a ring of material will be “cleaned” along the planet’s orbit. This might generate a Gaussian profile in the distribution of particles as assumed in our simulation, but this stage might in practice be too short to have a high chance of observability. We have also simulated a more realistic scenario using a top-hat particles distribution, i.e. assuming that the gap has been already fully cleaned by the planet. Our results show that the chances of detecting the gap might be slightly higher with a top-hat distribution, but in general this feature does not seem to be very relevant in the final results.
- In the simulations we are assuming a nearly perfect data reduction process which might not be the case in a real analysis process.
- An alternative kind of analysis of the data might include performing a PSF deconvolution, but this technique basically shares the same major limitation we have encountered here, i.e. it requires a very good knowledge of the PSF of the system.
- At the two shortest wavelengths (I and H-band) our simulations could profit from the use of a coronagraph. Given the expected pixel size of the instrument at these wavelengths it might be possible to place a coronagraph with a small enough focal plane mask to avoid masking the region of the gap, which would be located at 10 AU = 23.8 pixel in the best case scenario. In this case it will be necessary to have a good understanding of the light transmission function in the transition region surrounding the coronagraph as well as a good account of the optical aberrations present in the entrance pupil, which might cause the stellar light to scatter.
- The simulations are limited by the number of PSF models that we are able to generate. In that sense we are simulating observations in the wavelengths for which we already have PSF models in hand, even though they might not be the most appropriate for this science case. For example, the use of a narrow-band filter at the wavelength of the PAH feature at 3.3 μm might provide better contrast between the star and the disk, and therefore enhance the chances of being able to detect a gap in the disk.

7.6 Concluding remarks

The set of simulations presented in this report aim to trace the detectability of planetary formation features in young low-mass circumstellar disk located in nearby star forming regions. Our working case consists of a T Tauri star surrounded by a face-on circumstellar disk with a certain set of physical properties described

in Section 7.4.3.1. The planet formation features are modelled by introducing a Gaussian or top-hat shaped gap in the continuity of the particles in the disk. In this framework the simulation consisted of testing the detectability of the disk itself as a function of the distance to the observer, and the detectability of the gap at two different radii from the parent star: 5 AU and 10 AU. The simulations were performed in 3 near-IR bands (I, H and K) and 3 mid-IR band (L, N and Q).

The results shown here suggest that given the expected observing conditions and high spatial resolution to be reached with a fully-adaptive telescope like the E-ELT we would be able to trace the existence of the modelled low-mass circumstellar disk out to a distance of at least 1000 pc, a priori without the need to use a coronagraph.

The detection of the planetary formation features produced by a 1 $M_{\rm Jup}$ planet is, on the other hand, a much more challenging task. The results of the simulations suggest that in order to detect these features a nearly perfect knowledge of the PSF of the system would be required. Our results indicate that the level of accuracy needed to detect the modelled gap of particles is within 10 %. We therefore expect to be able to trace such a fine structure only if precise PSF calibration procedures can be developed. The chances of detection are certainly more promising for wider gaps located at larger radii from their parent star.

In terms of the telescope instrumentation, this science case sets only one significant requirement. Within the current set of instrument studies there is only one foreseen imager for the mid-IR wavelength range. The current specifications in terms of pixel scale for this instrument in the L-band are not enough to profit from the advantages of this wavelength range in the study of circumstellar disks. A smaller pixel scale, closer to the scale determined by the size of the diffraction limited core of the PSF, will certainly make a contribution to the study of the inner regions of circumstellar disks.

References

Beckwith S.V.W., Sargent A.I., Chini R.S., Guesten R., 1990, AJ, 99, 924
Beckwith S.V.W., Sargent A.I., 1996, Nature, 383, 139
Bodenheimer P., Pollack J.B., 1986, Icarus, 67, 391
Burrows C.J., et al., 1996, ApJ, 473, 437
Coffey D., Bacciotti F., Woitas J., Ray T.P., Eislöffel J., 2004, ApJ, 604, 758
Hartmann L., Cassen P., Kenyon S.J., 1997, ApJ, 475, 770
Hogerheijde M.R., van der Tak F.F.S., 2000, A&A, 362, 697
Grazian A., Fontana A., De Santis C., Gallozzi S., Giallongo E., Di Pangrazio F., 2004, PASP, 116, 750
Jedrzejewski R.I., 1987, MNRAS, 226, 747
Kalas P., Liu M.C., Matthews B.C., 2004, Science, 303, 1990
Kalas P., Graham J.R., Clampin M.C., 2005, Nature, 435, 1067
Kalas P., Graham J.R., Clampin M.C., Fitzgerald M.P., 2006, ApJ, 637, 57
McCaughrean M.J., O'dell C.R., 1996, AJ, 111, 1977
Meyer M.R., Backman D.E., Weinberger A.J., Wyatt M.C., 2007, Protostars & Planets V, 573
Pinte C., Ménard F., Duchêne G., Bastien P., 2006, A&A, 459, 797
Schneider G., Smith B.A., Becklin E.E., Koerner D.W., Meier R., Hines D.C., et al., 1999, ApJ, 513, 127
Strom S.E., Edwards S., Skrutskie M.F., 1993, Protostars & Planets III, 837
Watson A.M., Stapelfeldt K.R., Wood K., Ménard F., 2007, Protostars & Planets V, 523
Wyatt M.C., 2006, ApJ, 639, 1153
Weinberger A.J., Becklin E.E., Schneider G., Smith B.A., et al., 1999, ApJ, 525, L53

8 S5-1: Characterising the lowest mass freely floating objects in star forming regions

Authors: A. Calamida, F. Comerón, H. Zinnecker

8.1 The science case

This proposal aims at the spectroscopic characterisation of the lowest mass freely floating objects selected in wide-area imaging surveys of young star forming regions, carried out with either survey telescopes or large format infrared imagers at 8-m telescopes. It is likely that the physical properties of such objects, as little massive as one or a few Jupiters, can be derived only from the comparison of spectra covering diagnostic atomic and molecular lines to the next generation of ultracool atmospheres. Therefore, the determination of the shape of the lowest-mass end of the sub-stellar initial mass function (IMF) may be only possible by using relatively high signal-to-noise spectroscopy that can only be provided by ELTs. Besides, spectroscopic monitoring will yield information on the meteorology of these objects, which are expected to display complex weather patterns. The goal of this project is to obtain near-simultaneous spectroscopy of the lowest-mass objects in star forming regions, in the wavelength interval ranging from the red (0.6 μm) spectral region dominated by the opacity in alkali lines, to the near-infrared (NIR, 4 μm) dominated by molecular opacities (H_2, H_2O, CH_4, NH_3).

By the time the E-ELT enters operations it is likely that complete IMFs down to giant-planet masses will be available for at least some of the nearest star forming regions thanks to the next generation of infrared wide-field surveys to be carried out with VISTA, by pencil-beam deep observations with instruments like HAWK-I, and by other surveys of specific regions already ongoing (Caballero et al. 2007). Those imaging surveys may be able to show the existence of a cutoff at the lower end of the IMF of freely floating objects and to check theoretical predictions against actual statistics, but their quantitative results will be limited by the lack of spectroscopic information allowing accurate determinations of their physical properties.

The faint end of the sub-stellar mass range represents an as yet uncharted region of the parameter space populated by objects with low temperatures (T < 1000 K) and low surface gravities ($\log g$ < 3, Baraffe et al. 2003, hereafter BA03). However, objects with similar temperatures, classified as late T dwarfs, are already known (Geballe et al. 2001; Saumon et al. 2007, and references therein) and have been theoretically modelled (Burrows et al. 2001; Marley et al. 2002; Burrows et al. 2003). These are typically evolved brown dwarfs (BDs) found in the field, or in orbit around higher mass companions. Their masses are a few tens of a Jupiter mass, and they have had time during their evolution to cool down to the low temperatures observed now. The atmospheres of such objects are rich in chemistry, mineralogy and meteorology, truly bridging the gap between the domain of stellar and planetary atmospheres. The formation of molecules and grains at such low temperatures is coupled to large-scale convection extending well into the atmosphere (Helling et al. 2008), sedimentation of grains below the photosphere, raindown of chemical species, and the formation of clouds and of large-scale weather patterns (Burgasser et al. 2002). Observationally, this means that the spectral features of BD atmospheres must display time variability. Photometric monitoring of T dwarfs has indeed revealed so far that aperiodic brightness variations with amplitudes of up to 0.3 mag and characteristic timescales ranging from several hours to a few days are frequent (Enoch et al. 2003; Artigau et al. 2003). This variability has been attributed to the formation and evolution of cloud decks.

The main difference between the already well known evolved T dwarfs and objects with similar temperatures in a star forming region concerns the mass and the surface gravity. Assuming that such objects do form at all, BA03 models predict that a T = 1000 K object in a 5 Myr young star forming region should have a mass M = 3 $M_{\rm Jup}$ and a surface gravity $\log g$ = 3.6, to be compared to $M = 55\ M_{\rm Jup}$ and $\log g$ = 5.3 for a 5 Gyr-old T dwarf of the same temperature. Observations of a handful of T dwarf-like objects in star forming regions (e.g. Zapatero Osorio et al. 2002) indicate colours and spectral features that are similar, but not identical, to those of field T dwarfs, whereas recent atmosphere models clearly identify gravity-dependent spectral features (Burrows et al. 2006). On the other hand, an intriguing possibility currently under theoretical investigation is that the spectrum may contain as well telltale signatures of the

formation mechanism. Such signatures would be the consequence of the differences in composition and internal structure between those objects formed in isolation by collapse and fragmentation, and those formed in an accretion disk around a star and subsequently ejected into interstellar space.

While they will be within the detection limits of imaging surveys, spectroscopy of these objects will require ELT capabilities due to their faintness even in the nearest star forming regions at a distance of $\sim$ 150 pc. Expected apparent magnitudes for a 5 Myr object of 1 $M_{\rm Jup}$ ($T_{\rm eff}$ $\sim$ 640 K) are I = 26.3, J = 22.8, H = 22.6, K = 23.0, L' = 19.6 (BA03). Interestingly, these same models predict a luminosity-temperature relation that is virtually independent of age over the entire range from 1 Myr up to 10 Gyr. The existence of such *pseudo-main sequence* provides a simple way to estimate the contamination of this region of the colour-magnitude diagram (CMD) of young aggregates due to non-members: objects with the same colours but brighter apparent magnitudes provide a direct measurement of the volume density of evolved objects of any age in the same colour range between us and the aggregate. One should note however that luminosity estimates at such young ages are at present extremely uncertain and may fall considerably below those indicative values, as recently shown by Marley et al. (2007).

Spectroscopy at moderate resolution is needed in order to analyse the composition of dust clouds that form in the atmospheres of these objects. The main spectral regions of interest are the J, H, K and L' bands, where the dominant sources of opacity (H_2O, NH_3, CH_4) appear, and also the far red (0.6–1.0 μm), where lines from neutral alkali elements (Li, Na, K, Cs, Rb) dominate the flux depression shortwards of 1 μm (Burrows et al. 2000) and together with refractory elements (Al, Ca, Cr, etc.) provide important diagnostic on temperature and pressure (Lodders 1999) that can be extremely useful to constrain the evolutionary tracks where the objects lie. Distances to nearby star forming regions, which currently are the dominant source of uncertainty in the observational determination of luminosities, will be known to an accuracy better than $\sim$1% by the time the E-ELT enters operations thanks to the results provided by GAIA. Time monitoring of ultracool objects will be important, as it will enable the observation of global weather pattern evolution and the development and disappearance of local atmospheric features, thus providing information on the atmospheric circulation, the timescale for the formation of local instabilities, or the role of rotation. The changing cloud coverage in these objects, the complex chemistry in their atmospheres, and the rotational modulation probably makes their spectrum evolve continuously, thus requiring integration times limited to a few hours in order to obtain meaningful snapshots. We expect rotation to be the main factor in limiting the duration of exposures, based on recent results by Zapatero Osorio et al. (2006) that show that field T dwarfs have rotation periods of 12.5 hours or less and that the spin-down timescale at such low masses is very slow. Furthermore, with the advances in the modelling of the atmospheres and interiors of these objects expected in the coming decade, it may be foreseen that the spectroscopic characterisation will be an even more valuable tool than it is today to relate the properties of ultracool young objects to their masses, thus becoming essential to determine the shape of the IMF at its lowest end.

8.2 Goals of the DRM simulations

This proposal was designed to explore the performance of the E-ELT under non-optimal atmospheric conditions. The assumption is that a programme of these characteristics will be executed in a survey-like manner at times when the conditions do not allow observations with the good adaptive optics performance required by most other science cases. It is therefore proposed that the simulations use as input atmosphere parameters typical or poor conditions and assume only the Ground Layer Adaptive Optics (GLAO) correction provided by the telescope.

The main questions that the simulations should address are:

- Given an exposure time limited to two hours by the object rotation period, what are the limiting masses that can be reached at the nearest star forming regions (Chamaeleon, Lupus, Ophiuchus, R CrA)?
- Given the expected E-ELT performance in the infrared and the colours of targets, what is the best spectral region for the spectroscopic study of these objects?

8.3 Metrics / figures of merit

The output should be the recovered spectrum for objects of different magnitudes at a reference band. There is no need to carry out simulations for objects with magnitude fainter than $H \sim 23$ mag, as it is unlikely that surveys with smaller telescopes will provide candidates much fainter than this magnitude. The recovered spectra will then be compared with the input spectra in order to determine the magnitude limit down to which meaningful measurements of key features in different parts of the spectrum can still be made. Comparisons will also be made to template spectra of different spectral types so as to determine the accuracy of the overall spectral classification as a function of the magnitude.

Based on these results, a comparison to theoretical evolutionary tracks will tell us the mass down to which the E-ELT can provide a reliable characterisation of the objects that we are considering, given typical parameters of nearby low-mass star forming regions (distance = 150 pc, age = 2 Myr).

8.4 DRM simulations

8.4.1 Methodology

In order to take into account the points of Sections 8.2 and 8.3, we first simulate spectra for the four proposed fiducial cases listed in Table 8.1, i.e. objects with temperatures in the range $400 < T_{\rm eff} < 2000$ K and masses $M \leq 10\ M_{\rm Jup}$. As a second step, we produce spectra for objects with similar temperatures but different surface gravities, in order to estimate the accuracy of the overall spectral classification as a function of magnitude. In the mid-infrared (MIR) case we also test different telescope sites, simulating spectra for a Paranal-like and a High & Dry site (see Section 8.4.3.2).

After generating the spectra in the I, J, H, K and L' bands we compare them to the input ones. Moreover, we compare spectra with the same temperature but different surface gravities, in order to estimate the lowest mass at which we can distinguish the two objects.

8.4.2 Pipeline

We need to simulate spectra of the lowest mass freely floating objects in star forming regions, as massive as one or a few Jupiters, assuming different input parameters. At first, we simulate the spectra for the four fiducial cases listed in Table 8.1. The template spectra are taken from F. Allard's AMES-Cond ($T_{\rm eff} < 1400$ K) and AMES-Dusty ($T_{\rm eff} > 1700$ K) model grids (Allard et al. 2000, see Section 8.4.3 for more details), while masses and apparent magnitudes are derived adopting a BA03 isochrone for t = 5 Myr and adding a distance modulus of μ_0 = 6.0 mag. The parameters of the fiducial cases are listed in Table 8.1.

Table 8.1: Parameters of the four fiducial cases.

Mass $M_{\rm Jup}$	$T_{\rm eff}$ K	$\log g$	m_I mag	m_J mag	m_H mag	m_K mag	m_L mag
Fiducial cases							
0.5	455	2.790	28.25	25.37	24.91	26.96	21.03
1	644	3.141	26.34	22.84	22.57	22.97	19.56
3	1,098	3.576	23.54	19.89	19.56	16.21	17.21
10	1,965	3.921	19.61	16.80	16.44	16.00	15.06
Comparison cases							
20	389	4.802	...	...	...	...	...
30	610	5.011	...	...	...	...	...
20	936	4.704	...	...	...	...	...
30	1979	4.715	...	...	...	...	...

Table 8.2: Parameters of simulated spectra in the different optical, NIR and MIR bands.

Band	Model grid	$\delta\lambda$ μm	$R^a_{\rm ini}$	$R^b_{\rm fin}$	$t_{\rm exp}$ s
I	Cond/Dusty	0.6–1.2	4500	4500	7200
J	Cond/Dusty	0.9–1.5	6000	4500	7200
H	Cond/Dusty	1.3–1.9	8000	4500	7200
K	Cond/Dusty	1.9–2.5	11 000	4500	7200
L'	Cond	3.5–4.1	7600	4500	7200
L'	Dusty	3.5–4.1	3800	3800	7200

[a]Spectral resolution synthetic input spectrum.
[b]Spectral resolution of the simulated spectrum.

Each synthetic spectrum is sorted by wavelength and the region of interest is selected in each band (I, J, H, K, L'). Table 8.2 lists the wavelength region, the initial spectral resolution of the synthetic input spectrum, and the final resolution of the simulated spectra for the different adopted bands and model grids. The final resolution is the same for all the spectra, i.e. $R = 4500$, except for the L' dusty spectrum, where $R = 3800$. These values are suitable for the measurement of the spectral features of interest, and the exposure time is fixed at t = 7200 s (see Section 8.2). We adopt the data for the atmosphere and telescope transmission, the atmospheric and telescope thermal emission, the sky continuum and the emission lines available in [RD1] for different mirror coatings (bare aluminium and protected silver/aluminium) and telescope sites (Paranal-like and High & Dry), and we smooth them to the same spectral resolution of the simulated spectra ($R = 4500$ or 3800).

We assume an integral field unit (IFU) instrument and we simulate the spectrum in the central spatial resolution element (spaxel), fixed at 40 mas for the optical-NIR cases, and 20 mas for the MIR case (see Section 8.4.3.2 for more details). Given that we are using only the telescope GLAO correction, most of the object flux falls indeed outside the central spaxel. In the simulations, therefore, we rescale the object flux for the fraction of ensquared energy (EE) that falls within the assumed central spaxel. The EE values adopted are listed in Table 8.3 and have been estimated interpolating the values given in [RD1]. Since we are considering only freely floating objects, we can assume that they are isolated, i.e. there is no contamination by the flux of neighbouring stars. In order to increase the signal-to-noise ratio (S/N), we sum the spectrum of the central spaxel with the spectra of the 8 surrounding spaxels. The fraction of the energy enclosed in this area is listed in the last column of Table 8.3. There is no substantial improvement in adding the flux of more surrounding spaxels (we tested the case of summing the spectra of 25 spaxels).

In order to generate the spectra we used a modified version of a code originally developed by G. Battaglia (see Section 12), which was adapted to our specific case.

Table 8.3: Some input parameters used for the simulation of the optical-NIR and MIR spectra (taken from [RD1]).

Band	$\lambda_{\rm eff}$ μm	Half-width μm	$2.5 \log f_{\lambda_0}$ erg/s/cm^2/Å	EE (central spx)[a] %	EE (total)[b] %
I	0.88	0.24	-22.69	0.97	7.7
J	1.22	0.213	-23.76	1.44	10.5
H	1.63	0.307	-24.86	2.15	14.7
K	2.19	0.39	-26.01	3.83	21.2
L'	3.80	0.65	-28.28	4.9	17

[a]Ensquared energy within the central spaxel.
[b]Ensquared energy within the central 9 spaxel.

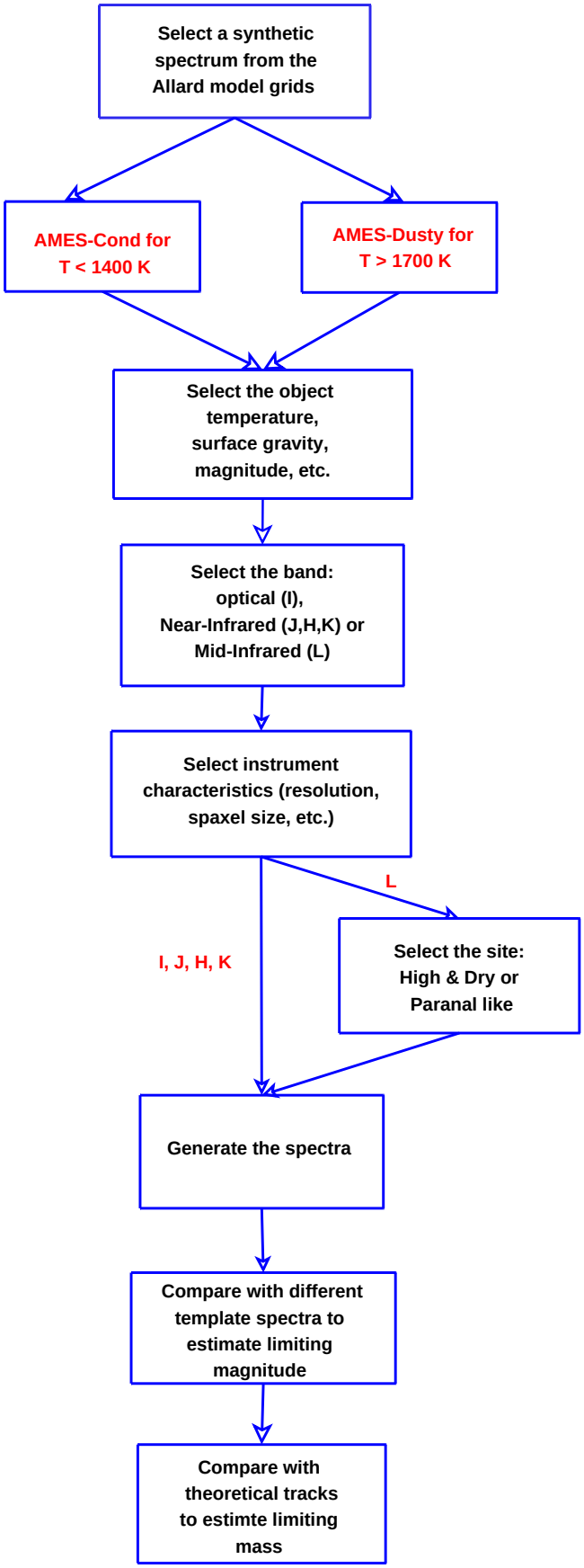

Figure 8.1: Flowchart of the simulations.

First of all, the code starts from the template spectrum and normalises the flux according to the apparent magnitude of the object, the zero point and the full-width of the selected filter (see Table 8.3 for the filter characteristics). Moreover, the flux is scaled for the percentage of EE falling in the reference radius of the assumed spaxel.

As a second step, the spectrum is multiplied by the atmospheric transmission, and then the atmospheric and telescope thermal emissions, the sky continuum and the emission lines are added according to the selected site and telescope/mirror parameters. The spectrum is multiplied also by the total system efficiency, including the telescope throughput, by the effective photon collecting area and the exposure time, and then is convolved to the final selected spectral resolution ($R = 4500$ or $R = 3800$, see Table 8.2), with 2 pixels per spectral resolution element. The dark current, photon noise and read-out noise are added at the end. At the same time, a sky spectrum is simulated with the same resolution and adopting the same input parameters, and then is subtracted from the final spectrum. This procedure is repeated to create a spectrum for each of the 9 spaxels. The final spectra are then summed weighting them by the relative errors.

The input parameters that the code takes into account and that can be modified are:

- telescope diameter and effective photon collecting area
- spaxel area
- spectral resolution
- GLAO PSF for all the bands
- mirror coating
- telescope site
- dark current and read-out noise
- instrument efficiency
- exposure time
- airmass

The final simulated spectra for each band are compared with the input ones convolved to the same spectral resolution. We then plot the spectra in the whole optical-NIR spectral range (0.6–2.5 μm), in order to estimate which is the best spectral region for the study of these objects, on the basis of the features that we can identify and their S/N ratio. Furthermore, we compare spectra for the same temperatures (see Table 8.1) and different surface gravities, focusing in particular on the I-band spectral range (0.6–1.0 μm), where the gravity effects are strong in the region of the alkali lines. In the MIR case, we compare the spectra of the four fiducial cases simulated assuming different telescope sites and different brightnesses for the objects.

The outline of the simulation pipeline is given in Fig. 8.1.

8.4.3 Inputs

8.4.3.1 Scientific data

The science input are template spectra of mid-T dwarfs, which are expected to adequately represent the spectra of the least massive objects of interest in this proposal, covering the 0.6–4 μm wavelength range. We assume a fixed integration time of 2 hours. This time span is limited by the rotation period of the target. Simulations have been carried out by scaling the flux of the template spectrum by different factors, which represent either objects of different luminosities within a given star forming region, or objects of a characteristic intrinsic luminosity belonging to star forming regions placed at different distances.

We adopted BA03 predicted magnitudes at 5 Myr to set the fiducial cases to be considered. The second question in Section 8.2 obviously involves a trade-off between the maximisation of the S/N ratio for a given integration time, and the scientific interest of the spectral features accessible in a given wavelength range. Table 8.1 summarises the range of properties to be considered, for a distance modulus of 6 mag, that is roughly applicable to all the regions listed in Section 8.2. Fig. 8.2 shows a K-band image of the four objects representing the fiducial cases, simulated assuming a GLAO PSF, a Paranal-like site, and a pixel scale of 40 mas.

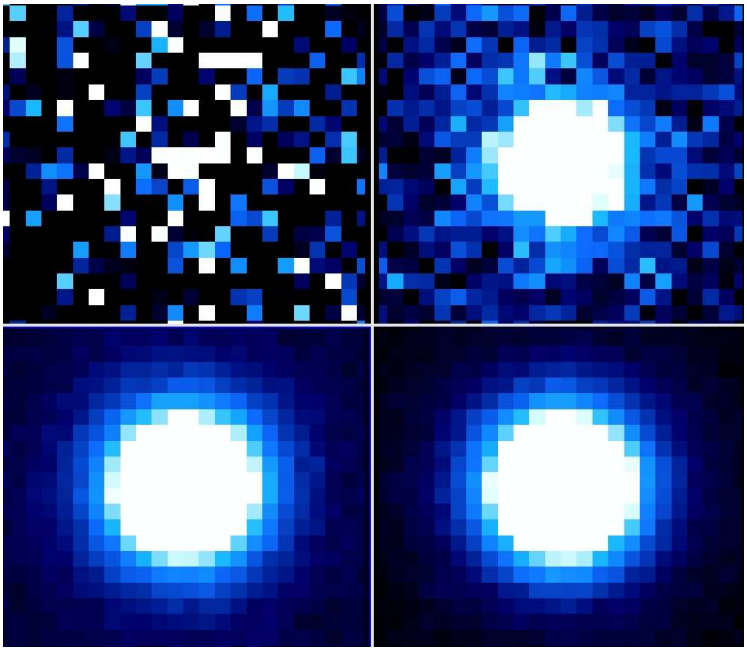

Figure 8.2: K-band images of the four objects corresponding to the fiducial cases (see Table 8.1), simulated with a GLAO PSF, a pixel size of 40 mas and $t_{\rm exp} = 3600, 1800, 1, 1$ s (top left to bottom right).

As a template we adopted F. Allard's synthetic spectra for BDs and extra-solar planets (Allard et al. 2000, see this web page for more information). In particular, we used the AMES-Cond model atmosphere grids for objects with $T_{\rm eff} < 1400K$ and the AMES-Dusty grids for objects with $T_{\rm eff} > 1700K$. Fig. 8.3 shows a synthetic input spectrum for an object with $T_{\rm eff} = 1098$ K and $\log g = 3.58$ in the optical-NIR wavelength range, convolved to the spectral resolution of $R = 4500$. Adopting a BA03 isochrone of t = 5 Myr and a distance modulus of 6 mag, these values correspond to a mass of 3 $M_{\rm Jup}$ and to an apparent magnitude of m_I = 23.54 (see Table 8.2). Fig. 8.4 shows a synthetic spectrum for the same input parameters but in the MIR wavelength range (m_L = 17.21 mag). Table 8.1 also lists the parameters of the other four cases we simulated to analyse the effects of gravity on the spectral features.

8.4.3.2 Technical data

The technical data used for the simulations were taken mainly from [RD1].

Site and background: We follow [RD1]. We use the background emission and atmospheric transmission model for the Paranal-like site for the optical-NIR case, while for the MIR case additional simulations are performed for the High & Dry site. We assume a zenith distance of 0° for all of the simulation runs.

Telescope: As described in [RD1]. We use both the bare aluminium and the protected silver/aluminium coatings.

Instrument: We assume an IFU with $R = 4500$ and with a spaxel size of 40 mas in the optical and NIR, and 20 mas in the MIR. The instrument efficiency is assumed to be $\sim$ 28% for both the optical-NIR and the MIR cases. This value does not include the detector QE, assumed to be 0.9 in the optical-NIR and 0.5 in the MIR (see [RD1]). The read-out noise of the detector was assumed to be 3 e^-/pix in the optical-NIR and 200 e^-/pix in the MIR (see [RD1]).

AO mode and PSFs: We perform the simulations assuming that conditions do not allow observations with good AO performance but only with the GLAO correction provided by the telescope. Hence we adopt the GLAO PSFs of [RD1], provided for the different bands (I_C, J, H, K, L'). We only use the on-axis PSFs, assuming that the object that we are observing is centred, and assume a zenith distance of 0°.

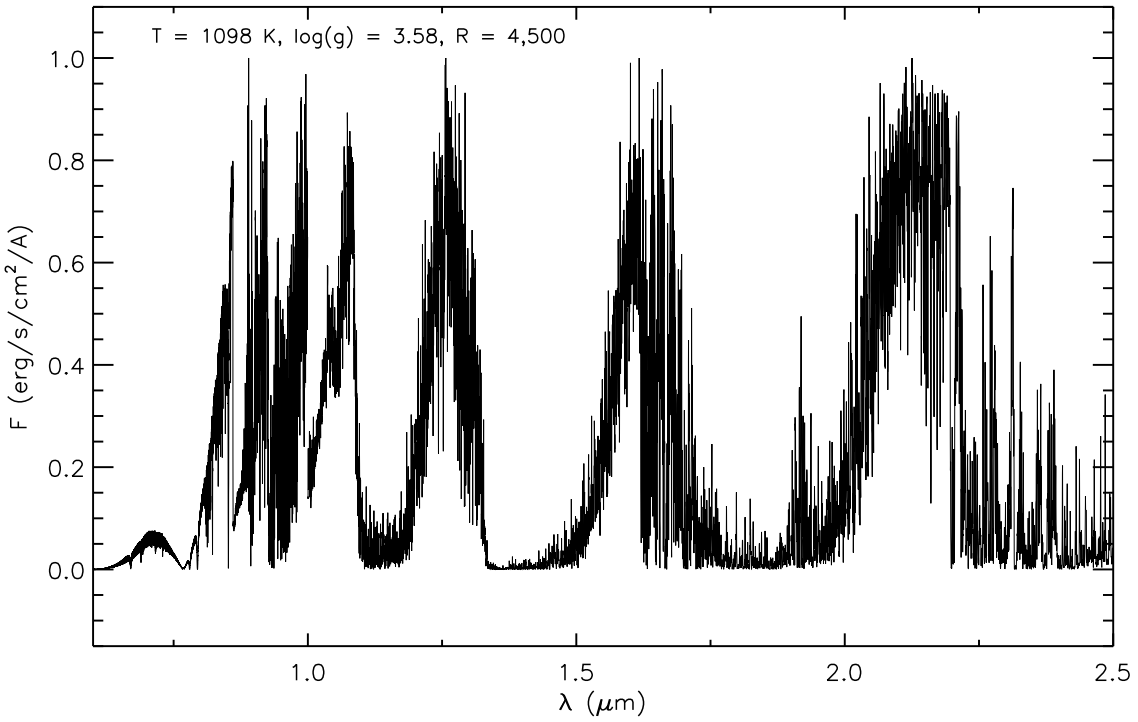

Figure 8.3: Synthetic spectrum in the optical-NIR wavelength range from the AMES-Cond model atmosphere grids (Allard et al. 2000) for $T_{\rm eff} = 1098$ K, $\log g = 3.58$, convolved to a resolution of $R = 4500$.

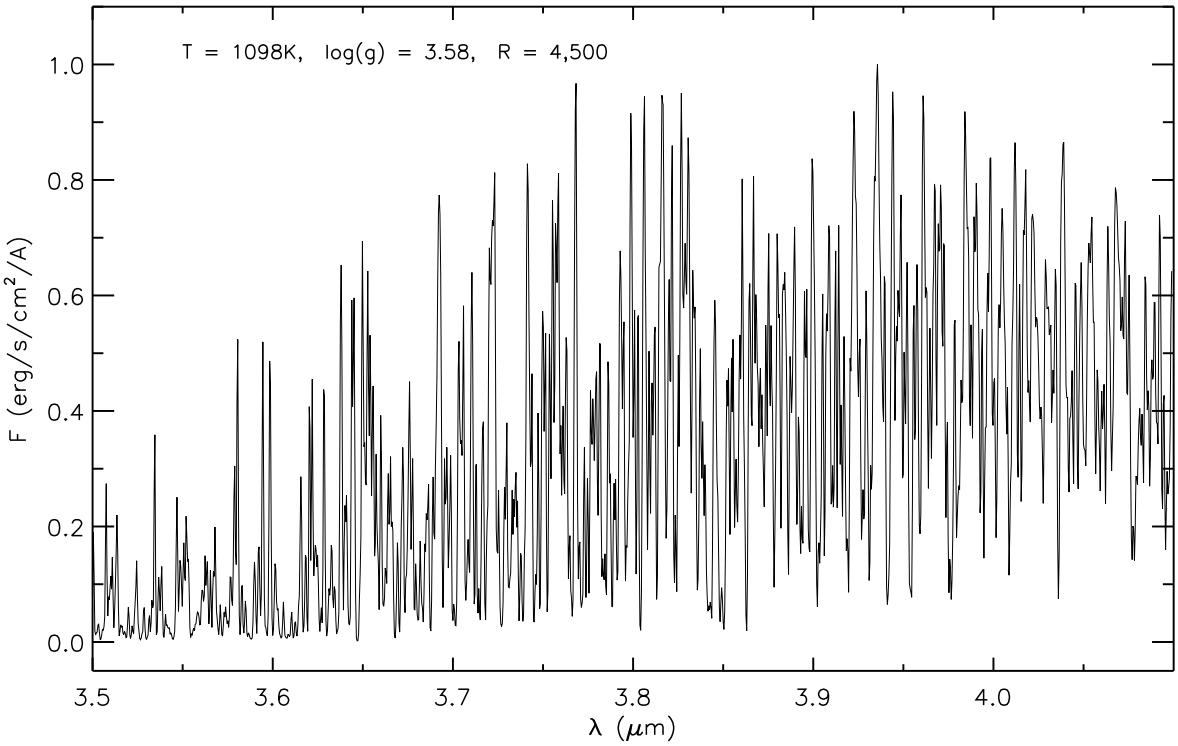

Figure 8.4: Same as Fig. 8.3 but for the MIR wavelength range.

8.4.4 Outputs

The simulation data products are 1-D spectra for each of the different runs.

Fig. 8.5 shows a simulated spectrum for an object with $T_{\rm eff} = 1098$ K and $\log g = 3.58$ (see the bottom left panel of Fig. 8.2) in the optical-NIR wavelength range, assuming a Paranal-like site and a silver/aluminium mirror coating. Fig. 8.6 shows the same spectrum simulated in the MIR wavelength range, assuming a High & Dry site and the same coating.

Fig. 8.7 shows a simulated spectrum for an object with $T_{\rm eff} = 644$ K and $\log g = 3.14$ (see the top right panel of Fig. 8.2) in the optical-NIR wavelength range, assuming a Paranal-like site and a silver/aluminium mirror coating. Fig. 8.8 shows the same spectrum simulated in the MIR wavelength range, assuming a High & Dry site and the same coating.

8.5 Results of simulations

8.5.1 Simulation runs

The different simulation runs and their input parameters are summarised in Table 8.4.

The exposure time is fixed at $t_{\rm exp} = 7200$ s for all the simulation runs since it is limited by the object rotational period ($T \approx 2$ h).

In reality, the observations should extend continuously for one or possibly two full nights to sample the rotational modulation of the objects. Then, they should be repeated once per week and once per month to sample the evolution of the cloud coverage and weather patterns. The total exposure time per instrument

Table 8.4: Input parameters for the different simulation runs.

Mass $M_{\rm Jup}$	$T_{\rm eff}$ K	$\log g$	Band	Site	Coating
Fiducial cases					
0.5	455	2.790	I	Paranal	Ag/Al
0.5	455	2.790	J	Paranal	Ag/Al
0.5	455	2.790	H	Paranal	Ag/Al
0.5	455	2.790	K	Paranal	Ag/Al
0.5	455	2.790	L'	Paranal	Ag/Al
1	644	3.141	I	Paranal	Ag/Al
1	644	3.141	J	Paranal	Ag/Al
1	644	3.141	K	Paranal	Ag/Al
1	644	3.141	H	Paranal	Ag/Al
1	644	3.141	L'	Paranal	Ag/Al
3	1098	3.576	I	Paranal	Ag/Al
3	1098	3.576	I	Paranal	Al
3	1098	3.576	J	Paranal	Ag/Al
3	1098	3.576	H	Paranal	Ag/Al
3	1098	3.576	K	Paranal	Ag/Al
3	1098	3.576	L'	Paranal	Ag/Al
3	1098	3.576	L'	High & Dry	Ag/Al
10	1965	3.921	I	Paranal	Ag/Al
10	1965	3.921	I	Paranal	Al
10	1965	3.921	J	Paranal	Ag/Al
10	1965	3.921	H	Paranal	Ag/Al
10	1965	3.921	K	Paranal	Ag/Al
10	1965	3.921	L'	Paranal	Ag/Al
10	1965	3.921	L'	High & Dry	Ag/Al
Comparison cases					
20	389	4.802	I	Paranal	Ag/Al
30	610	5.011	I	Paranal	Ag/Al
20	936	4.704	I	Paranal	Ag/Al
20	936	4.704	I	Paranal	Al
20	936	4.704	J	Paranal	Ag/Al
30	1979	4.715	I	Paranal	Ag/Al
30	1979	4.715	J	Paranal	Ag/Al
30	1979	4.715	H	Paranal	Ag/Al
30	1979	4.715	K	Paranal	Ag/Al

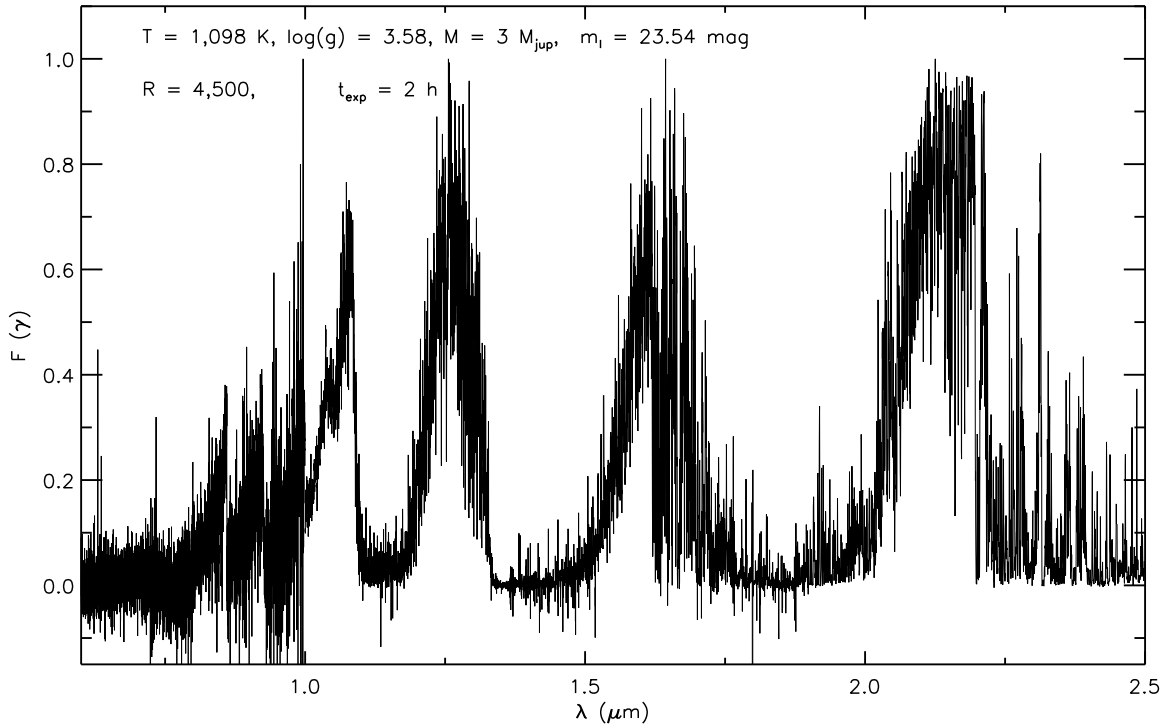

Figure 8.5: Simulated spectrum for an object with $T_{\rm eff} = 1098$ K and $\log g = 3.58$ in the optical-NIR wavelength range. The mass and the apparent magnitude are estimated by assuming a BA03 isochrone for $t = 5$ Myr and adding a distance modulus μ_0 = 6.0 mag.

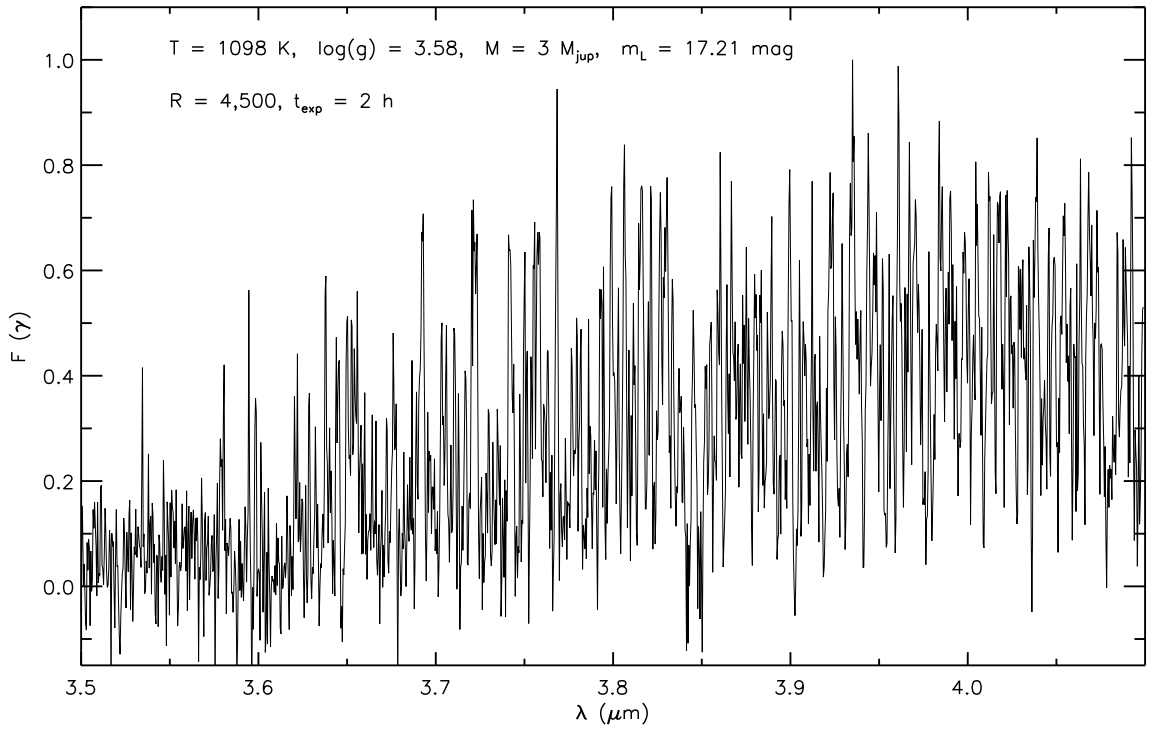

Figure 8.6: Same as Fig. 8.5 but for the MIR wavelength range.

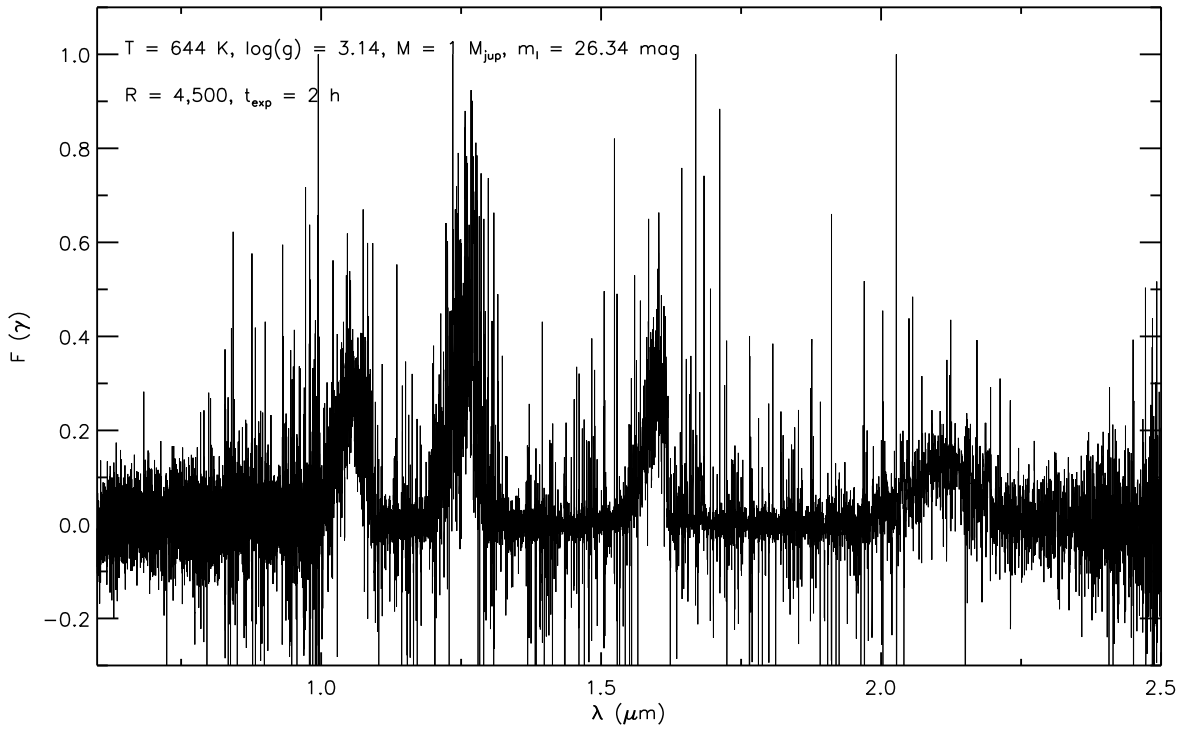

Figure 8.7: Simulated spectrum for an object with $T_{\rm eff} = 644$ K and $\log g = 3.14$ in the optical-NIR wavelength range. The mass and the apparent magnitude are estimated by assuming a BA03 isochrone for $t = 5$ Myr and adding a distance modulus μ_0 = 6.0 mag.

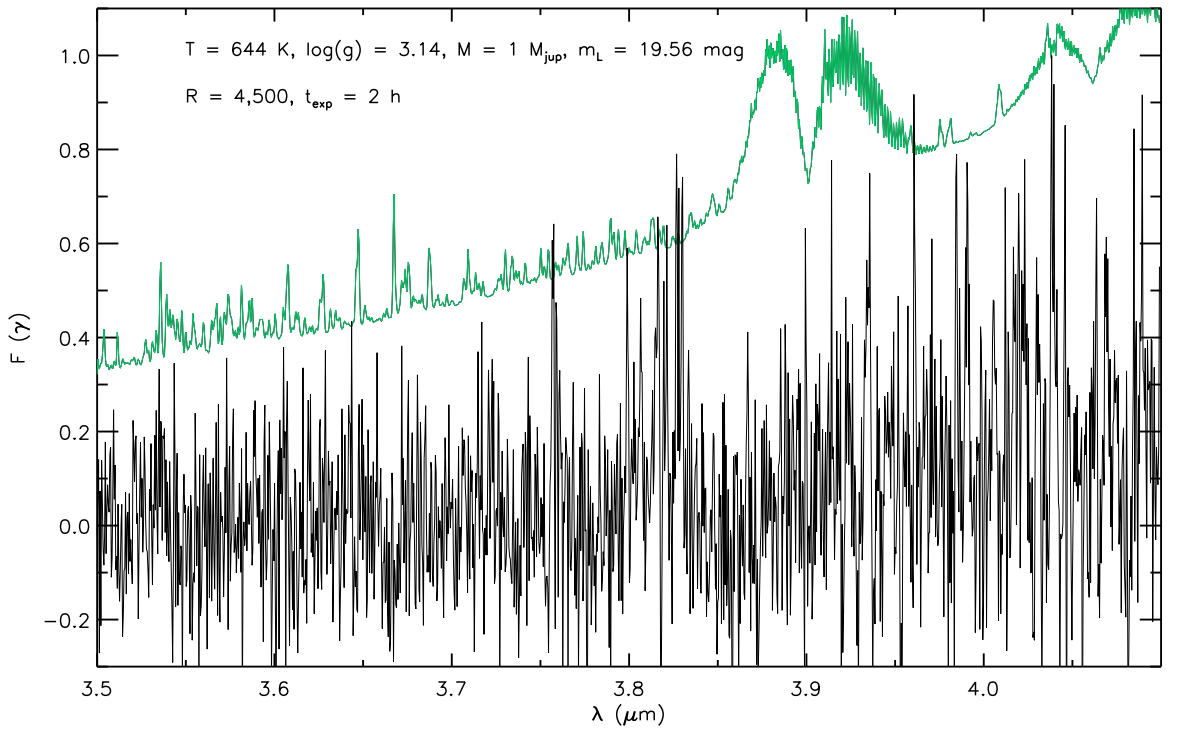

Figure 8.8: Same as Fig. 8.7 but for the MIR wavelength range. The green line shows the spectrum before the sky subtraction.

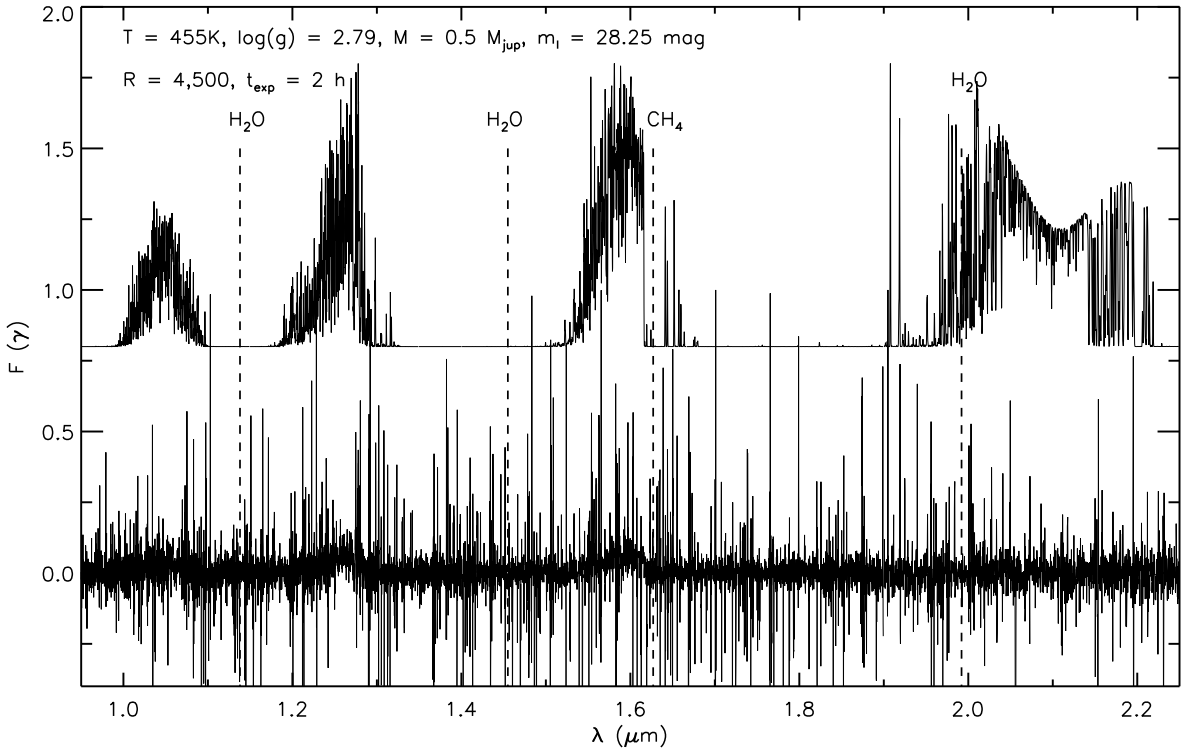

Figure 8.9: Simulated spectrum (bottom) in the NIR wavelength range for an object with $T_{\rm eff} = 455$ K and $\log g = 2.79$ compared to the synthetic input spectrum (top) from the AMES-Cond model atmosphere. The mass and the apparent magnitude are estimated by assuming a BA03 isochrone for $t = 5$ Myr and adding a distance modulus μ_0 = 6.0 mag. The water vapour bands and the CH_4 absorption bands are labelled with dashed lines.

setting (NIR + MIR) is estimated to be $T_{\rm tot} \approx 100$ h, in order to observe approximately 10 objects in a given star forming region.

We tested the different mirror coatings for two fiducial cases in the I-band, since in this region ($0.7 < \lambda < 0.9$ μm) the mirror reflectivity of the bare aluminium coating shows a pronounced dip.

The site has been tested only for the MIR case, since the telescope thermal emission depends on the site temperature and strongly affects observations in the MIR wavelength range.

8.5.2 Analysis

We now compare the simulated spectra in the I, J, H, K, L' bands with the input synthetic ones. Fig. 8.9 shows a simulated spectrum for an object with $T_{\rm eff} = 455$ K and $\log g = 2.79$ (see the top left panel of Fig. 8.2) in the NIR wavelength range, assuming a Paranal-like site, a silver/aluminium mirror coating and an exposure time of 2 hours, compared to the input synthetic spectrum from the AMES-Cond model atmosphere.

The simulated spectrum is clearly dominated by noise, with S/N $\approx$ 1–3. On the other hand, some spectral features are detectable in the NIR wavelength range, such as the water vapour bands around 1.2 μm, 1.5 μm and 2.0 μm, and the CH_4 absorption band around 1.6 μm. These features are marked with dashed lines in Fig. 8.9, and the corresponding S/N is indicated in Table 8.5.

Figs. 8.10 and 8.11 show the simulated spectra in the NIR wavelength range of an object with $T_{\rm eff} = 1098$ K, $\log g = 3.58$ and an object with $T_{\rm eff} = 644$ K, $\log g = 3.14$. In both cases we adopted a Paranal-like site, a silver/aluminium mirror coating and an exposure time of 2 hours. The water vapour bands around

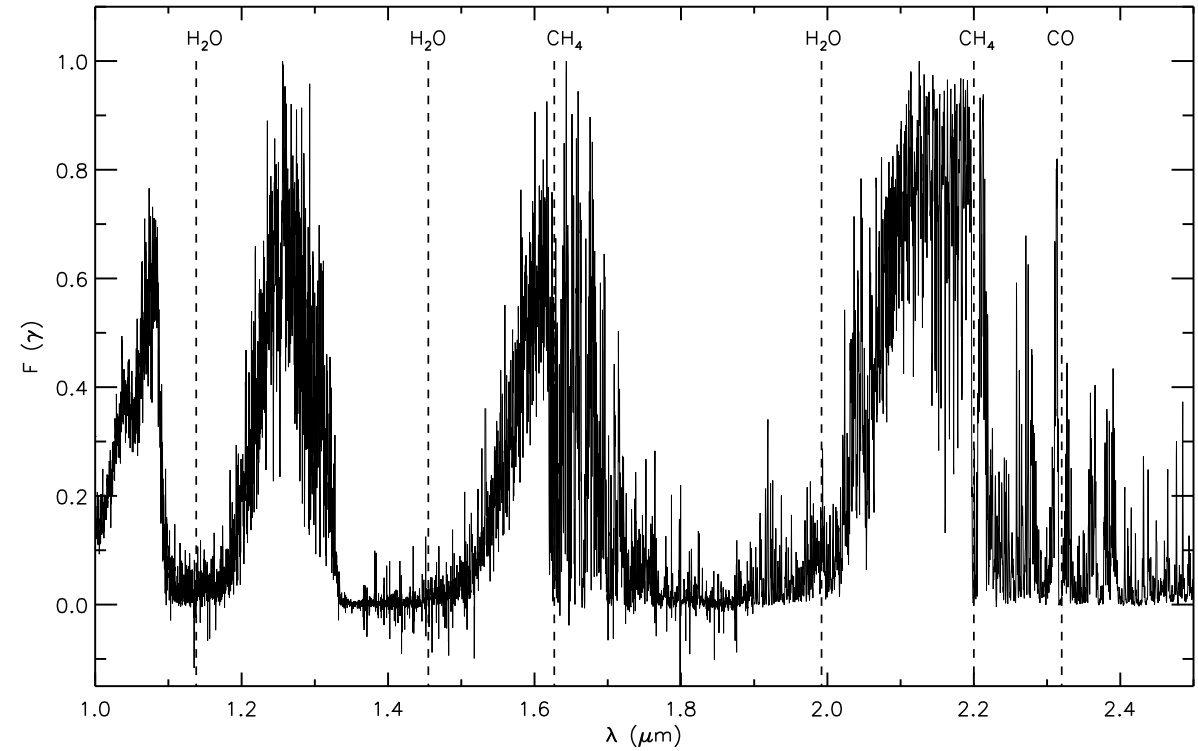

Figure 8.10: Simulated spectrum in the NIR wavelength range for an object with $T_{\rm eff} = 1098$ K and $\log g = 3.58$, for a Paranal site, a silver/aluminium coating, and $t_{\rm exp}$ = 2 h. The water vapour bands and the CH_4 absorption bands are labelled with dashed lines.

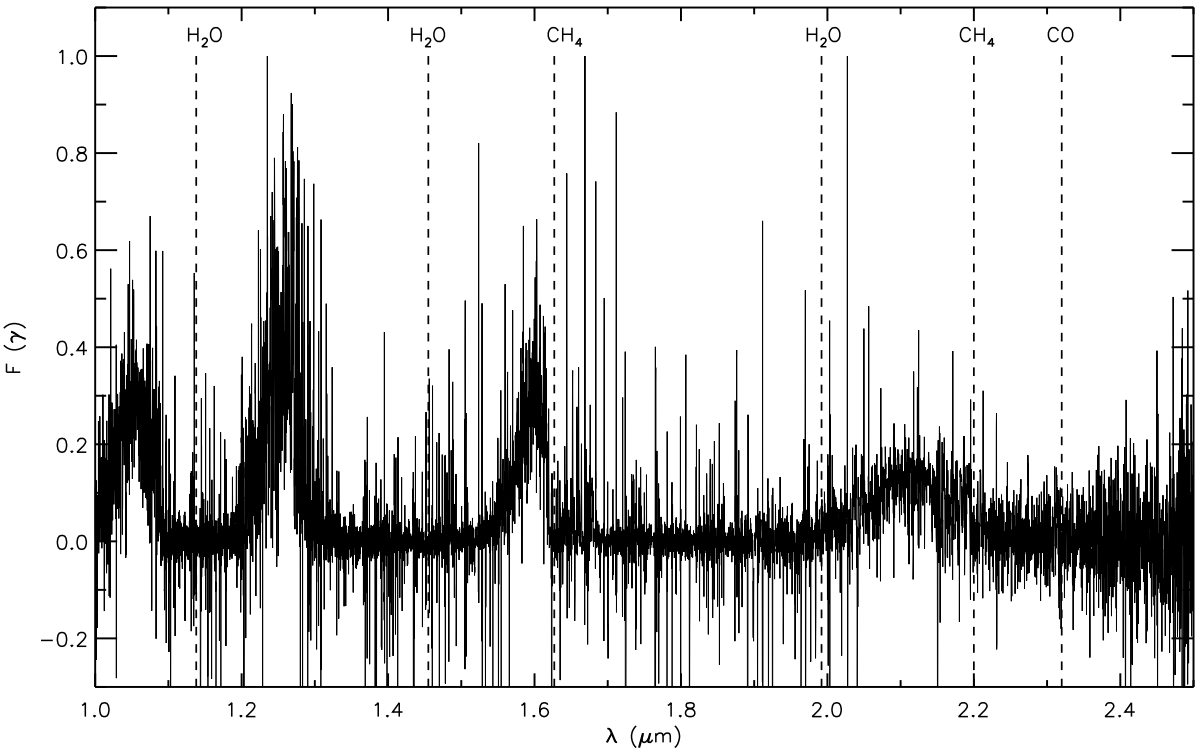

Figure 8.11: Same as Fig. 8.10 but for an object with $T_{\rm eff} = 644$ K and $\log g = 3.14$.

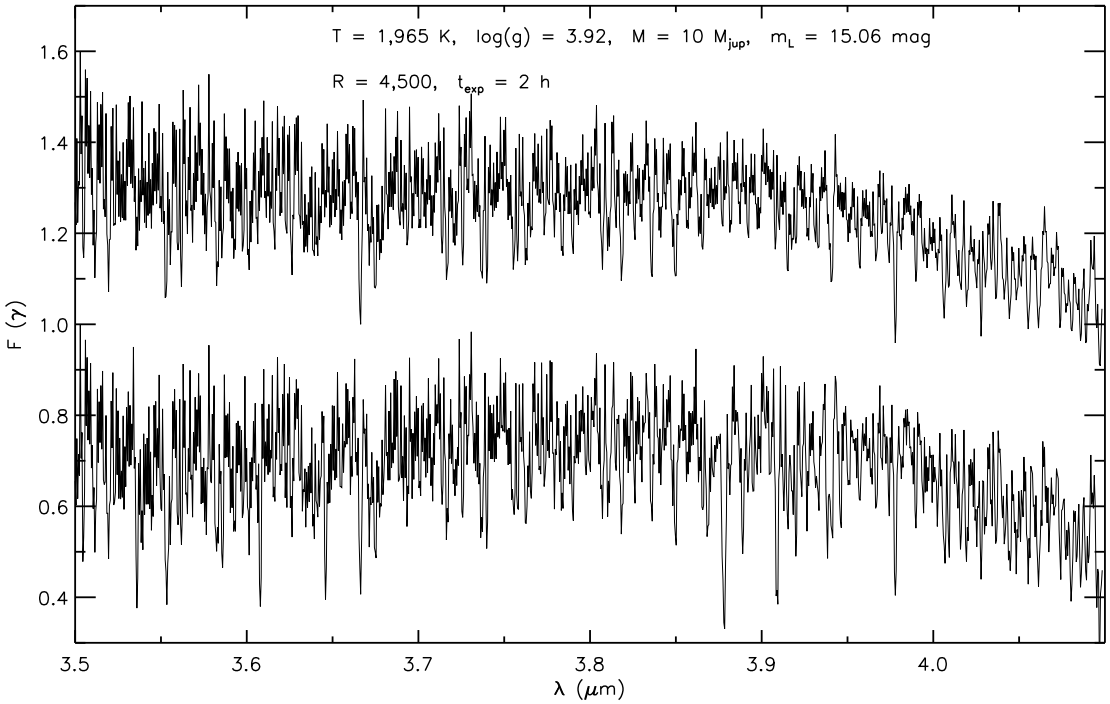

Figure 8.12: Simulated spectrum (bottom) in the MIR wavelength range for an object with $T_{\rm eff} = 1965$ K and $\log g = 3.92$, for a High & Dry site, a silver/aluminium coating, and $t_{\rm exp}$ = 2 h, compared to the synthetic input spectrum (top) from the AMES-Dusty model atmosphere. The mass and the apparent magnitude are estimated by assuming a BA03 isochrone for $t = 5$ Myr and adding a distance modulus μ_0 = 6.0 mag. Note that most of the features detectable in the simulated spectrum (bottom) are real.

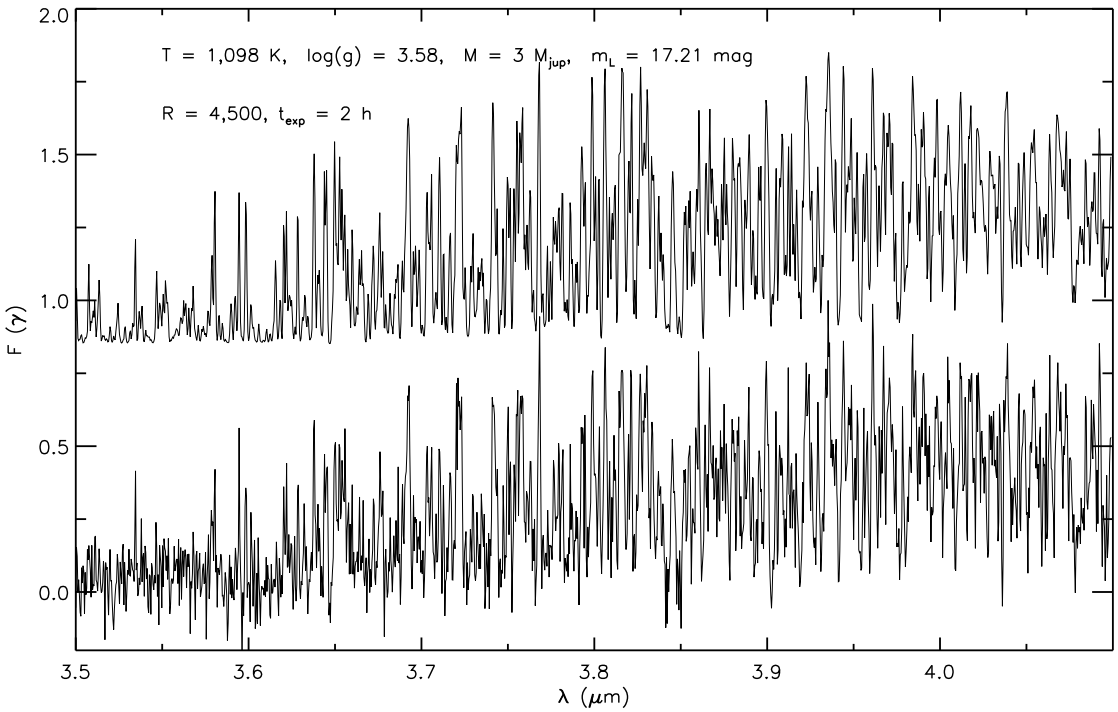

Figure 8.13: Same as Fig. 8.12 but for an object with $T_{\rm eff} = 1098$ K and $\log g = 3.58$, compared to the synthetic input spectrum from the AMES-Cond model atmosphere.

1.2 μm, 1.5 μm and 2.0 μm, the CH_4 absorption bands around 1.6 μm and 2.2 μm, and the CO absorption line are marked with dashed lines in both figures and the S/N of these features is listed in Table 8.5.

These figures and the table show that the best spectral region to observe the H_2O and CH_4 absorption bands is the K-band, for objects with masses $M \geq 1 M_{\rm Jup}$.

Fig. 8.12 shows a simulated spectrum (bottom) in the MIR wavelength range for an object with $T_{\rm eff} = 1965$ K and $\log g = 3.92$, produced assuming a High & Dry site, a silver/aluminium coating and $t_{\rm exp}$ = 2 h, compared to the input AMES-Dusty model atmosphere (top). It is worth noticing that most of the features detectable in the simulated spectrum are real. The S/N at the central wavelength ($\lambda \approx 3.8$ μm) is $\approx$ 15 (see Table 8.5).

For temperatures lower than ≈ 1000 K, the sky and the telescope thermal emission dominate the spectra, and only a few spectral features are detectable at least in the case of $T_{\rm eff} = 1098$ K and $\log g = 3.58$ (see Fig. 8.13), where the S/N is $\approx$ 5 at the central wavelength.

Table 8.5: S/N ratio of key features of the simulated spectra.

Mass $M_{\rm Jup}$	$T_{\rm eff}$ K	$\lambda \sim 0.9$ μm	H_2O 1.2 μm	H_2O 1.5 μm	CH_4 1.6 μm	H_2O 2 μm	CH_4 2.2 μm	$\lambda \sim 3.8$ μm
0.5	455	$\approx$ 1-3	$\approx$ 2	$\approx$ 2	$\approx$ 2	$\approx$ 2	$\approx$ 2	$\approx$ 2-3
1	644	$\approx$ 1-3	$\approx$ 5	$\approx$ 5	$\approx$ 5	$\approx$ 5	$\approx$ 10	$\approx$ 2-3
3	1098	$\approx$ 10	$\approx$ 10	$\approx$ 10	$\approx$ 10	$\approx$ 15	$\approx$ 15	$\approx$ 5
10	1965	$\approx$ 20-30	$\approx$ 30	$\approx$ 30	$\approx$ 40	$\approx$ 30	$\approx$ 30	$\approx$ 15

8.5.3 Compliance with figures of merit

Fig. 8.14 shows a comparison between the spectra of two objects with the same temperature ($T_{\rm eff} \sim 2000$ K) but different surface gravities ($\log g \sim 4$ and ~ 5). We assumed that the objects have the same apparent magnitudes in the I, J, H, K bands (see Table 8.1).

After a qualitative analysis of the spectra we can argue that we would not be able to distinguish the two objects observing them in the NIR wavelength range only. The gravity effects are indeed stronger in the optical red region. Fig. 8.15 shows a zoom of the two spectra in the 0.7–0.95 μm wavelength range: the differences in the profile of the neutral alkali elements lines (Li, Na, K, Cs, Rb) are now clearly detectable.

The same experiment has been performed for objects with temperatures around $T_{\rm eff} \sim 1000$ K ($\log g \sim$ 3.5 and $\sim$ 4.5, Fig. 8.16), $T_{\rm eff} \sim 600$ K ($\log g \sim 3$ and ~ 5, Fig. 8.17), and $T_{\rm eff} \sim 400$ K ($\log g \sim 3$ and ~ 5), in the optical wavelength range only (see also Table 8.1).

These figures show that in spite of the very low S/N in these spectra (≤ 10) we are still able to distinguish between objects with different gravities down to $T_{\rm eff} \sim 1000$ K. For a surface gravity $\log g$ = 3 this corresponds to a mass of ~ 3 $M_{\rm Jup}$ and to an apparent magnitude $m_I \sim$ 23.5 mag, assuming a BA03 isochrone for an age t = 5 Myr and a distance of $\sim$150 pc. For temperatures lower than 1000 K (Fig. 8.17) we detect mostly noise in the spectra, and we are not able to disentangle objects with different surface gravities.

8.5.4 Sensitivity to input parameters

We performed different simulation runs changing various input parameters, as already mentioned in Sections 8.4.2 and 8.5.1.

The most critical input parameter for the optical wavelength region is the mirror coating because the bare aluminium coating shows a pronounced drop in the wavelength region 0.7–0.9 μm.

Therefore, we tested the two mirror coatings for the case of the objects with $T_{\rm eff} = 1098$ K and $T_{\rm eff} = 1965$ K in the I-band. We focused in the region $0.7 < \lambda < 0.9$ μm, where the difference in the alkali line profiles due to the surface gravity are detectable.

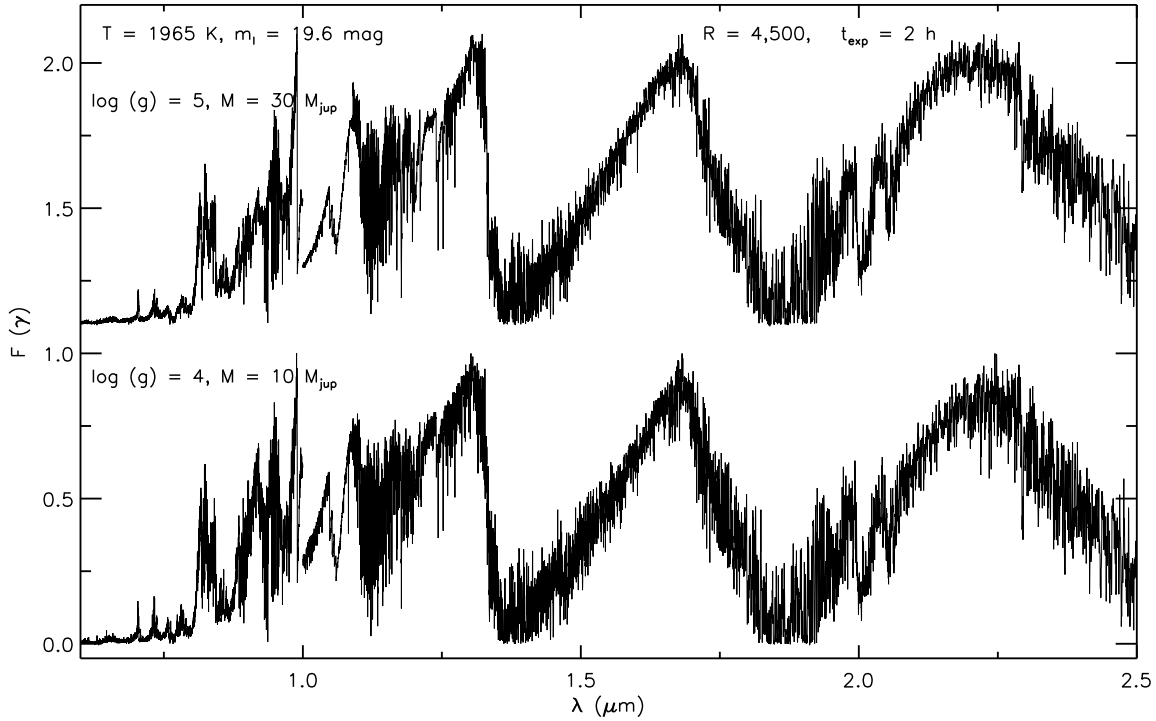

Figure 8.14: Comparison between the simulated spectra of two objects with the same temperature ($T_{\rm eff} \sim$ 2000 K) but different surface gravities ($\log g \sim$ 4 and 5) in the optical-NIR wavelength range. The same apparent magnitude has been assumed for the objects and is labelled in the figure.

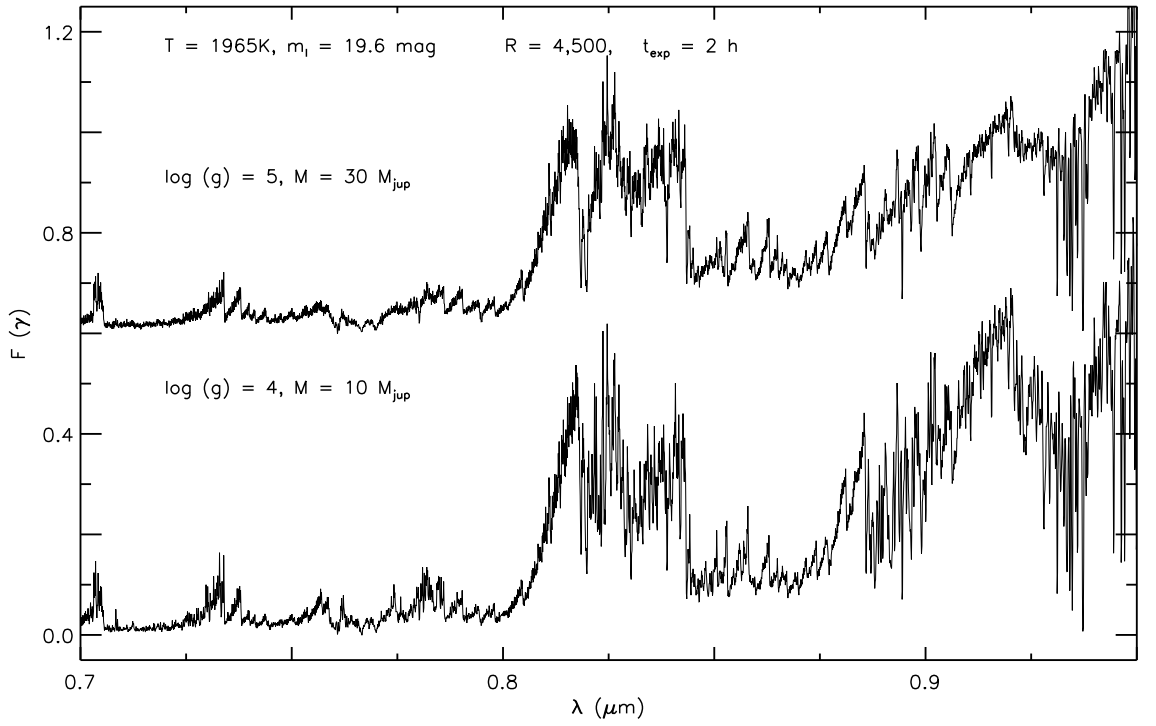

Figure 8.15: Same as Fig. 8.14 but zoomed in the optical wavelength region, where the alkali lines can be observed.

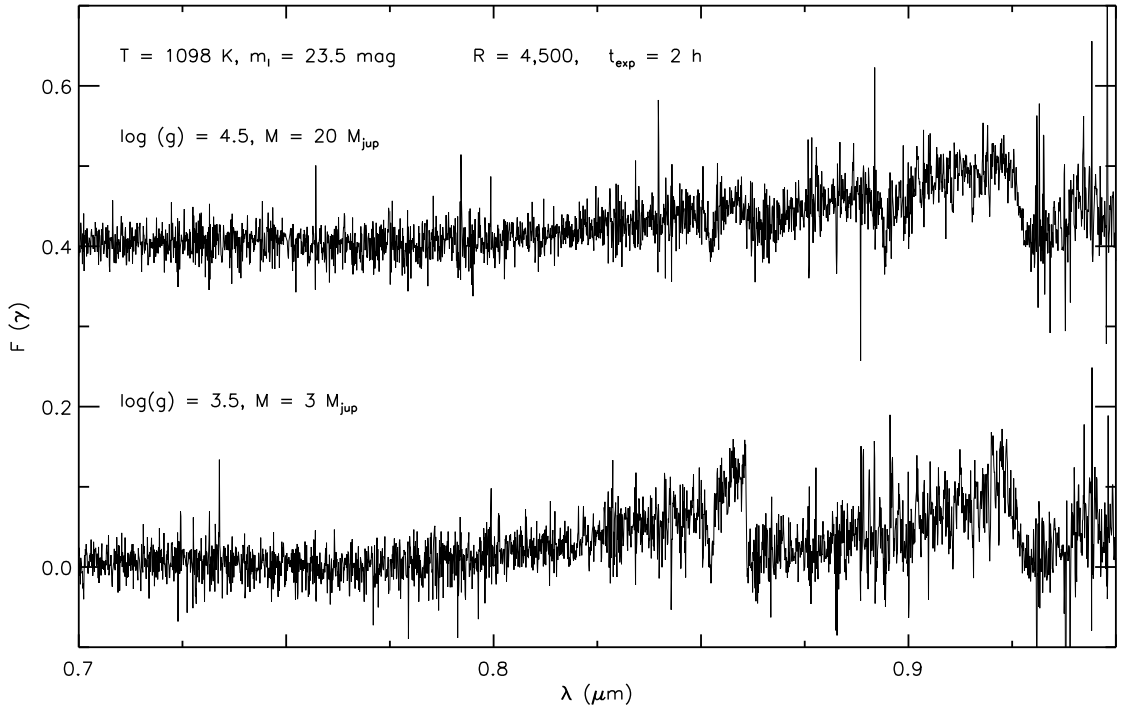

Figure 8.16: Comparison between the simulated spectra of two objects with the same temperature ($T_{\rm eff} \sim 1000$ K) but different surface gravities ($\log g \sim$ 3.5 and 4.5) in the optical wavelength range. The same apparent magnitude has been assumed for the objects and is labelled in the figure.

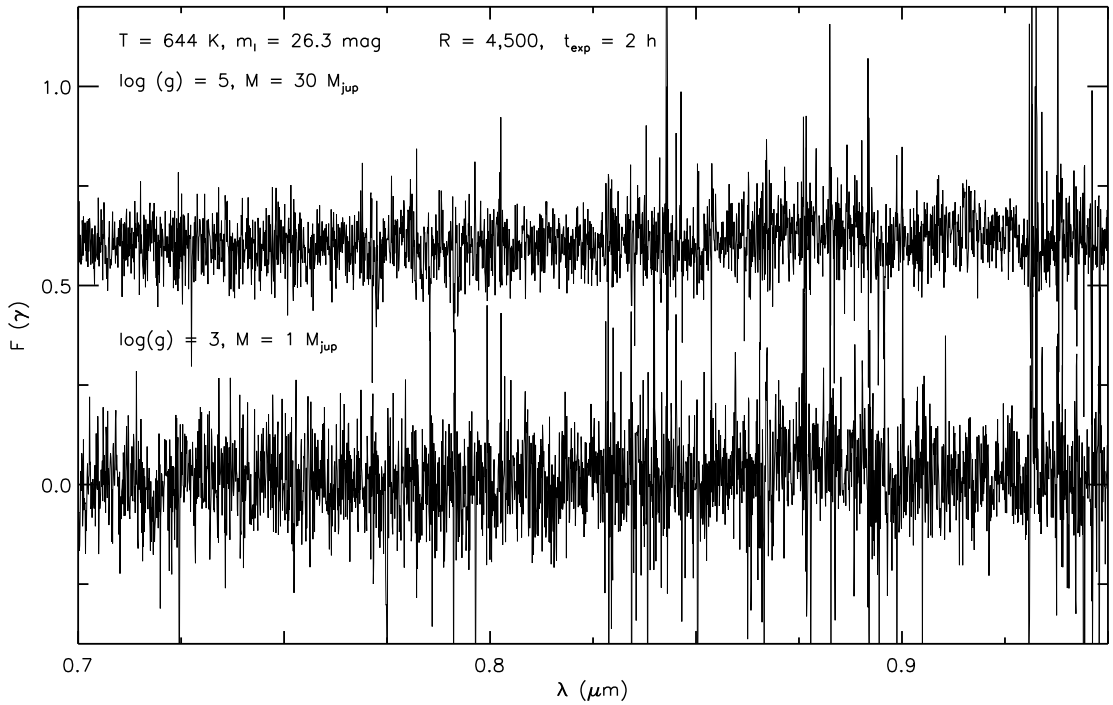

Figure 8.17: Comparison between the simulated spectra of two objects with the same temperature ($T_{\rm eff} \sim 600$ K) but different surface gravities ($\log g \sim$ 3 and 5) in the optical wavelength range. The same apparent magnitude has been assumed for the objects and is labelled in the figure.

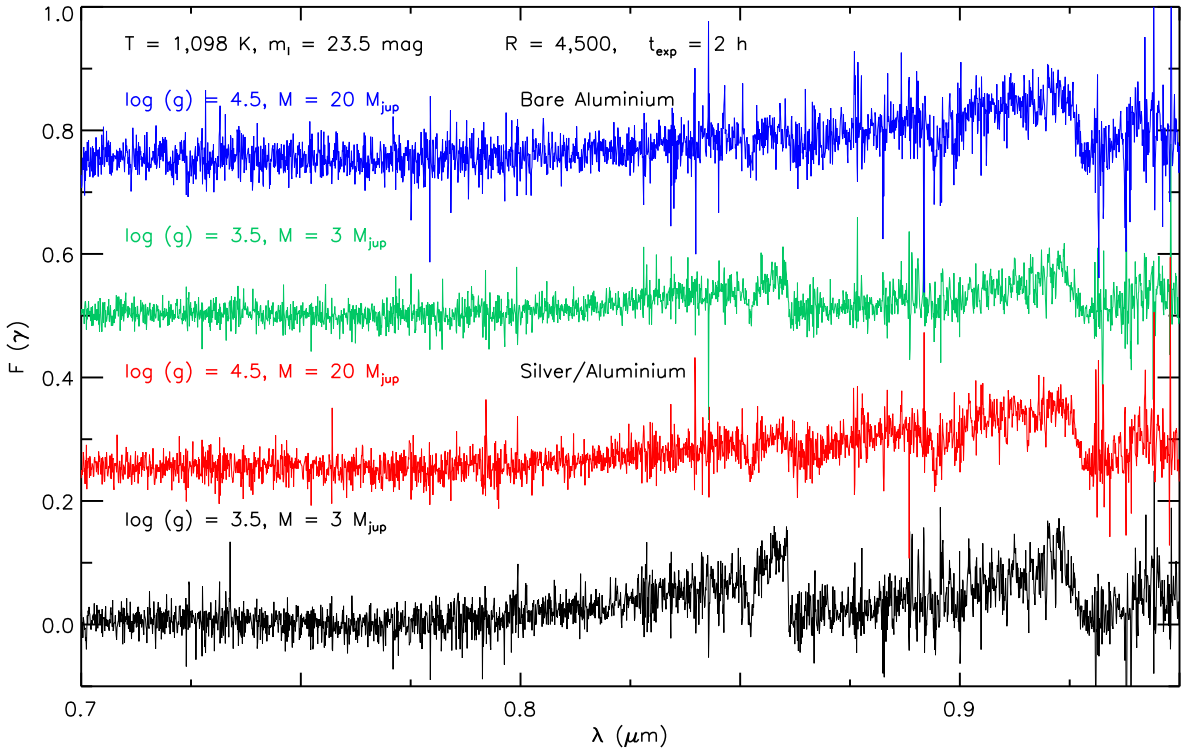

Figure 8.18: Comparison between four spectra for two objects with $T_{\rm eff} \sim 1098$ K and $\log g = 3.5$ or $\log g = 4.5$, respectively, simulated assuming a bare aluminium (green and blue lines) or a silver/aluminium (black and red) coating. The same apparent magnitude has been assumed for the objects and is labelled in the figure. The normalisation is the same for all the spectra.

Fig. 8.18 shows the comparison between four spectra for two objects with $T_{\rm eff} \sim 1098$ K and $\log g = 3.5$ or $\log g = 4.5$, respectively, simulated assuming a bare aluminium (green and blue lines) or a silver/aluminium (black and red) coating, an apparent magnitude of m_I = 23.5 mag (μ_0 = 6 mag) and an exposure time of 2 hours.

The differences between the spectra simulated with two different gravities is still detectable but less clear in the case of the bare aluminium coating. It is worth noticing that the change in the profile of some lines, such as around $\lambda \approx 0.85$ μm, in the spectrum simulated for $\log g = 3.5$ and a bare aluminium coating (green line in Fig. 8.18) mimics the change that occurs in the spectrum of an object simulated for $\log g = 4.5$ and a silver/aluminium coating (red line). However, this effect should disappear after the spectra are flux calibrated.

We simulated spectra assuming a different telescope site, i.e. Paranal-like or High & Dry, only for the MIR case, since the telescope thermal emission depends on the site temperature and strongly affects observations in the MIR wavelength range. Fig. 8.19 shows the binned spectrum of an object with $T_{\rm eff} = 1965$ K and $\log g = 3.92$, simulated assuming a Paranal-like site (green and black solid lines) or a High & Dry site (blue and red), for two different assumed apparent magnitudes, $m_L \sim 15$ mag (signal-dominated case) and $m_L \sim 17$ mag (noise-dominated). We can argue from a qualitative comparison with the input synthetic spectrum (cyan) that in noise-dominated conditions the difference between the two simulated spectra (red and black) is negligible, and almost all the spectral features disappear. In signal-dominated conditions, there is nearly no difference between the two simulated spectra, and almost all the spectral features are detectable.

Fig. 8.20 then shows the binned spectrum of the same object but simulated for an apparent magnitude of $m_L \sim 16$ mag, assuming a Paranal-like site (black solid line) or a High & Dry site (red) compared to the synthetic input spectrum (green). This is a case where it makes a difference to observe at a High & Dry

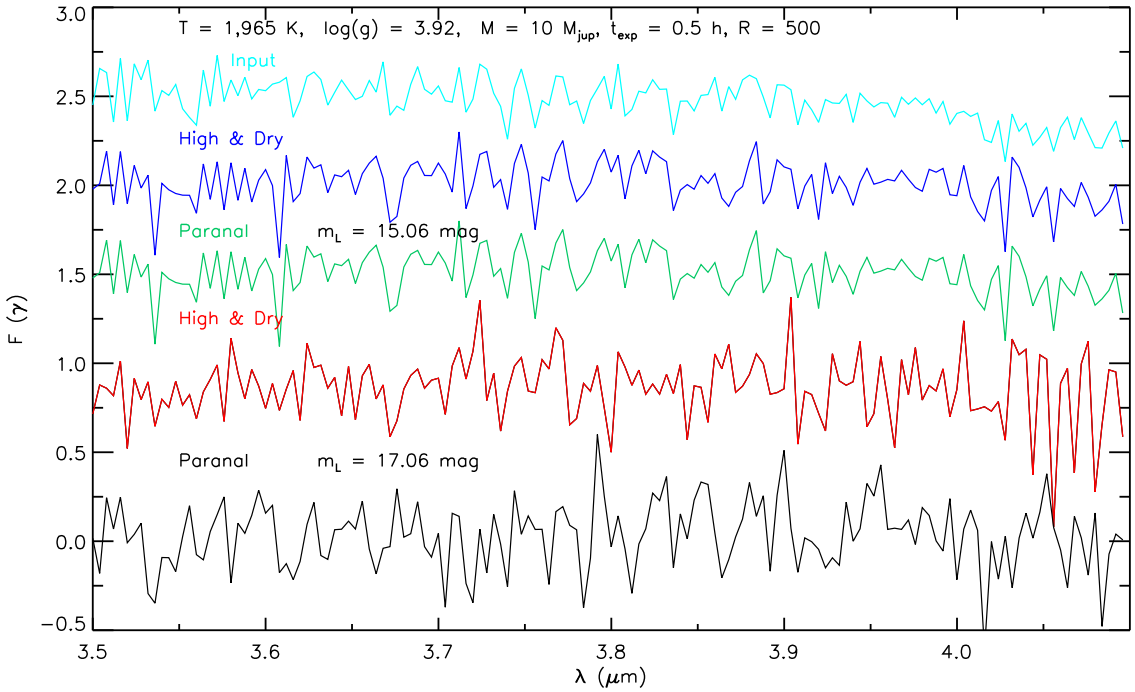

Figure 8.19: Comparison between the binned spectrum of an object with $T_{\rm eff} = 1965$ K and $\log g = 3.92$ simulated assuming a Paranal-like site (green and black solid lines) or a High & Dry site (blue and red), for two different assumed apparent magnitudes, $m_L \sim 15$ mag (signal-dominated conditions) and $m_L \sim 17$ mag (noise-dominated), and the synthetic input spectrum (cyan).

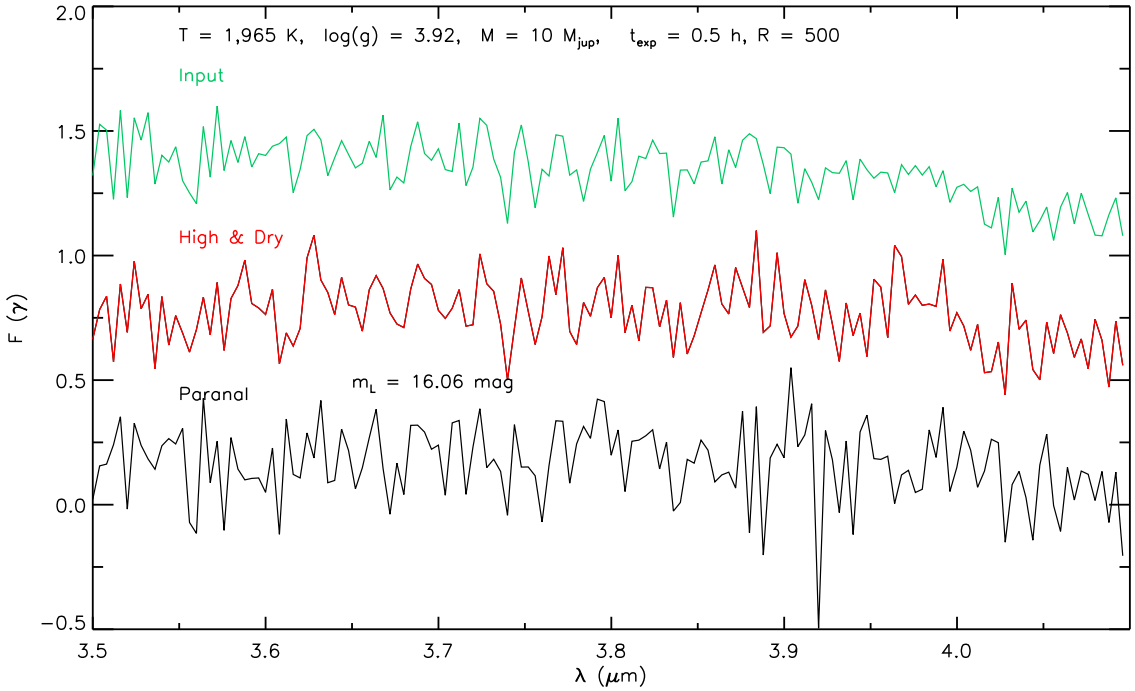

Figure 8.20: Comparison between the binned spectra of an object with $T_{\rm eff} = 1965$ K and $\log g = 3.92$ simulated assuming a Paranal-like site (black solid line) or a High & Dry site (red), for the same apparent magnitude $m_L \sim 16$ mag, and the synthetic input spectrum (green).

site: some spectral features are still detectable in the spectrum simulated for the High & Dry site, while almost no features are detectable in the spectrum simulated for the Paranal-like site.

It is noteworthy that in the wavelength range 3.5 $< \lambda <$ 4.1 μm low-resolution spectroscopy ($R \approx 4000$) with an E-ELT and only GLAO correction of a planetary mass object ($M \leq 10\ M_{\rm Jup}$) is not affected by the telescope site in noise ($m_L \geq 17$ mag) and signal-dominated ($m_L \leq 15$ mag) conditions, but the quality of the spectra improves when passing from a Paranal-like to a High & Dry telescope site, when we are observing in 'intermediate' conditions, i.e., for this case, $15 \leq m_L \leq 17$ mag.

8.5.5 Calibration requirements

There are no special calibration requirements to execute this program.

8.5.6 Limitations

This proposal was designed to explore the performance of the E-ELT under non-optimal atmospheric conditions. The assumption is that a program of these characteristics will be executed in a survey-like manner at times when the conditions do not allow observations with the good AO performance required by most other science cases.

The simulation results are then limited by the adoption of atmosphere parameters typical of poor conditions and by the use of only the GLAO correction provided by the telescope.

A possible improvement of the results will be the adoption of the LTAO correction. In the case of the K-band, for instance, where we better observe the water vapour and the CH_4 absorption bands, the object flux in the central spaxel (40 mas) would increase from $\approx$ 4% (GLAO) to 52% (LTAO). A similar improvement will occur in the L'-band region (spaxel of 20 mas, 5% to 51%).

8.6 Concluding remarks

The main conclusions are the following:

Optical-NIR wavelength region:

- The simulations showed that, given the limitations of a short exposure time ($t_{\rm exp} = 7200$), low-resolution ($R = 4000$), GLAO correction, and site (Paranal-like), the best spectral region to observe the H_2O and CH_4 absorption bands is the K-band, for objects with masses $M \geq 1\ M_{\rm Jup}$ and a S/N $\geq$ 10.
- The simulations showed that in spite of the low S/N (< 10) we are able to distinguish between objects with the same effective temperature but different surface gravities down to $T_{\rm eff} \sim 1000$ K in the optical wavelength region (0.7 $< \lambda <$ 0.95 μm). For a surface gravity $\log g$ = 3 this corresponds to a mass of $\sim 3\ M_{\rm Jup}$ and to an apparent magnitude $m_I \sim$ 24 mag, assuming a BA03 isochrone for an age t = 5 Myr and a distance of $\sim$150 pc.
- The differences between the spectra simulated with the same temperature and two different surface gravities is still detectable but less clear when assuming a bare aluminium mirror coating instead of a protected silver/aluminium coating. It is worth noticing that the change in the profile of some lines due to gravity mimics the change that occurs when the spectrum of the object is simulated adopting a different mirror coating (see Fig. 8.18). However, this effect should disappear after normalising the flux of the spectra with a telluric reference star.

MIR wavelength region:

- For temperatures higher than $T_{\rm eff} \sim 1000$ K ($m_L < 17$ mag), most of the spectral features are detectable in the simulated spectra of planetary mass objects ($M > 3\ M_{\rm Jup}$). The choice of the telescope site does not affect the simulations when the spectra are signal-dominated, i.e. for $m_L \leq$

15 mag. On the other hand, the quality of the spectra improves when passing from a Paranal-like to a High & Dry site.

- For temperatures lower than $T_{\rm eff} \sim 1000$ K ($m_L \geq 17$ mag), the sky and the telescope thermal emission dominate the simulated spectra (noise-dominated conditions). Only a few main spectral features are detectable at least in the case of $T_{\rm eff} = 1098$ K and $\log g = 3.58$. The choice of the telescope site does not affect the simulations since the spectra are noise-dominated.

References

Allard F., et al., 2000, ApJ, 539, 366
Artigau E., et al., 2003, in "Brown Dwarfs", IAU Symp. 211, APS Conf. Ser.
Baraffe I., et al., 2003, A&A, 382, 563 (BA03)
Burgasser A.J, et al., 2002, ApJ, 571, L151
Burrows A., et al., 2000, ApJ, 531, 438
Burrows A., et al., 2001, Rev. Mod. Phys., 73, 719
Burrows A., et al., 2003, ApJ, 596, 587
Burrows A., et al., 2006, ApJ, 640, 1063
Caballero J.A., et al., 2007, A&A, 470, 903
Enoch M.L., et al., 2003, AJ, 126, 1006
Geballe T.R., et al., 2001, ApJ, 556, 373
Helling C., et al., 2008, A&A, 485, 547
Lodders K., 1999, ApJ, 519, 793
Marley M.S., et al., 2002, ApJ, 568, 335
Marley M.S., et al., 2007, ApJ, 655, 541
Saumon D., et al., 2007, ApJ, 656, 1136
Zapatero Osorio M.R., et al., 2002, ApJ, 578, 536
Zapatero Osorio M.R., et al., 2006, ApJ, 647, 1405

9 S5-2: The centres of massive dense young clusters: deep E-ELT infrared imaging and 3D spectroscopy

Authors: A. Calamida, M. Nonino, H. Zinnecker

9.1 The science case

We propose to use the 42 m E-ELT at 2–5 μm to probe the number density and brightness of deeply embedded massive stars and proto-stars just formed in dense Galactic proto-cluster clouds (ultra-compact H II regions, hot cores, outflows and maser sources), penetrating as much as 200 mag of visual extinction. The combination of astrometric, proper motion (1 mas/yr) and spectroscopic, radial velocity ($R \sim 10^4$) data are crucial to study dynamical processes associated with cluster formation, such as tight binary formation and gravitational interactions followed by stellar ejections. Integral field spectroscopy is needed for these dense and severely crowded clusters (up to 1000 objects per arcsec^2 at K = 25–30 mag).

Massive stars ($M > 20\ M_\odot$) represent powerful engines of change/evolution within their Galactic and extragalactic environment. When they die and explode as supernovae, they dramatically deposit lots of kinetic energy and chemically enriched material in the interstellar medium. Yet despite their enormous overall importance for Galactic dynamical and chemical evolution, their origins and birth processes remain poorly understood (Bally & Zinnecker 2005; Zinnecker & Yorke 2007).

In this DRM report, we show how the E-ELT will help to understand the formation and early dynamical evolution of massive stars embedded in highly obscured, very compact H II regions. The wavelength range of 2–5 μm (K, L and M bands) is required to penetrate the visual extinction (up to A_V = 200 mag) in the dense molecular clumps ($N_{H_2} = 3 \times 10^{23}$ cm^{-2}) where massive stars are born. Note that J-band (1.2 μm) and H-band (1.6 μm) observations do not peer into such dense dusty birthplaces, while K-band observations do ($A_J = 0.28\,A_V$, $A_H = 0.18\,A_V$, $A_K = 0.11\,A_V$, with $A_V \lesssim$ 200 mag; based on the interstellar extinction law of Rieke & Lebofsky 1985). The L and the M-band ($A_L = 0.06\,A_V$, $A_M = 0.02\,A_V$) data are needed to probe the hot and warm circumstellar matter (disks/envelopes) by means of the detection of infrared excess ($K-L$, $L-M$); the Pfund γ (3.76 μm) and Brackett α (4.05 μm) recombination lines also fall into these L and M-band, respectively, and allow us (in conjunction with their flux ratio to Brackett γ at 2.17 μm and via Menzel Case B recombination line theory) to determine the extinction towards resolved ionising stellar sources, i.e. massive stars, in ultra-compact (size $\sim$0.1 pc) and hyper-compact (size $\sim$0.01 pc) radio H II regions, which indicate the deeply embedded sites of high-mass star formation.

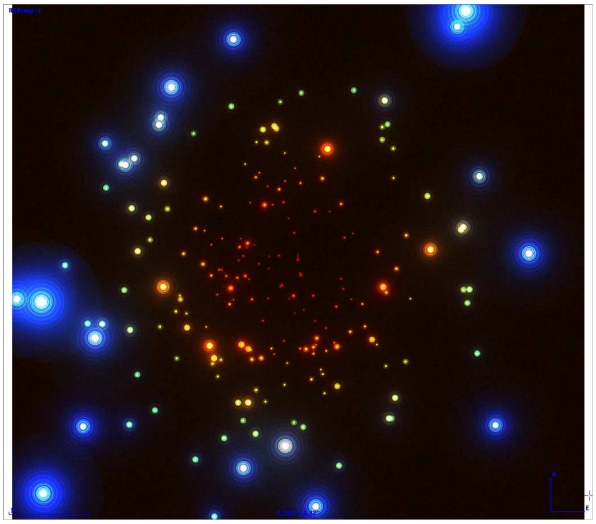

Figure 9.1: K, L, M composite image of a dense massive cluster in the Galactic centre.

Once we know the extinction we can correct the K-band magnitudes and can determine the luminosities and masses of the bright central cluster stars. These massive stars are born on the main sequence and are not dominated by accretion luminosity anymore (Zinnecker & Yorke 2007). They typically originate in dense, crowded stellar clusters or Trapezium systems, with linear star-to-star separations of the order of ∼1000–5000 AU (inferred from aperture synthesis radio observations of ultra and hyper-compact H II regions). At the typical distances to massive star forming regions of 2–10 kpc, such linear separations translate into angular separations of 0.1–0.5 arcsec. Considering that there are up to a dozen embedded massive stars in such dense proto-cluster clumps implies a central surface density of 20 to 1000 objects per arcsec^2! To convincingly resolve such crowded stellar systems, we need at least 10 mas angular resolution observations, the same as ALMA in the sub-mm regime. Thus, these observations need 5 times better angular resolution at 2 μm than JWST will provide. Note also that WFC3 on HST does not cover the required 2–5 μm domain but stops at 1.6 μm. HST/WFC3 may thus detect and resolve near-infrared (NIR) sources in less extincted (A_V = 50–100 mag), less crowded regions (at a later stage of evolution), while JWST can go to higher extinction but only in more nearby proto-clusters at 500–1000 pc, such as the Orion-KL region, not at 5–10 kpc, i.e. the Galactic centre (GC) region. The number of massive star formation sites is very limited in the solar neighbourhood and grows with distance squared. Thus, the E-ELT (together with ALMA) will be the telescope of choice when it comes to studying the variety of high-mass birthplaces in dense incipient clusters in the inner Galaxy. In the more nearby massive star formation regions (1–2 kpc) the angular resolution of the E-ELT may help to detect short-period binary and multiple systems by way of precise astrometric signatures in the reference frame of bright cluster stars and study their mass ratios. Close binary evolution of massive stars depends to a great deal on the initial binary mass ratio and is relevant as input for understanding the origin of low-mass and high-mass X-ray binaries. Finally, gravitational interactions will occur in such dense proto-clusters and some heavy-weights may be ejected at speeds in excess of 50 km s^{-1} (Blaauw's runaway stars). 50 km s^{-1} is 10 mas/yr at 1 kpc or 1 mas/yr at 10 kpc; this again shows that we need 10 mas spatial resolution to catch these dynamical ejection phenomena in the distant clusters (e.g. near the GC).

To summarise: we plan to exploit the E-ELT near and mid-IR enhanced sensitivity and angular resolution to peer through huge amounts of dust extinction, and to take direct, nearly-diffraction limited IR images in order to determine the stellar density in the centres of deeply embedded, massive clusters (e.g. the progenitor of NGC 3603), before the gas has been expelled and the cluster hatches. This will allow us to infer if stellar collisions are likely in the formation of the most massive stars and perhaps, in combination with ALMA sub-mm observations, if such clusters form as a result of global collapse of molecular clouds.

As an example, we derive the scientific requirements for the E-ELT to resolve a cluster of four massive stars (Trapezium system) embedded in a dense, hyper-compact H II region with radius 1000 AU at 2 kpc distance (diameter 1″) behind 200 mag of visual extinction (22.5 mag extinction in the K-band at around 2 μm and 4.5 mag in the M-band at around 5 μm; see the relations given above). High proper motion (several mas/yr), including dynamical ejections at velocities in excess of 30 km s^{-1} due to close encounters (∼20 AU), in such multiple systems can be monitored over a 1–2 year timeline. We can also work out that we need to reach a limiting apparent magnitude of K = 28 mag (with signal-to-noise ratio S/N = 5 in 1 hour), to see deeply embedded early O-stars assuming an absolute O-star magnitude of M_K = -6 mag and a distance modulus of 11.5 (2 kpc). These requirements can be met by a giant telescope of 42 m diameter, if a natural NIR bright guide star can be found within 10–30″, perhaps in the adjacent compact H II region, on which the wavefront sensing can be performed. Alternatively, a laser guide star must be used.

Follow-up integral field spectroscopy of the incipient cluster will also be required. The first goal here would be to derive the extinction to the individual massive stars by measuring the flux ratio of the Brackett γ (2.17 μm) to Brackett α (4.05 μm) line, which changes with extinction from its intrinsic Menzel Case B value (1/3). Secondly, we expect to detect collimated H_2 jets (2.12 μm) if the stars are in a phase of active disk accretion. Also, CO bandhead emission and absorption (both the fundamental at 4.6 μm and the first overtone at 2.3 μm) must be searched for to determine the temperature of the warm gas near the stellar photosphere and its kinematics (rotation, infall). A velocity resolution of 30 km s^{-1} is required, i.e. $R = 10^4$. This would also allow us to detect radial velocity variations and orbital motion of massive close binaries, making use of NIR photospheric absorption lines. Many visible massive OB stars are double-

lined short-period spectroscopic binaries, but the question is whether they are born as close binaries or if they form by disk-assisted grazing collisions (Moeckel & Bally 2007).

The requirements on the integral field spectroscopy are as follows: we would hope to use 4000×4000 IR arrays, which at 5 mas/pixel provide a field of view FoV = 20″ for imaging, thus catching the wider cluster core (0.1 pc). For Integral Field Unit (IFU) spectroscopy, we can concentrate the unit on the central or other 2″×2″ fields, requiring 400×400 = 160 000 fibres. Given the desired spectral resolution of 10 000, we need 2000 pixels to disperse the whole K-band. If we use only 1/2 of the K-band (2.10–2.18 μ for H_2 and Brγ, and 2.29–2.31 μm for 2-0 CO bandhead), we can fit 4 spectra in one spectral row. As we have 4000 spectral rows, we can fit 16 k spectra or spaxels on one IR array detector, hence we need 160 000/16 000 = 10 IR chips to match our scientific requirements. Similar for the M-band. This will be feasible by the time E-ELT will be working.

Finally we mention that ALMA will have a similar resolution of 10 mas at sub-mm wavelengths as the E-ELT in the near and thermal IR. Hence, while E-ELT will be particularly good at detecting OB stars at the end of their main accretion phase, ALMA should see OB proto-stars, i.e. massive stellar objects in their early or main mass assembly phase. Together E-ELT and ALMA are a powerful combination to reveal one of the most hidden but most important secrets of stellar astrophysics: the origin of massive stars.

Our E-ELT measurements will be directed towards regions with signs of active massive star formation, such as the Orion BN/KL hidden source(s), Spitzer IR dark clouds with UCHII regions as well as hot molecular cores with methanol (or H_2O/OH) masers, all known to pinpoint massive stars in the process of formation. These measurements will have the following immediate objectives: (i) See the photospheres of deeply embedded stars and proto-stars and determine the stellar number density in these dense molecular cloud clumps. This is important to estimate the likelihood of stellar collisions, a process under discussion for the growth of very massive stars (Bonnell, Bate & Zinnecker 1998). (ii) Check if the cluster centre is mass segregated at birth, i.e. whether all bright massive stars are actually located near the cluster centre, a profound prediction of the theory of massive star formation by competitive accretions including the merging of sub-clusters (Bonnell, Bate & Vine 2003). (iii) From proper motion astrometric measurements at two epochs (1 year apart) find evidence for dynamical interactions, such as gravitational slingshot effects and stellar ejections at speeds characteristic of OB runaway stars, i.e. some 40–100 km s^{-1} corresponding to (1–2 mas/yr) at 4–8 kpc distances of typical massive star forming regions. (iv) Integral field spectroscopic follow-up, as described above, to complement the proper motion data with radial velocity information, also relevant to catch the close spectroscopic massive binary population at birth. It is the interactions with these “hard” binaries that give rise to the ejection of single massive stars from the cluster (e.g. de Wit et al. 2005 and runaway O-star references therein). (v) Try to observe the sub-fragmentation of a centrally-peaked, dense molecular core destined to form a massive star in its centre by identifying column density fluctuations on very small, sub-arcsec scales, using the NICE method of Alves, Lombardi & Lada (2007), which measures the near-infrared colour excess of background stars through the cloud core in question (here K vs. K_S, or even $K-L$). Such data would settle an ongoing debate between two schools of massive star formation: Dobbs & Bonnell (2005) vs. Krumholz & McKee (2005). The former authors predict that fragmentation occurs, the latter disagree and predict a monolithic collapse without fragmentation.

9.2 Goals of the DRM simulations

The main questions to be answered by the simulations would (ideally) be:

1. How deep can we reach in the K, L, M bands to observe the embedded stars in highly obscured proto-clusters? Can we penetrate 100 mag of extinction or even 200 mag? At which K-magnitude can we reach $S/N \sim 10$ photometry for stars in very young deeply embedded clusters near the GC? Or in more nearby clusters?

2. How does the clumpy nature of extinction affect the results? How to best model the clumpy extinction? How does the simulated embedded cluster appear for a centrally peaked extinction (compared to uniform extinction)?

3. How many stars can we retrieve from an input mass spectrum? How does severe extinction influence the various output colour-magnitude diagrams (CMDs)? Which CMDs are most useful? K vs. $K-L$?
4. Which S/N can we reach for high to medium resolution spectroscopy ($R = 10\,000$) to classify the embedded, heavily obscured massive stars (i.e. $A_V \geq 100$ mag)?
5. Can we detect close massive binaries through K-band spectroscopy of their (fast, $\sim$200 km s^{-1}) orbital motion?
6. Which accuracy can we hope to achieve for K-band astrometry (to estimate stellar proper motions)?
7. Can we estimate the extinction to the embedded ultra-compact H II region, by measuring the ratio of recombination lines (i.e. Brγ vs. Brα)?

9.3 Metrics / figures of merit

The input for the imaging simulations will be catalogues of stellar positions and magnitudes, according to the extinction distribution and the amount of mass segregation assumed.

The simulations will need to perform source detection and photometry, producing output catalogues of recovered stellar positions and magnitudes. A study of the stellar completeness will be performed and results will be compared for the different assumed cases.

The main tool to be used in deciding whether or not the proposal achieves its goals will be the comparison between the input and output luminosity function in each of the K, L and M bands, and the magnitude scatter ($\sqrt{\frac{1}{N}\sum(\mathrm{input}-\mathrm{recovered})^2}$) for the different reddening distributions. We will then be able to check how much the extinction affects the results.

K vs. $K-L$, K vs. $K-M$ and L vs. $L-M$ CMDs will be produced and compared for the different simulation runs.

The output of the spectroscopic simulations will be the recovered 1D and 3D spectra for stars in the young embedded cluster at a reference band (K, 2.1–2.2 μm). The S/N of a given reference line (e.g. Brγ) will be estimated for stars with different masses and compared for different assumed extinction distributions in the cluster.

9.4 DRM simulations

9.4.1 Methodology

In order to take into account the points of Sections 9.2 and 9.3 we split the imaging simulations in several cases assuming different extinction distributions and mass segregation possibilities.

After simulating images in the K, L, M bands of young dense massive clusters with different extinction distributions and mass segregation, we perform crowded field profile fitting photometry on the data. We then compare the obtained luminosity functions with the input ones, in order to determine the recovered fraction of stars. This will allow us to check if we would be able to estimate the stellar density in the centres of deeply embedded massive clusters, and how much mass segregation will affect our results. We will also produce different IR CMDs and estimate the limiting magnitude we are able to reach with reasonable exposure times.

We then simulate 1D spectra for cluster stars selected on the K-band simulated images. We consider only two cases, i.e. a cluster with a uniform extinction distribution with A_V = 100 and 150 mag, respectively, and no mass segregation. We then produce 3D spectra in the wavelength range 2.1–2.2 μm, the region where we expect to detect the Brγ line (2.17 μm) and collimated H_2 jets (2.12 μm) if present.

9.4.2 Pipeline

9.4.2.1 Imaging

We simulate images of young massive dense star clusters assuming different extinction distributions and different input parameters. At first, we need to generate a catalogue with positions and magnitudes in the K, L, M bands. For this purpose we adopt a distribution of stellar masses drawn from the Salpeter Initial Mass Function (IMF) truncated at the high mass end to 130 $M_{\odot}$ and at the low mass end to 8 $M_{\odot}$. The normalisation factor is chosen such that 200 objects are simulated and scattered at random according to a Gaussian spatial distribution with a full width of 6″. At the distance of the GC ($\mu_0 \sim 14$ mag) this is equivalent to 0.2 pc.

Masses are transformed to K, L, M absolute magnitudes by adopting the Geneva evolutionary tracks (Lejeune & Schaerer 2001), with solar chemical composition ($Z = 0.02$) and for an age of $\log t = 5.9$ yr (Schaller et al. 1992; Meynet et al. 1994). A distance modulus for the GC of μ_0 =14 mag is added to the absolute magnitudes, plus a random depth for each star drawn from a uniform probability distribution ranging from 0 to 0.1 mag.

We neglect field star and galaxy contamination given the small FoV we are interested in (3″ radius) and the high stellar density of these young massive clusters.

After selecting the filter we rescale the star positions according to the assumed pixel scale. We simulate images adopting the pixel scale of the corresponding PSF for that band. Due to the need for excellent on-axis AO performance and the small angular extent of these dense clusters at the distance of the GC, we use LTAO PSFs. These are optimally sampled, i.e. the diffraction limited core of the PSF is covered by 2–3 pixel.

In order to generate the images we adopt standard IRAF packages and the Large Binocular Camera Image Simulator (LBCImSim, Grazian et al. 2004). This simulator (see the LBC web page) has been developed to test the performance of the LBC, but it can be used to simulate optical and NIR images for any other telescope, such as the E-ELT. This simulator has been modified to work also in the MIR regime. We proceed in two steps:

1. We first generate an artificial image, in the standard FITS format, using the IRAF MKOBJECTS task, with one second exposure time and with no noise. The image will contain the sources convolved with the LTAO PSF.

2. We then feed this image to the LBCImSim simulator that produces an output image adding the noise and other observational artifacts, such as system efficiencies, flat-field inaccuracies, sky background. It then simulates the observational strategy, i.e. the detector integration time (DIT) and the number of exposures (NDIT).

Images of the sky background are also generated following the same steps and then subtracted from each single scientific image. The resulting FITS format images are then ready to be analysed.

The data analysis is carried out using DAOPHOT IV/ALLSTAR (Stetson 1987, 1991) and ALLFRAME (Stetson 1994), and personal IDL routines. DAOPHOT is one of the most popular software package for stellar photometry. It was developed by Peter B. Stetson at the Dominion Astronomical Observatory, more than twenty years ago, and has undergone a continuous evolution and sophistication. DAOPHOT is used to identify the brightness peaks on the image and to perform synthetic aperture photometry. We then select a few bright, isolated and non-saturated stars to reconstruct an empirical PSF and use ALLSTAR to perform profile fitting photometry on the single images. All the images are at last analysed together with the task ALLFRAME (for more details see Section 9.5.2). We end up with a catalogue with positions, apparent magnitudes and relative errors, for each image. In order to create a unique star catalogue and estimate the weighted mean magnitudes we use DAOMATCH/DAOMASTER (Stetson 1994).

We are now able to estimate the recovered fraction of stars by comparing the obtained luminosity function with the input one, and to estimate the magnitude scatter (see Section 9.5.2).

The outline of the simulation pipeline for the imaging case is given in Fig. 9.2.

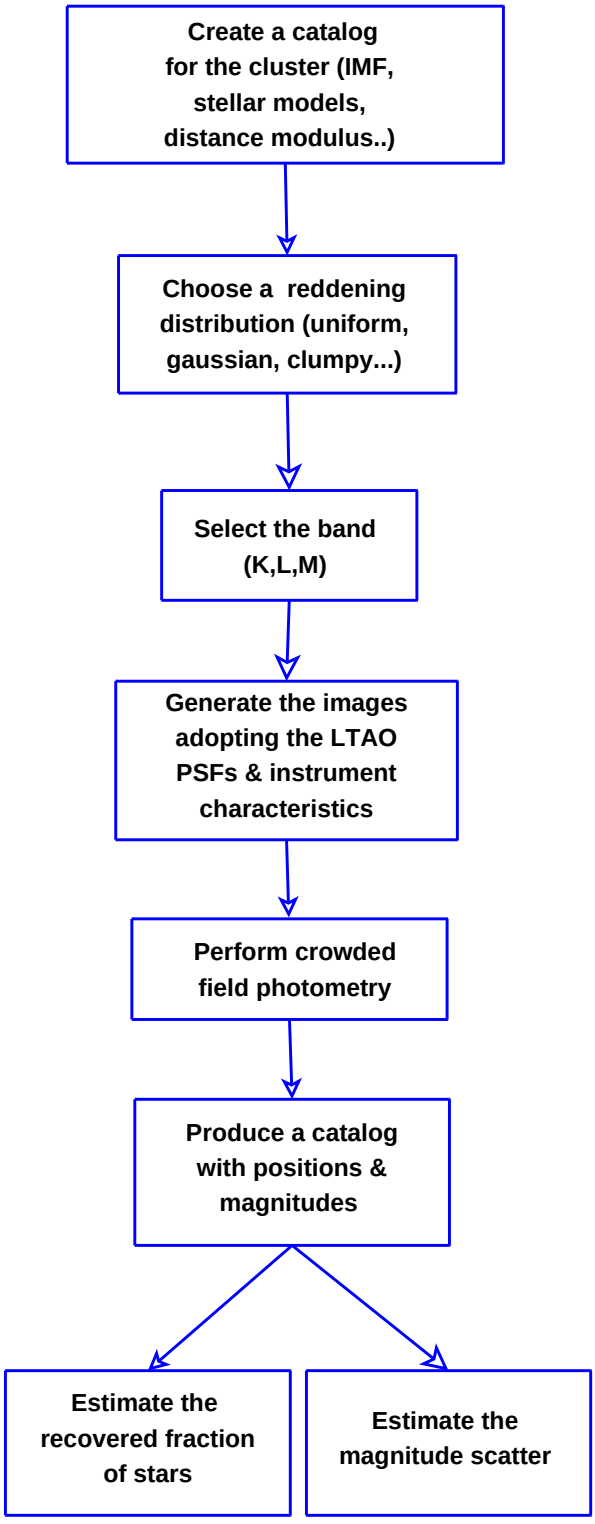

Figure 9.2: Flowchart of the imaging simulations.

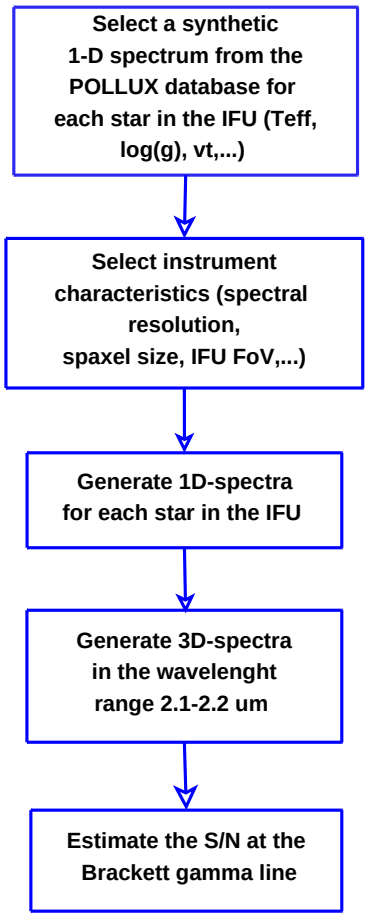

Figure 9.3: Flowchart of the spectroscopic simulations.

9.4.2.2 Spectroscopy

We need to simulate integral field spectra of massive stars in young dense clusters in the GC, assuming different input parameters. The template 1D spectra are taken from the POLLUX database, operated at GRAAL (Université Montpellier II, CNRS) with the support of the PNPS and INSU (see Section 9.4.3.1 for more details).

We select a region of the star cluster on the simulated K-band images assuming an IFU with a FoV of $1.6'' \times 1.6''$ and a spatial resolution element (spaxel) of 40 mas. In the considered FoV thus fall 37 stars, for which we have masses, temperatures, gravities and absolute magnitudes derived adopting a Salpeter IMF and the Geneva stellar models (see Section 9.4.2.1). The apparent magnitudes have been derived by adding a distance modulus of μ_0 = 14.0 mag (GC) plus a random population depth between 0 and 0.1 mag and an individual extinction for each star according to the reddening distribution assumed (see Section 9.4.3.1 for more details).

The synthetic 1D spectrum for each star in the IFU is sorted by wavelength and the region of interest is selected (2.1–2.2 μm). Using a code developed by G. Battaglia (see Section 12), which we have adapted to our specific case, we then simulate the observed 1D spectrum.

First of all, the code starts from the template spectrum and normalises the flux according to the apparent magnitude of the object, the zeropoint and the full-width of the selected filter (see [RD1] for the filter characteristics). As a second step, the spectrum is multiplied by the atmospheric transmission, and then the atmospheric and telescope thermal emissions, the sky continuum and the emission lines are added according to the selected site and telescope/mirror parameters. The spectrum is multiplied also by the total system efficiency, including the telescope throughput, by the effective photon collecting area and the

exposure time, and then is convolved to the final selected spectral resolution (R = 10 000), with 2 pixel per spectral resolution element. The photon and read-out noise are added at the end. At the same time, a sky spectrum is simulated with the same resolution and adopting the same input parameters, and then is subtracted from the final spectrum.

The input parameters that the code takes into account and that can be modified are:

- telescope diameter and effective photon collecting area
- spaxel area
- spectral resolution
- mirror coating
- telescope site
- dark current and read-out noise
- instrument efficiency
- exposure time
- airmass

In spite of the adoption of the LTAO PSF, ∼50% of the star flux falls indeed outside of the central spaxel (see Table 9.2). In the simulations, therefore, we rescale the object flux to the percentage of encircled energy (EE) falling in the reference radius of the assumed spaxel (i.e. for a spaxel of 40 mas, R ∼ 28 mas). Since there are 37 stars in the considered FoV of the IFU, we estimate the fraction of the star flux that falls in each of the 1600 spaxels. We assume that this flux is ≈ 0 when R > 85 mas, i.e. the radius of a region of 3×3 spaxels centred on the star. The EE values at different radii are estimated adopting the profile fits of the LTAO K-band PSF described in [RD1] (see Section 9.4.3.2). We then add up in each spaxel of the IFU the fluxes coming from all the stars. We are now able to generate the final datacube adding up the simulated 1D spectra according to the star positions in the IFU and their flux.

The outline of the simulation pipeline is given in Fig. 9.3.

9.4.3 Inputs

9.4.3.1 Scientific data

Below we list the scientific input for the imaging simulations:

- A distribution of stellar masses drawn from a Salpeter IMF truncated at 130 $M_\odot$ at the high mass end and at 8 $M_\odot$ at the low mass end. The normalisation factor is chosen such that 200 stars are simulated.
- Masses are transformed to K, L, M absolute magnitudes using the Geneva evolutionary tracks for a solar chemical composition ($Z = 0.02$) and an age of $\log(t/\mathrm{yr}) = 5.9$ (see Fig. 9.4). A distance modulus of μ_0 = 14 mag plus a random population depth between 0 and 0.1 mag is added. We adopt the reddening law of Rieke & Lebofsky (1985) and we obtain $A_K = 0.11\,A_V$, $A_L = 0.06\,A_V$, and $A_M = 0.02\,A_V$. A uniform background of $K \sim$ 21–22 mag arcsec^{-2} is also added to the images, in order to take into account the "sea" of 1 $M_\odot$ stars behind an extinction of A_V = 15 mag.
- The simulated stars are scattered at random according to a Gaussian spatial distribution with a full width of 6″ (0.2 pc at GC distance). We also simulate clusters with mass segregation, assuming that all stars with masses larger than 30 $M_\odot$, in the first case, and 20 $M_\odot$, in the second case, have a probability of 50% to be in the cluster centre, i.e. inside the inner 0.1 pc. The mass segregation cases are simulated only for a uniform extinction distribution with A_V = 150 mag.
- Extinction distributions:
  1. A constant extinction is assumed for the whole cluster. In particular, we consider three cases: A_V = 100, 150 and 200 mag.
  2. A Gaussian extinction profile with a peak of A_V = 150 mag, falling off with a FW = 6″ is simulated.

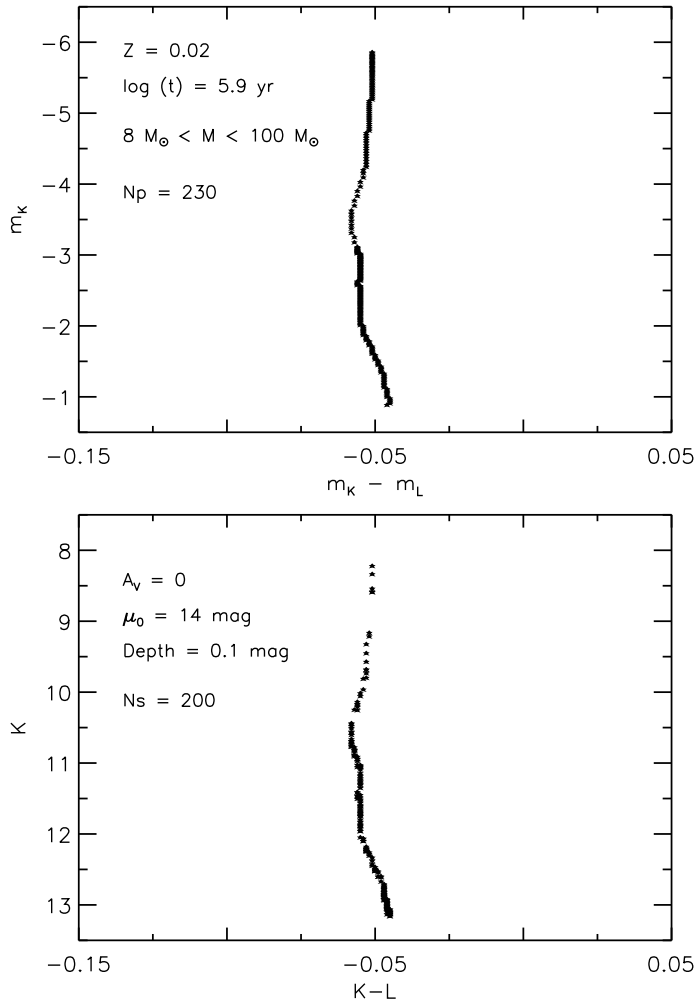

Figure 9.4: Top: isochrone from the Geneva database (Schaller et al. 1992; Meynet et al. 1994) with an age of $\log(t/\mathrm{yr}) = 5.9$ and a solar chemical composition plotted on the $M_K, M_K - M_L$ CMD. The isochrone has been cut to the range 8–100 $M_\odot$. Bottom: $K, K - L$ CMD of the 200 simulated cluster stars. Masses are distributed according to a Salpeter IMF. A distance modulus of μ_0 = 14 mag and a random population depth between 0 and 0.1 mag have been added to the absolute magnitudes.

3. A clumpy extinction on top of a uniform extinction of A_V = 150 mag is assumed. Clumps of ≈ 0.01 pc in the GC region would introduce a fluctuating spatially random extinction component in the cluster with an amplitude of an additional A_V = 50 mag over scales of 0.25″, for a total of 576 clumps in the 6″×6″ FoV (see Fig. 9.5).
4. A clumpy extinction on top of a Gaussian extinction profile with a peak of A_V = 150 mag and a FW = 6″. Clumps have an amplitude of A_V = 50 mag over scales of 0.25″.

The science input for the spectroscopic case are the K-band images simulated for different reddening distributions and template spectra for hot OB stars.

The template spectra are taken from the POLLUX database. High resolution synthetic spectra have been computed using the best available models of atmosphere (CMFGEN, ATLAS and MARCS), spectral synthesis codes (CMF_FLUX, SYNSPEC and TURBOSPECTRUM) and atomic linelists from the VALD database. Spectral types from O to M are represented for a large set of parameters: $T_{\rm eff}$, $\log g$, [Fe/H] (for more details visit the POLLUX web site).

Fig. 9.6 shows a synthetic input spectrum for a star with $T_{\rm eff}$ = 35 000 K, $\log g = 4.25$, v_t = 10 km s^{-1} and R = 100 000 in the K-band wavelength range.

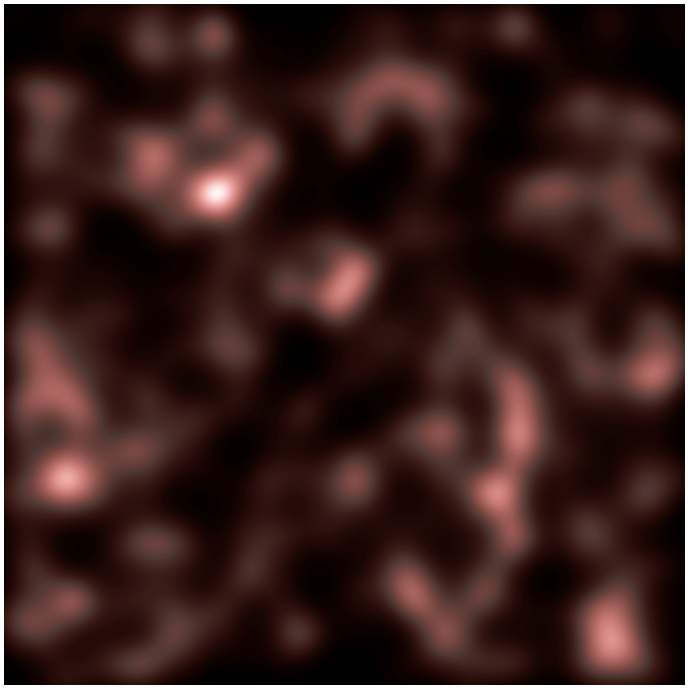

Figure 9.5: Clumpy medium simulated for a young massive dense cluster with an amplitude of A_V = 50 mag over scales of 0.25″. The total FoV is 6″×6″.

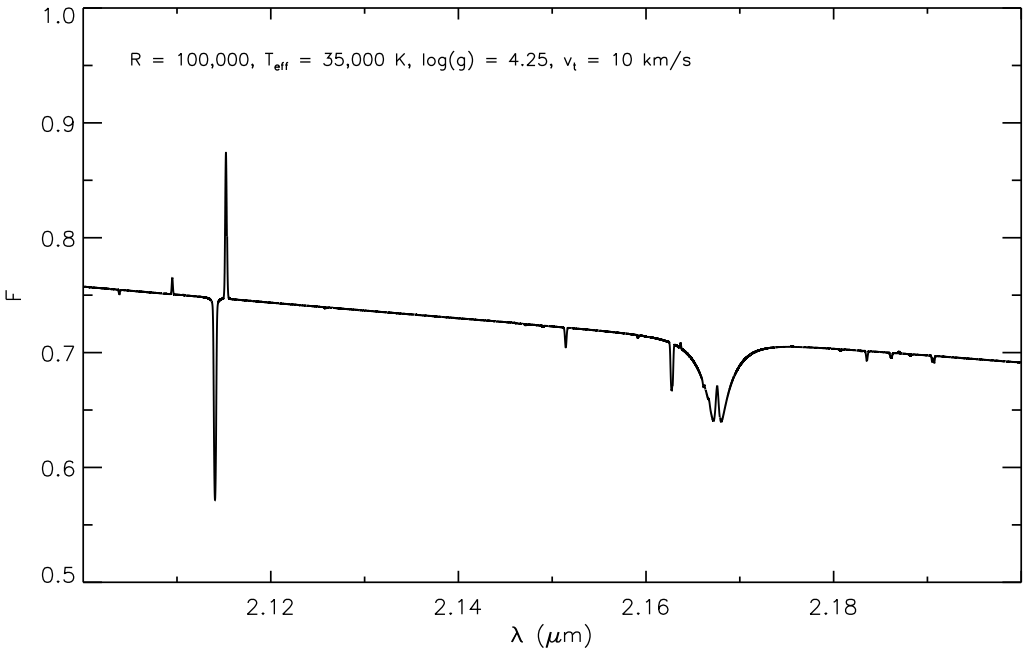

Figure 9.6: Synthetic spectrum from the POLLUX database with R = 100 000 for a star with $T_{\rm eff}$ = 35 000 K, $\log g$ = 4.25, v_t = 10 km s^{-1}, in the 2.1–2.2 μm wavelength range.

9.4.3.2 Technical data

The technical data used for the simulations were mainly taken from [RD1].

Site and background: We use the background model for a Paranal-like site as described in [RD1]. The zenith distance is assumed to be 0 for all simulation runs. For the imaging we are only concerned with broad-band background surface brightness values. We adopt the values of K = 13, L = 5.3 , M = 1.3 mag arcsec^{-2} from [RD1]. For the spectroscopy we follow the background model of [RD1].

Telescope: As described in [RD1]. We assume the protected silver/aluminium coating.

Instrument: For the imaging we assume two diffraction limited imagers, operating in the wavelength range 0.7–2.5 μm for the K-band simulations, and in the range 3–5 μm for the L and M-band simulations. We adopt the LTAO PSF pixel scales, i.e. 3.5 mas/pixel for the K, 5.6 mas/pixel for the L, and 7.7 mas/pixel for the M-band.

For the spectroscopy we assume an IFU instrument with a spaxel size of 40 mas. The IFU has a FoV = 1.6″×1.6″, requiring 40×40 fibres. In order to disperse half of the K-band ($2.1 < \lambda < 2.2$ μm) at the spectral resolution R = 10 000 we then need 1000 pixels.

Instrumental efficiencies were not available at the time this work was carried out. We choose to use constant values of 0.6 and 0.25 for the total telescope and instrument efficiencies (including the detector QE) for the imaging and IFU instruments, respectively.

The read-out noise (RON) of the detector was assumed to be 3 and 200 e$^-$/pix for the NIR and MIR, respectively, as in [RD1].

Adaptive Optics and PSFs: Since this science case requires high spatial resolution we have assumed LTAO for the simulations. In each band we only use the on-axis PSF, assuming that it does not vary significantly across the apparent target diameter of ∼6″. We use the analytic PSF fits described in [RD1] instead of the actual PSF images. These have the advantage of being free of speckles and improving the speed of the simulation code. Table 9.2 lists the energy contained in a single spaxel of 40 mas and in a region of 3×3 spaxels centred on the star. We assume that the contribution of the star flux to spaxels of the IFU at distances larger than $\sqrt{2} \times 1.5\times$ (spaxel size) is negligible.

9.4.4 Outputs

The simulation data products for the imaging case are images in the standard FITS format and tables with the stellar positions and recovered magnitudes in the different bands.

For each simulation run we generate a set of images, according to the S/N that we want to achieve. These images are dithered by a random number of pixels ($\leq$ 40 pix) in the X and/or Y direction. An image of the sky background is subtracted from each simulated scientific image. We then stack all the images of a set in a median in order to increase the S/N of the faint highly reddened stars, making their detection more straightforward. The left panel of Fig. 9.7 shows a zoom of the median of 25 K-band images ($t_{\rm exp}$ = 1 h) of the simulated young massive dense cluster for the case of a uniform reddening distribution with A_V = 150 mag. The right panel shows a zoom of the median of the subtracted images after two ALLFRAME runs (see Section 9.5.2 for more details).

For each simulation run we then produce a catalogue with star positions and magnitudes to be compared with the input one. We also generate RGB images combining three K, L and M-band images and rescaling them to the best spatial resolution, i.e. to the K-band image. Fig. 9.8 shows an RGB image for the case of a Gaussian reddening distribution.

The simulation data products for the spectroscopic case are 1D spectra for each star present in the assumed FoV and 3D spectra in the wavelength range 2.1–2.2 μm for two reddening distributions. Fig. 9.9 shows a 1D spectrum for a star with $T_{\rm eff}$ = 35 000 K and $\log g$ = 4.25, $M = 22\ M_\odot$, and K = 22 mag, simulated assuming a Paranal-like site, a silver/aluminium mirror coating, and an exposure time of $t_{\rm exp}$ = 1200 s.

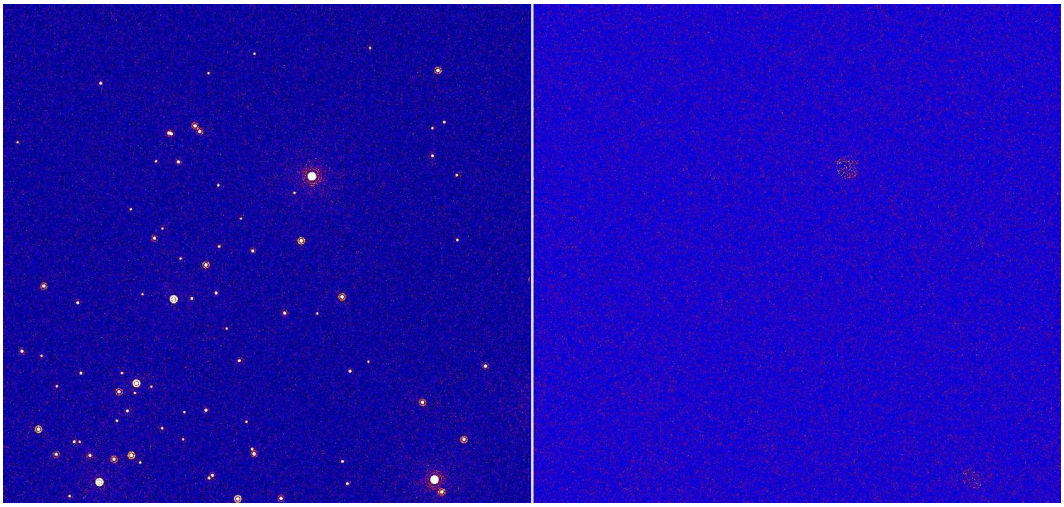

Figure 9.7: Left: Zoom of a K-band median image of the simulated massive dense cluster for a total exposure time of 25 h, a pixel scale of 3.5 mas/pixel, and a uniform reddening distribution with A_V = 150 mag over the whole cluster. Right: Zoom of the median of the subtracted K-band images for the same parameters.

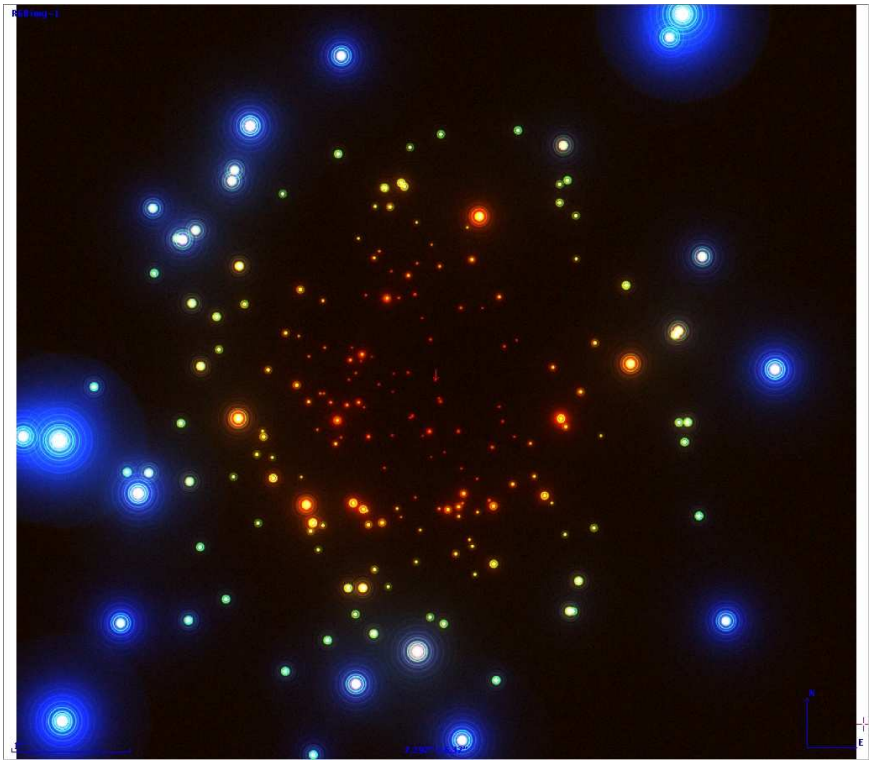

Figure 9.8: Simulated RGB image of a massive dense cluster for a Gaussian reddening distribution with a peak of A_V = 150 mag and a FW = 6″.

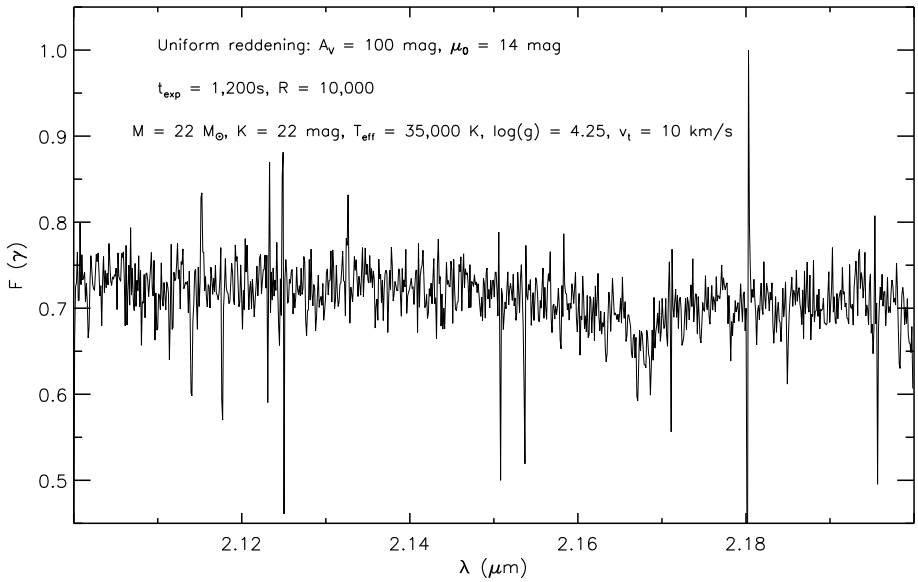

Figure 9.9: Sky subtracted simulated 1D spectrum for a star with $M = 22\ M_{\odot}$, $K = 22$ mag, $T_{\rm eff}$ = 35 000 K, $\log g$ = 4.25, and v_t = 10 km s^{-1} in the 2.1–2.2 μm wavelength range. The star's magnitude was derived adopting a Geneva isochrone for $\log(t/{\rm yr})$ = 5.9 and $Z = 0.02$, and assuming a distance modulus of μ_0 = 14 mag and a uniform reddening distribution with A_V = 100 mag.

9.5 Results of simulations

9.5.1 Simulation runs

The different simulation runs for the imaging case and their input parameters are summarised in Table 9.1.

In our imaging simulations we first test how the presence of different extinction patterns in a young massive dense star cluster affects the observations. Furthermore, we include two degrees of mass segregation to check if we would be able to detect its presence in a highly extincted dense cluster.

We use the E-ELT Imaging Exposure Time Calculator [RD1] to estimate the best observing times required to reach the faintest limiting magnitudes (in the case of a uniform reddening distribution with A_V = 150 mag, the faintest star has K = 29.7, L = 22.2, and M = 16.4 mag), and not to saturate the brightest stars of the cluster in the two MIR bands (in the same case, the brightest star has L = 17.3, and M = 11.5 mag). For each reddening distribution case we then have to find the best observing strategy. The exposure times for the different simulation runs are listed in columns 5, 6 and 7 of Table 9.1.

The total observing time, for instance, for a massive dense cluster with a uniform reddening distribution with A_V = 150 mag is:

$$T_{\rm tot} = 25(K) + 1.1(L) + 0.05(M) \approx 26\ {\rm h} + {\rm sky} + {\rm overheads}. \tag{9.1}$$

The same observing time is adopted in the case of a clumpy reddening distribution, i.e. clumps of amplitude A_V = 50 mag over a uniform extinction A_V = 150 mag. It is worth noticing that in this case the faintest star has K = 49, L = 33, and M = 19.7 mag, but the brightest star has always L = 17.5, and M = 11.5 mag. Therefore, the observations are particularly challenging for the coexistence of very bright, mostly saturated stars in the images, with very reddened (the extinction can be as high as $A_V \sim 300$ mag, i.e. $A_K \sim 30$, $A_L \sim 20$, and $A_M \sim 6$ mag), and thus very faint stars.

The different simulation runs for the spectroscopic case and their input parameters are summarised in Table 9.2. In our spectroscopic simulations we test how the level of extinction in a young massive dense star cluster affects the observations.

Table 9.1: Input parameters for the different imaging simulation runs.

Band	Reddening[a]	A_V mag	Pixel scale mas/pix	DIT s	NDIT	$T_{\rm tot}$ h	Mass seg.[b]
K	Uni	100	3.5	2.7	30	0.1	—
K	Uni	150	3.5	60	60	25	—
K	Uni	150	3.5	60	60	25	$M > 20\ M_\odot$[c]
K	Uni	150	3.5	60	60	25	$M > 30\ M_\odot$[c]
K	Gauss	150[d]	3.5	1.2	85	0.7	—
K	Gauss/Clu	150+50[e]	3.5	1.2	85	0.7	—
K	Clu	150+50[f]	3.5	60	60	25	—
L	Uni	100	5.6	0.5	80	0.1	—
L	Uni	150	5.6	1.33	120	1.1	—
L	Uni	150	5.6	1.33	120	1.1	$M > 20\ M_\odot$[c]
L	Uni	150	5.6	1.33	120	1.1	$M > 30\ M_\odot$[c]
L	Uni	200	5.6	1.56	115	1.25	—
L	Gauss	150[d]	5.6	0.55	110	0.42	—
L	Gauss/Clu	150+50[e]	5.6	0.55	110	0.42	—
L	Clu	150+50[f]	5.6	1.33	120	1.1	—
M	Uni	100	7.7	0.04	75	0.017	—
M	Uni	150	7.7	0.05	75	0.020	—
M	Uni	150	7.7	0.05	75	0.020	$M > 20\ M_\odot$[c]
M	Uni	150	7.7	0.05	75	0.020	$M > 30\ M_\odot$[c]
M	Uni	200	7.7	0.056	125	0.050	—
M	Gauss	150[d]	7.7	0.04	75	0.017	—
M	Gauss/Clu	150+50[e]	7.7	0.04	75	0.017	—
M	Clu	150+50[f]	7.7	0.056	125	0.050	—

[a] Assumed reddening distribution.
[b] Presence of mass segregation.
[c] Masses of stars that have a probability of 50% to be in the inner 0.1 pc of the cluster.
[d] Extinction of the peak of the Gaussian.
[e] Extinction of the peak of the Gaussian plus amplitude of the clumps.
[f] Uniform extinction plus amplitude of the clumps.

Table 9.2: Input parameters for the different 1D spectroscopy simulation runs.

$\delta\lambda$ μm	$R^a_{\rm in}$	$R^b_{\rm fin}$	$t_{\rm exp}$ s	A_V[c] mag	FoV arcsec^2	Spaxel size mas	EE (central spx)[d] %	EE (total)[e] %
2.1–2.2	100 000	10 000	1200	100	1.6×1.6	40	52.3	61.0
2.1–2.2	100 000	10 000	7200	150	1.6×1.6	40	52.3	61.0

[a] Spectral resolution of the synthetic input spectra.
[b] Spectral resolution of the simulated spectra.
[c] Constant extinction value for the whole cluster.
[d] Ensquared energy contained in the central spaxel of 40 mas.
[e] Ensquared energy contained in a region of 3×3 spaxels centred on the star.

9.5.2 Analysis

9.5.2.1 Imaging

The data reduction is performed with DAOPHOT IV/ALLSTAR/ALLFRAME and personal IDL routines following these steps:

- We first identify the brightness peaks on each image: a threshold must be defined in order to choose a level of significance of the brightness enhancement. This threshold depends on the RON (3 e^-/pix for the K-band, or 200 e^-/pix for the L and M bands) of the detector, and the mean sky value of the image. These quantities are used to define the random error per pixel, σ:

$$\sigma = \sqrt{\overline{\mathrm{sky}} + 2\,\mathrm{RON}} \tag{9.2}$$

  A scaling factor decided by the user fixes the threshold for the star search. This is done by the routine FIND, which performs a rough estimate of the centroid and flux of the stars. To avoid the inclusion of non-stellar objects (extended sources, cosmic rays, cosmetic defects) in the catalogue, DAOPHOT defines two shape parameters, the *sharpness* and the *roundness*, that can be used for a first selection of the sources. This selection is particularly useful to reject fake detections on the images, such as the diffraction rings of the brightest stars (see Fig. 9.8).

- We then perform aperture photometry with the routine PHOT on all the objects detected, estimating the flux in an aperture of radius 5–6 pixels. This radius must be small enough to contain negligible contamination by the light from neighbouring stars. The contribution of the local sky background is estimated in a circular annulus around the object (r_1=7, r_2=20 pixel) under the assumption that the spatial scale variation of the background is smaller than the aperture, and then subtracted. There are two obvious drawbacks affecting this approach: (i) the crowding of the observed massive dense clusters means that the star profiles overlap, so that it is very difficult to take into account the contribution of nearby objects; (ii) the measurement of the more reddened stars in the clusters, characterised by a very limited number of counts, is dominated by the background error, fluctuations and statistics.

- We thus need to perform profile fitting photometry. We select a few bright, isolated and non-saturated stars in order to reconstruct an empirical PSF. DAOPHOT models the PSF with a mixed technique, in the sense that it uses the analytic formula plus a numerical map of the residuals. A Moffat function with β=2.5 is used for the analytical part, using a fitting radius of 2–3 pixels depending on the image pixel scale. We use this value in order to decrease the effect of the crowding, especially when fitting the very faint objects. The PSF is defined in a region with an area of radius 20 pixels (PSF radius, see Table 9.1). Given the list of selected PSF stars, the routine PSF fits a simple analytic formula with non-linear least squares. Then, all the observed residuals from the fit are interpolated on a common grid and averaged into a numerical look-up table of residuals. We assume that the PSF is constant all over our field of view (6″×6″). The comparison between the image of the empirical PSF and of a star for the case of the K-band with uniform background is shown in Fig. 9.10. It is worth noticing that the technique of the analytic function plus the numerical matrix works very nicely in reconstructing the PSF. On the other hand, our PSF does not perfectly reproduce the shape of the stars in the L and M band images: a central brightness peak is left in the subtracted image. This residual flux is of the order of 2–3% and we take it into account when estimating an aperture magnitude correction (see discussion below). The result could be improved, for example, feeding directly to DAOPHOT the analytical fit of the simulated PSF.

- We use this PSF and the aperture magnitude catalogue to perform profile fitting photometry on the single images with the task ALLSTAR. This program tries to exploit all the information available, performing a simultaneous fit of all the sources. During each iteration, it subtracts all the stars from a working copy of the image, according to the best estimates of the position and the magnitude of each object, computes the increments to the positions and magnitudes of the subtraction residuals around each position, and then checks each star to see whether it has converged or not. When a star has

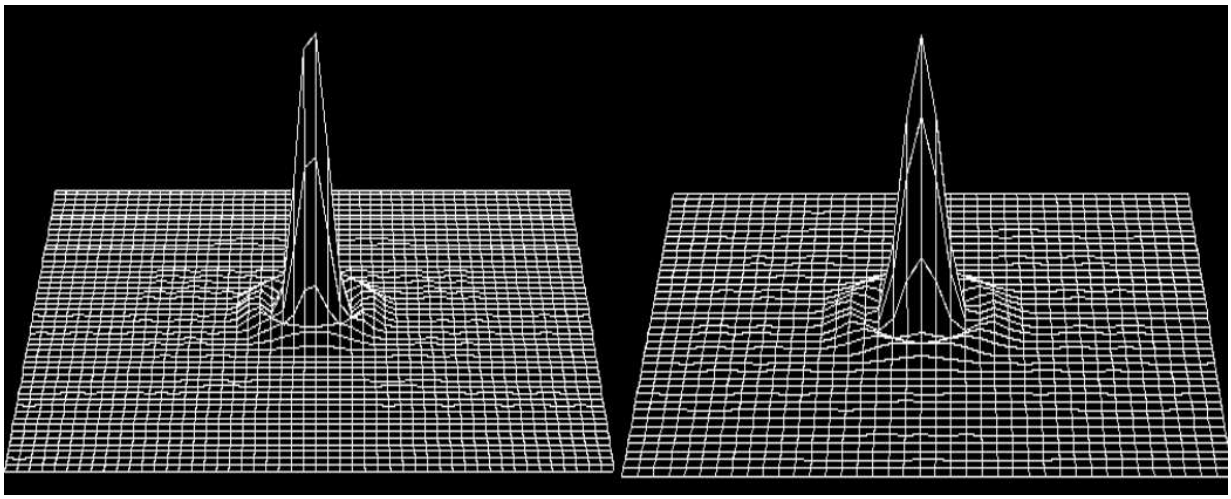

Figure 9.10: Left: 3D-profile of a star in a K-band image. Right: 3D-profile of the empirical K-band PSF reconstructed by DAOPHOT.

converged, it is permanently subtracted from the working image and the result is saved. A non-linear least square fitting technique is applied to derive the parameters of each object. ALLSTAR writes out also the final subtracted image, which can be visually inspected in order to check the quality of the fit and look for any faint lost or discarded objects.

- Once that all the images have been reduced with ALLSTAR, a single catalogue containing the position and the magnitude for each detected star is available for each frame. In order to produce a unique catalogue with common positions and weighted mean magnitudes we use the tasks DAOMATCH and DAOMASTER. These are two programs that estimate the coordinate transformations between different star catalogues and allow to rescale them to a single reference coordinate system. Once that all the star lists are rescaled, it is possible to create a catalogue with weighted mean magnitudes. We use this as an input catalogue to perform profile fitting photometry simultaneously on all the images with the task ALLFRAME. ALLFRAME starts with the master list of stars, which are in the x, y coordinate system of a reference image, inverts the coordinate of each star in each frame, and performs a local fit using the individual PSF. By applying a non-linear least squares fitting technique, during every iteration a new estimate of the position and magnitude of each star is provided. The position is rescaled in the reference frame coordinate system, and then averaged. This ensures that, in case the transformations are accurate enough, the centroid of the stars is known with great precision. This is probably the key point of this package: the accuracy in the centroid determination allows an optimal fit of stellar profiles, which ensures precise estimates of the magnitudes. Moreover, by using the same list of stars for all the frames, it is possible to reach a deeper magnitude limit, in the sense that the a priori knowledge of the existence of a star in a specific position makes it possible to fit it even if the S/N is very poor and, therefore, the fit on the individual image has failed. ALLFRAME writes out the final subtracted images too.

- After the ALLFRAME run, we stack all the subtracted images in a median image, and perform a search for any remaining stars in it. The S/N ratio of faint objects is raised, making their detection much easier. Unfortunately, in the search for these very faint stars, we also identify many spurious objects. In order to reject them, we run a program, SEPARATION (Stetson et al. 2003), that estimates the degree of crowding of each star of the catalogue. In most cases, we select all stars with *sep* > 5, i.e. stars that have required a correction of less than 1% for light contributed by known neighbours. The newly identified objects are then added to the input catalogue and ALLFRAME is run again. This procedure can be repeated until no more detectable objects, according to the assumed FIND threshold, are found in the subtracted median image.

In order to compare the final photometric catalogue with the input one, we have to transform the recovered PSF instrumental magnitudes to absolute magnitudes. We first estimate the aperture correction selecting a few bright and isolated stars on the frame. The corrections used in our simulation runs are listed in the fourth column of Table 9.3.

After applying the aperture correction we have to correct the flux for the fraction of energy outside of the PSF area. The EE values adopted in the different simulation runs are listed in Table 9.3. We interpolated the values given in [RD1] for the LTAO PSFs to obtain the EE corresponding to our PSF area.

The result of the process described above is a catalogue with stellar positions, input and recovered magnitudes for each simulation run.

Table 9.3: EE and aperture corrections adopted for the different bands.

Band	Pixel scale mas/pix	PSF radius mas	Ap. corr.[a] mag	EE[b] %	EE corr.[c] mag
K	3.5	64	-0.10	58.0	-0.60
L	5.6	101	-0.12	77.0	-0.28
M	7.7	93	-0.15	63.0	-0.50

[a] Aperture correction applied to the PSF magnitudes.
[b] Fraction of energy enclosed in the PSF area of radius equal to the PSF radius.
[c] EE correction applied to the PSF magnitudes.

9.5.2.2 Spectroscopy

Table 9.4 lists the parameters of the 37 stars included in the assumed IFU for the two simulation runs. Masses have been estimated adopting a Salpeter IMF (see Section 9.4.3.1) and converted to absolute magnitudes and colours using a Geneva isochrone for $\log(t/\mathrm{yr})$ = 5.9 and $Z = 0.02$ (see Section 9.4.3.1). Apparent magnitudes were derived assuming a distance modulus for the GC of μ_0 = 14 mag, a random population depth between 0 and 0.1 mag, and a uniform reddening distribution of A_V = 100 or 150 mag over the whole cluster.

Fig. 9.11 shows the position of the IFU on a simulated K-band image for the case of a uniform reddening distribution of A_V = 100 mag. The FoV contains a sample of 37 stars, the most massive having 100 $M_\odot$.

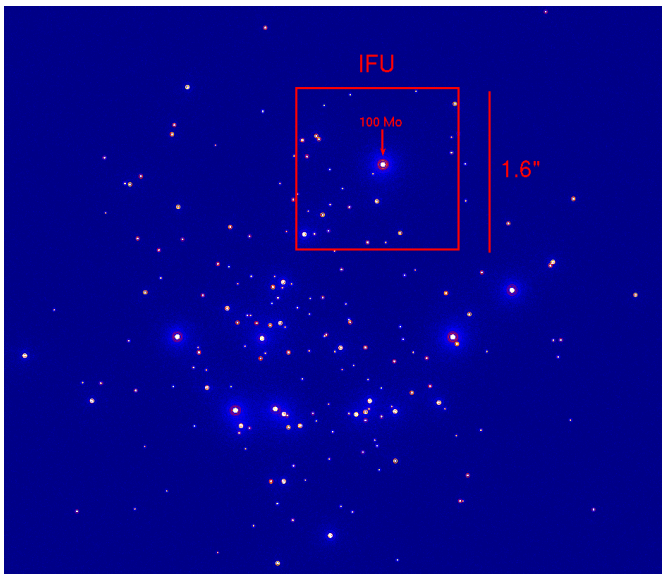

Figure 9.11: Position of the IFU on a simulated K-band image for the case of a uniform reddening distribution of A_V = 100 mag. The IFU is centred on a 100 $M_\odot$ star and has a FoV = 1.6″×1.6″.

Table 9.4: Parameters of the 37 stars included in the IFU.

ID[a]	K^b (A_V = 100) mag	K^c (A_V = 150) mag	Mass $M_\odot$	$T_{\rm eff}$ K	$\log g$
200	19.23	24.68	100.63	48 500	4.00
4	24.15	29.62	8.22	27 500	3.75
5	24.09	29.65	8.22	27 500	3.75
10	24.12	29.56	8.32	27 500	3.75
16	24.06	29.53	8.52	27 500	3.75
18	24.04	29.53	8.62	27 500	3.75
28	23.96	29.45	8.82	27 500	3.75
30	24.02	29.47	8.82	27 500	3.75
33	24.01	29.43	8.92	27 500	3.75
40	23.87	29.41	9.22	27 500	3.75
42	23.90	29.42	9.32	27 500	3.75
48	23.87	29.37	9.42	27 500	3.75
50	23.81	29.31	9.52	27 500	3.75
59	23.80	29.28	9.92	27 500	3.75
62	23.75	29.24	10.02	27 500	3.75
66	23.59	29.12	10.52	27 500	3.75
69	23.52	29.10	10.82	27 500	3.75
80	23.47	28.91	11.42	27 500	3.75
84	23.32	28.86	12.12	27 500	3.75
97	23.16	28.66	13.22	27 500	3.75
98	23.11	28.62	13.31	27 500	3.75
101	22.98	28.44	14.01	27 500	3.75
102	22.93	28.46	14.21	30 000	4.25
108	22.79	28.36	14.91	30 000	4.25
110	22.79	28.28	15.21	30 000	4.25
113	22.69	28.15	15.91	30 000	4.25
116	22.63	28.16	16.11	30 000	4.25
122	22.55	28.11	16.80	32 500	4.25
128	22.47	28.00	17.50	32 500	4.25
129	22.47	27.99	17.50	32 500	4.25
142	22.11	27.64	20.78	35 000	4.25
150	21.95	27.48	22.47	35 000	4.25
159	21.76	27.26	24.85	40 000	4.00
165	21.79	27.22	25.85	40 000	4.00
167	21.70	27.26	25.85	40 000	4.00
172	21.53	27.00	29.82	40 000	4.00
192	20.39	25.96	51.59	42 500	4.00

[a] Star identification number.
[b] Apparent K magnitude for a uniform reddening distribution of A_V = 100 mag.
[c] Apparent K magnitude for a uniform reddening distribution of A_V = 150 mag.

Table 9.5: S/N at the Brγ line ($\sim$2.17 μm) for a sample of stars in the simulated IFU for two reddening distributions.

Mass $M_\odot$	A_V[a] mag	K^b mag	S/N	$t_{\rm exp}$ s
100	100	19.2	>100	1200
30	100	21.5	$\sim$60	1200
13	100	23.1	$\sim$20	1200
8	100	24.1	$\sim$15	1200
100	150	24.7	$\sim$10	7200
30	150	27.0	$\sim$3	7200
13	150	28.6	—	7200
8	150	29.6	—	7200

[a] Constant extinction value for the whole cluster.
[b] Apparent K magnitude for the A_V value in the previous column.

First of all we analysed the 1D spectra of stars included in the IFU for the two simulation runs. We estimated the S/N using the IRAF routine SPLOT at the Brγ line (see Table 9.5). We are interested in this line since the flux ratio of the Brγ at 2.17 μm and the Brα at 4.05 μm would allow us to determine the extinction towards resolved stellar sources, i.e. massive stars, in these compact clusters (via Menzel Case B recombination line theory).

9.5.3 Compliance with figures of merit

9.5.3.1 Imaging

For each simulation run we produced a CMD in the different IR colours ($K-L$, $K-M$, and $L-M$) and a plot of the magnitude scatter $S = \sqrt{\frac{1}{N}\sum(\mathrm{input}-\mathrm{recovered})^2}$) between the input and recovered magnitude. We then estimated the number of recovered stars and compared it for simulations with different reddening distributions and mass segregation degrees.

Fig. 9.12 shows an RGB composite image obtained combining K, L, M-band images of a simulated young dense massive cluster for the case of a uniform reddening distribution with A_V = 100 mag. Figs. 9.13 and 9.14 are the same RGB image but for the cases of a uniform reddening distribution with A_V = 150 and 200 mag, respectively. Note that all the stars become fainter and redder. The noise increases too.

Figs. 9.15 and 9.16 show two RGB images for the case of a clumpy reddening distribution with an amplitude of A_V = 50 mag and a scale of 0.25″ over a uniform extinction of A_V = 50 mag and A_V = 150 mag, respectively. Note how some stars fade to invisibility if they happen to be behind the extinction clumps.

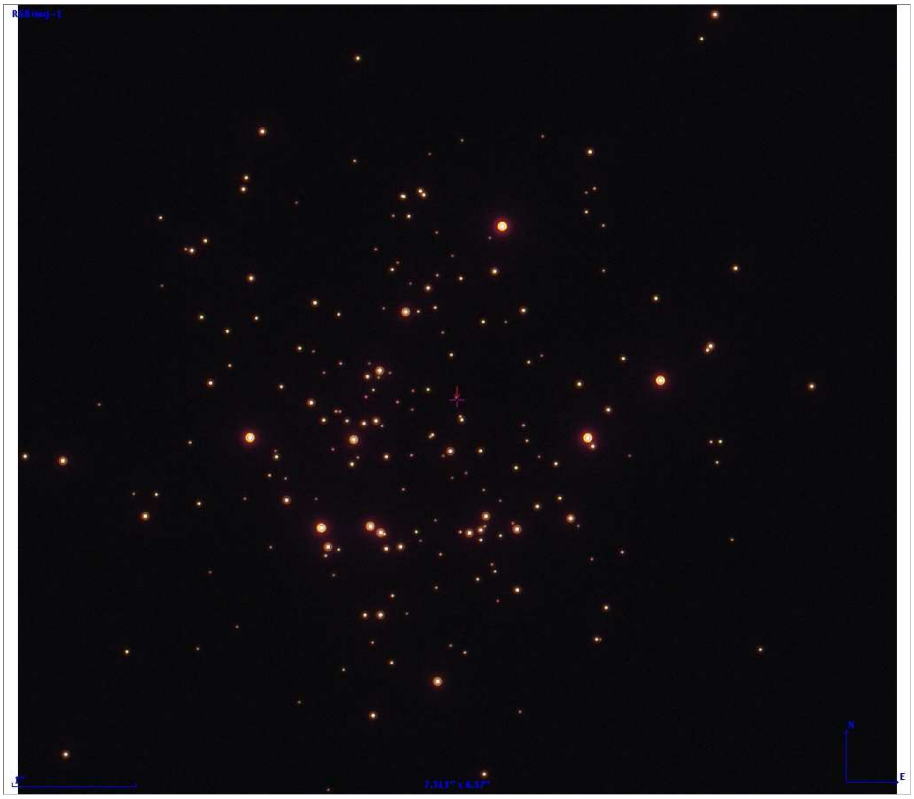

Figure 9.12: K, L, M-band RGB image of a simulated young dense massive cluster for a uniform reddening distribution of A_V = 100 mag over the whole cluster. The FoV is 7″×7″ and the faintest star in the image is $K \sim 24$, $L \sim 19$ and $M \sim 15$ mag.

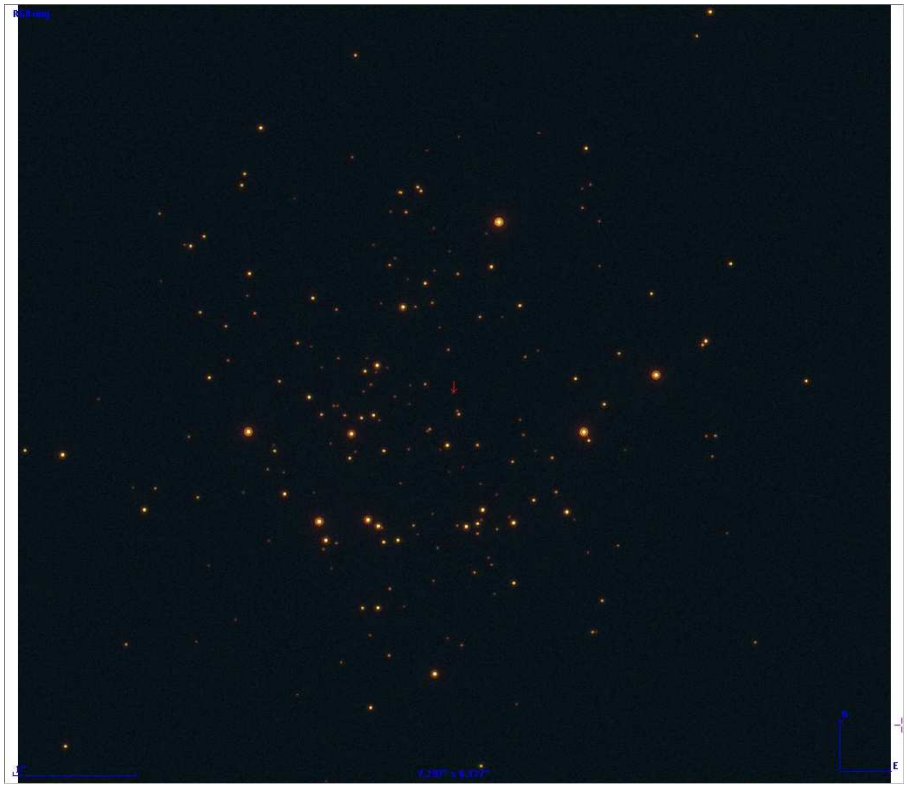

Figure 9.13: As Fig. 9.12 for A_V = 150 mag. The faintest star in the image is $K \sim 30$, $L \sim 22$ and $M \sim 16$ mag.

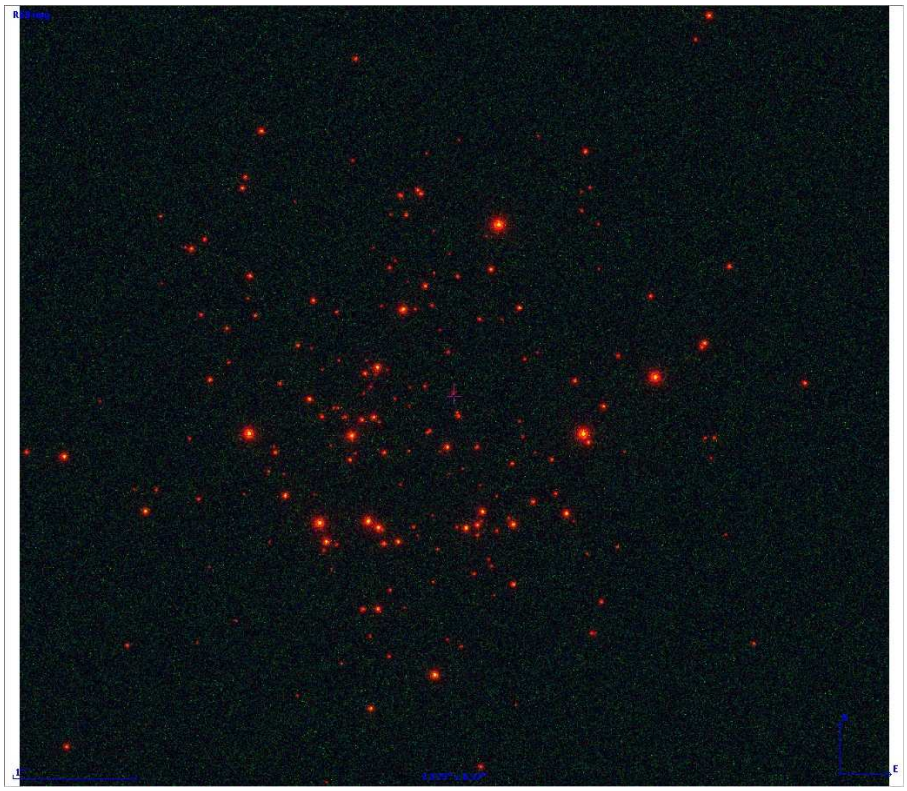

Figure 9.14: As Fig. 9.12 for A_V = 200 mag. The faintest star in the image is $K \sim 35$, $L \sim 25$ and $M \sim 17$ mag.

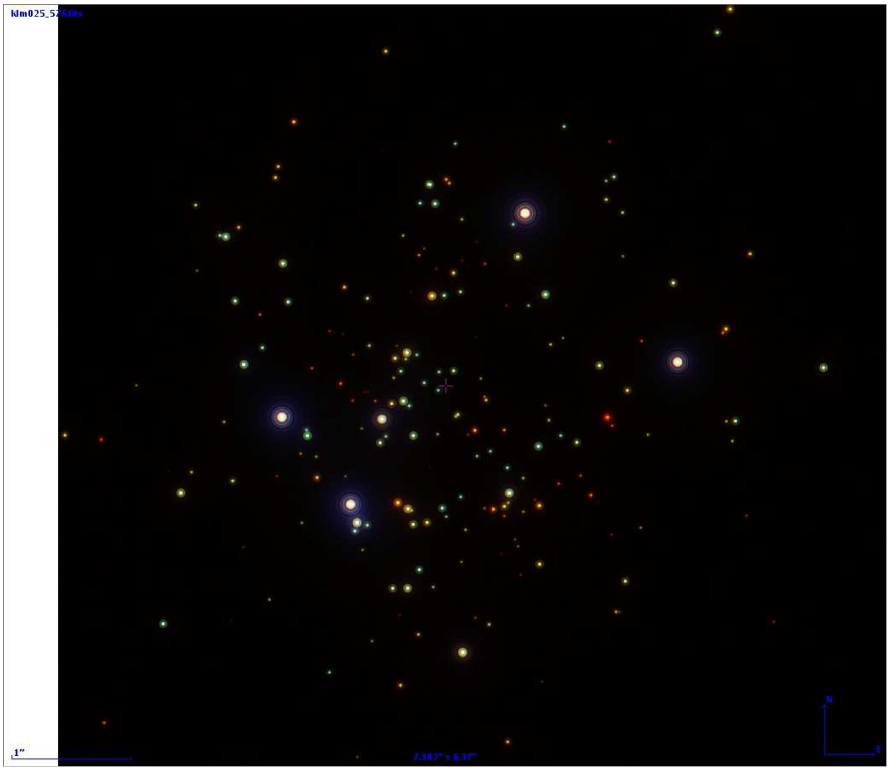

Figure 9.15: K, L, M-band RGB image of a simulated cluster with a clumpy reddening distribution with clumps of amplitude A_V = 50 mag and scale 0.25″ on top of a uniform extinction of A_V = 50 mag. The faintest star in the image is $K \sim 38$, $L \sim 27$, $M \sim 18$ mag.

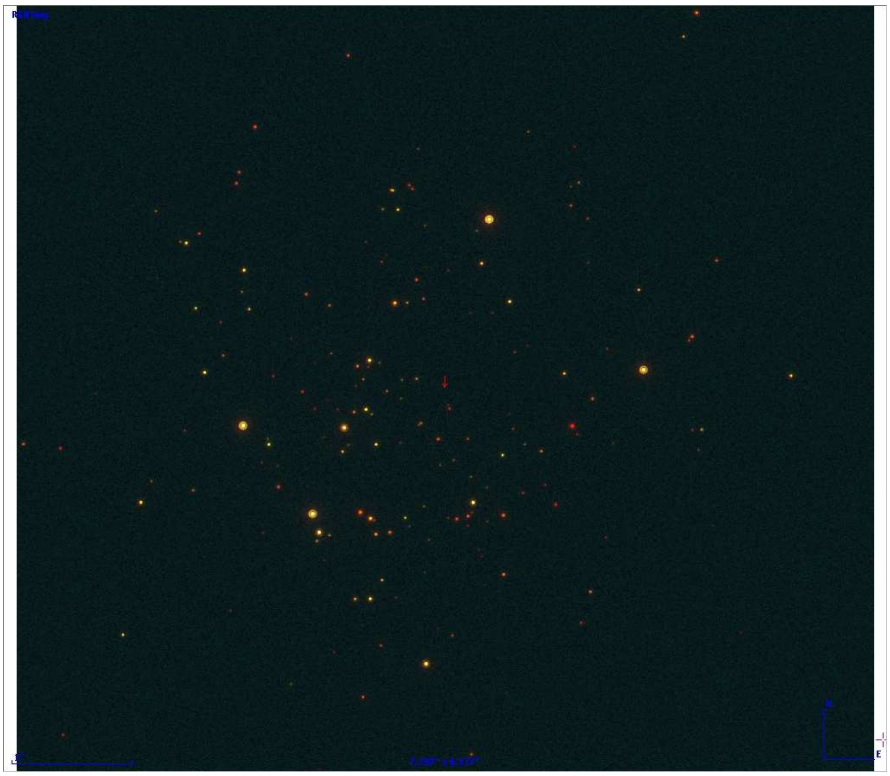

Figure 9.16: As Fig. 9.15 for a clumpy reddening distribution on top of a uniform extinction of A_V = 150 mag. The faintest star in the image is $K \sim 49$, $L \sim 33$, $M \sim 20$ mag.

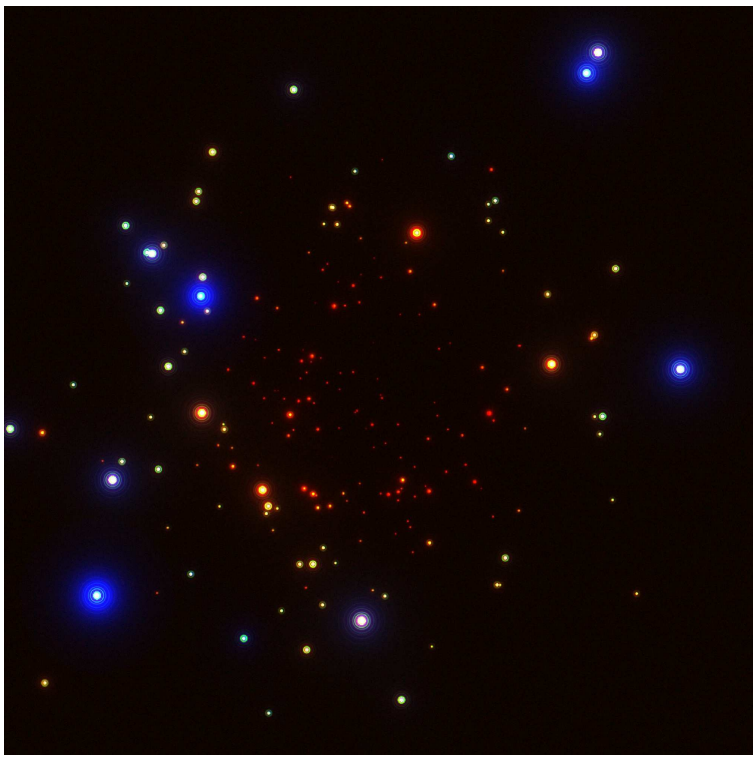

Figure 9.17: As Fig. 9.15 for a clumpy reddening distribution on top of a Gaussian reddening distribution with a peak of A_V = 150 mag and a FW of 6″. The faintest star in the image is $K \sim 43$, $L \sim 29$ and $M \sim 18$ mag.

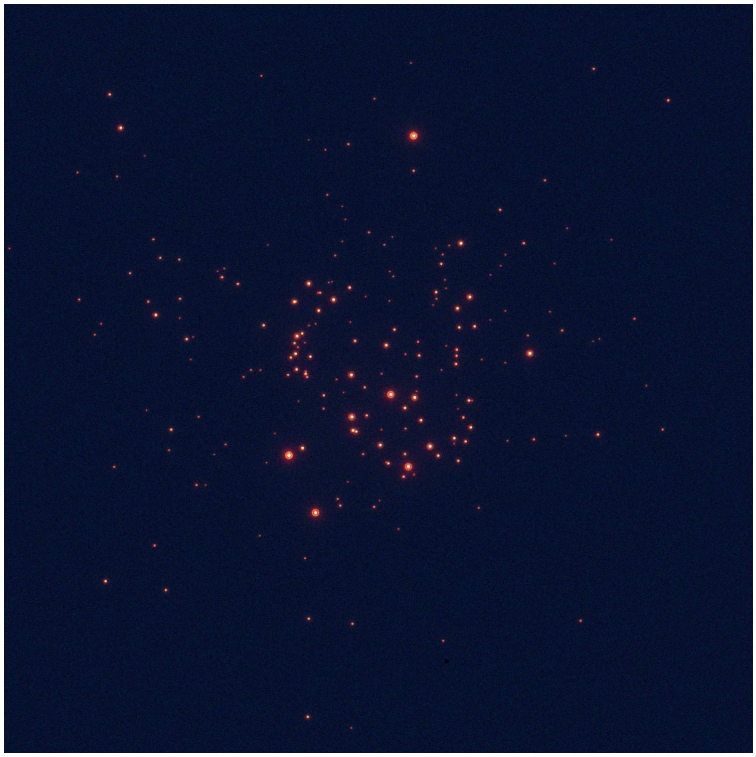

Figure 9.18: As Fig. 9.12 for a uniform reddening distribution of A_V = 150 mag over the whole cluster. Stars with masses $M > 20$ $M_\odot$ have a probability of 50% to be in the inner 0.1 pc of the cluster. The faintest star in the image is $K \sim 30$, $L \sim 22$ and $M \sim 16$ mag.

Fig. 9.8 showed an RGB image for the case of a Gaussian reddening distribution with a peak of A_V = 150 mag and a FW of 6″. Note how stars in the centre of the cluster are highly reddened while in the outskirts the extinction is approaching zero and stars are mostly blue.

Fig. 9.17 shows an RGB image for the case of a clumpy reddening distribution with an amplitude of A_V = 50 mag and a scale of 0.25″ over a Gaussian reddening distribution with a peak of A_V = 150 mag and a FW of 6″. Several stars in the outskirt of the cluster now fade to invisibility given the presence of extinction clumps (cf. Fig. 9.8).

Fig. 9.18 shows an RGB image for the case of a uniform reddening distribution with A_V = 150 mag but mass segregation is introduced in the simulations: stars with masses M $>$ 20 $M_\odot$ have a probability of 50% to be in the inner 0.1 pc of the cluster.

The $K, K-L$ CMD of a young massive cluster simulated with a uniform reddening distribution of A_V = 150 mag is shown in the left panel of Fig. 9.19. The number of recovered stars and the error bars are also shown in the figure. The magnitude of the faintest recovered star is $K \sim 29.7$ mag with $S/N \sim 10$. The magnitudes of the faintest recovered stars with the relative S/N and the magnitudes at which the scatter is 0.1 mag are listed in Table 9.6 for all simulation runs.

A similar CMD is shown in the right panel of Fig. 9.19 but for a cluster with a clumpy reddening distribution with clumps of amplitude A_V = 50 mag and scale 0.25″over a uniform extinction of A_V = 150 mag. This figure and Table 9.6 show that for the same total exposure time ($t_{\rm exp}$ = 25 h) we can reach the same limiting magnitude with the same S/N for the two different reddening distributions but we loose the most reddened stars, i.e. we recover only 34% of stars in the K-band, 87% in the L-band and 97% in the M-band. The L-band magnitude scatter for this simulation run is shown in Fig. 9.20. For a total exposure time of 1 h we are able to reach $L \sim 23$ mag, with $S/N \sim 3$.

We now summarise the results listed in Table 9.6 and discuss the influence of the different adopted input parameters.

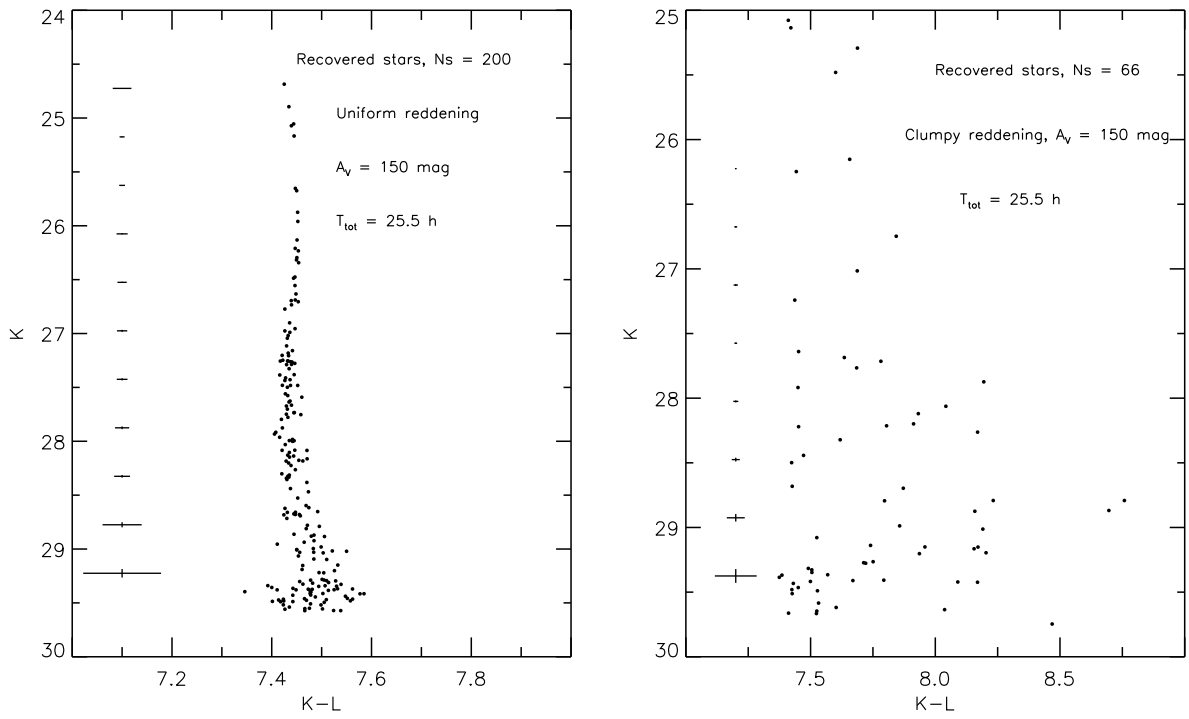

Figure 9.19: Left: $K, K-L$ CMD of a simulated young dense massive cluster for a uniform reddening distribution with A_V = 150 mag. Right: same for a clumpy reddening distribution with a clumps of amplitude A_V = 50 mag and scale 0.25″ over a uniform extinction of A_V = 150 mag.

Table 9.6: Results for the different imaging simulation runs.

Band	Reddening[a]	A_V mag	Limiting mag.[b] mag	S/N[c]	Scatter[d] mag	Recovered fraction %
K	Uni	100	24.2	170	...	100
K	Uni	150	29.7	10	29.7	100
$K^{e}_{\rm seg20}$	Uni	150	29.7	10	29.7	100
$K^{f}_{\rm seg30}$	Uni	150	29.7	10	29.7	100
K	Gauss	150[g]	27.8	4	27.4	89
K	Gauss/Clu	150+50[h]	27.4	20	...	66
K	Clu	150+50[i]	29.7	10	29.7	34
L	Uni	100	19.2	130	...	100
L	Uni	150	22.1	15	...	100
$L^{e}_{\rm seg20}$	Uni	150	22.1	15	...	100
$L^{f}_{\rm seg30}$	Uni	150	22.1	15	...	100
L	Uni	200	23.3	4	22.7	25
L	Gauss	150[g]	21.6	10	21.6	94
L	Gauss/Clu	150+50[h]	21.7	5	21.4	87
L	Clu	150+50[i]	22.7	5	22.3	64
M	Uni	100	15.5	40	...	100
M	Uni	150	16.4	40	...	100
$M^{e}_{\rm seg20}$	Uni	150	16.4	40	...	100
$M^{f}_{\rm seg30}$	Uni	150	16.4	40	...	100
M	Uni	200	17.4	35	...	100
M	Gauss	150[g]	16.2	40	...	94
M	Gauss/Clu	150+50[h]	17.6	23	...	97
M	Clu	150+50[i]	18.3	20	...	97

[a] Assumed reddening distribution.
[b] Magnitude of the faintest recovered star.
[c] S/N at the limiting magnitude reached.
[d] Magnitude at which $S = 0.1$ mag.
[e] Mass segregation is included, i.e. stars with $M > 20\ M_{\odot}$ have a probability of 50% to be in the inner 0.1 pc of the cluster.
[f] Mass segregation is included, i.e. stars with $M > 30\ M_{\odot}$ have a probability of 50% to be in the inner 0.1 pc of the cluster.
[g] Extinction of the peak of the Gaussian.
[h] Extinction of the peak of the Gaussian plus amplitude of the clumps.
[i] Uniform extinction plus amplitude of the clumps.

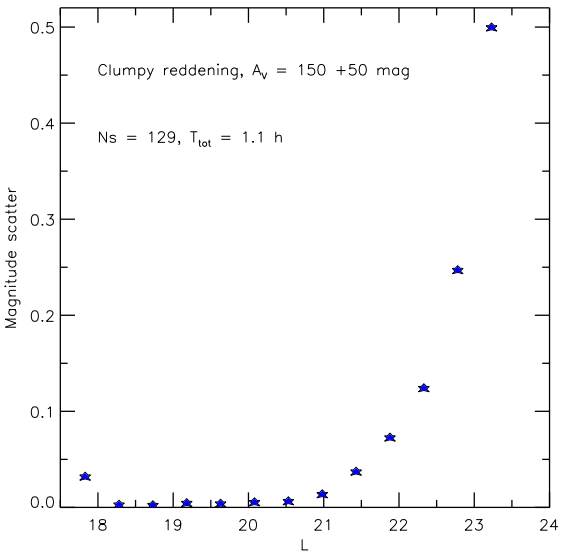

Figure 9.20: L-band magnitude scatter plotted against the L-band magnitude for a cluster with a clumpy reddening distribution. The number of recovered stars and the total exposure time are shown in the figure.

Reddening distribution:

1. Constant extinction for the whole cluster: when the reddening is less than $A_V < 150$ mag we are able to recover all the stars in the cluster with a $S/N > 10$. With a reddening of A_V = 200 mag, we detect 25% of stars in the L-band ($S/N > 5$) and 100% of stars in the M-band ($S/N > 35$).
2. Gaussian extinction profile: we recover more than 90% of stars in all IR bands. The magnitude scatter increases and we loose stars compared to the uniform reddening distribution case since we need to decrease the exposure time (see Table 9.1) given the presence of much brighter stars in the image (due to the extinction approaching 0 in the outskirts of the cluster).
3. Clumpy extinction on top of a uniform extinction of A_V = 150 mag: the magnitude of the faintest recovered star is $K \sim 29.7$ ($S/N \sim 10$), $L \sim 22.7$ ($S/N \sim 5$), and $M \sim 18.3$ mag ($S/N \sim 20$). We recover 34% of stars in the K-band, 64% in the L-band and 97% in the M-band.
4. Clumpy extinction on top of a Gaussian extinction profile: the magnitude scatter increases compared to the uniform reddening distribution case, given the decreased exposure time, and we recover 66% of stars in the K-band, 87% in the L-band, and 97% in the M-band.

Mass segregation: The presence of mass segregation at the level of 50% of stars with masses $>$ 20–30 $M_\odot$ being in the inner cluster core does not affect the results of the simulations.

9.5.3.2 Spectroscopy

For each simulation run we produced 1D spectra for all the stars included in the IFU and a 3D spectrum in the wavelength range 2.1–2.2 μm. Fig. 9.21 shows three 1D spectra for stars with masses 100, 30 and 13 $M_\odot$ for the case of a uniform reddening distribution with A_V = 100 mag, $t_{\rm exp}$ = 1200 s and R = 10 000. The qualitative analysis of the spectra indicates that we would be able to clearly detect the Brγ line at $\lambda \sim$ 2.17 μm for all the stars in the IFU ($8 < M < 100$ $M_\odot$). The S/N of the line for stars with four different masses and for the two reddening cases are listed in Table 9.5. However, this table and Fig. 9.22 indicate that in the case of a uniform reddening distribution with A_V = 150 mag, we would barely be able to detect the line in a star of 100 $M_\odot$ (top panel of Fig. 9.22) with an exposure time $t_{\rm exp}$ = 7200 s. The S/N of the Brγ line is already < 5 at M = 30 $M_\odot$, and we are essentially measuring noise. We also simulated 1D spectra for this case with longer exposure times but the S/N does not improve enough for reasonable values.

We did not simulate spectra for the other reddening distributions, such as the clumpy one, since most of stars would be even fainter in the K-band.

Fig. 9.23 shows a constant wavelength slice at λ = 2.164 μm, i.e. at the Brγ line, of a simulated 3D data cube for an IFU with FoV = 1.6″×1.6″ and spaxel size of 40 mas. The data cube contains a sample of 37 stars selected from a K-band simulated image for a uniform reddening distribution with A_V = 100 mag.

Fig. 9.24 shows the same for a uniform reddening distribution with A_V = 150 mag. In this case most of the stars have disappeared in the background noise.

9.5.4 Sensitivity to input parameters

The most critical input parameter in the imaging simulations is the reddening. If the extinction is larger than A_V = 150 mag we are not able to observe clusters at the distance of the GC in the K-band. On the other hand, the M-band would allow us to obtain almost complete observations of these GC clusters, but in the L-band we would loose ∼75% of stars. We would be able to observe GC cluster in the K, L, M bands if clumpy extinction is present, but we would loose a substantial fraction of stars depending on the clump scale (0.25″ in our simulations).

Mass segregation at the level of 50% of all stars with masses larger than 20 or 30 $M_\odot$ lying within the central 0.1 pc of the cluster does not affect the observation in the K, L, M bands.

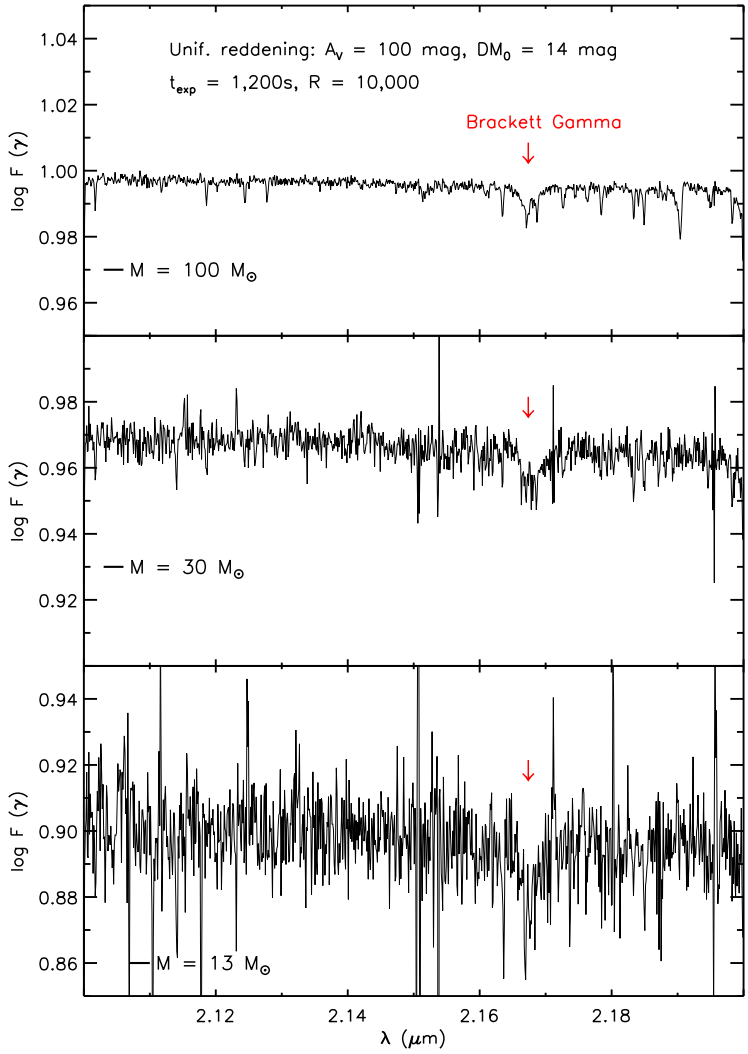

Figure 9.21: Sky subtracted simulated 1D spectrum for stars with different masses, as indicated in the figure, in the 2.1–2.2 μm wavelength range, for a uniform reddening distribution with A_V = 100 mag over the whole cluster.

The reddening is also the most critical input parameter for the spectroscopic simulations. As already shown in Section 9.5.3, for a uniform reddening larger than A_V = 100 mag, we are not able to accurately ($S/N > 10$) measure the Brγ line of stars in a cluster at the distance of the GC. The same result is also valid for the other reddening distributions.

9.5.5 Calibration requirements

No special calibration requirements are needed. The target fields are expected to contain sufficiently bright stars near the centre to allow an on-target accurate determination of the PSF, which will be taken as constant due to the small angular size of the target aggregates. These stars will also be included in wide-field surveys to be carried out in the coming years by survey telescopes like VISTA or in dedicated support observations by instruments such as HAWK-I at the VLT, thus providing built-in flux zero-points in the target fields.

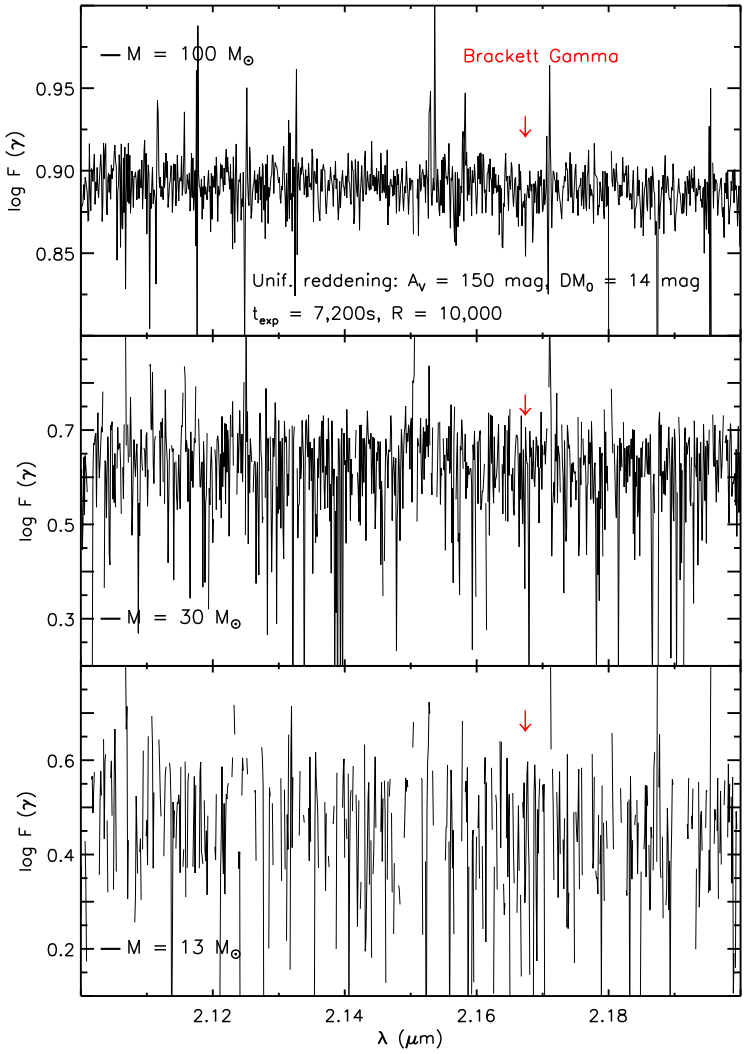

Figure 9.22: Same as Fig. 9.21 but for a uniform reddening distribution with A_V = 150 mag over the whole cluster.

9.5.6 Limitations

For the imaging case we did not perform simulations to estimate the accuracy of the proper motion measurements. Detailed and extensive simulations regarding the astrometric accuracy achievable with MICADO on the E-ELT have been performed by Trippe et al. (2010).

For the spectroscopic case we did not simulate spectra in the MIR wavelength range to check if we could detect the Brα line at 4.05 μm and use the ratio Brγ/Brα to estimate the individual reddening of cluster stars.

A possible improvement for the spectroscopic case would be to include the spectrum of the H II region in which the stars are embedded.

9.6 Concluding remarks

Imaging: The simulations confirmed that in order to penetrate an extinction as high as A_V = 200 mag we need to observe in the L or M-band. With the E-ELT and a total exposure time of $\sim$1 h + overheads,

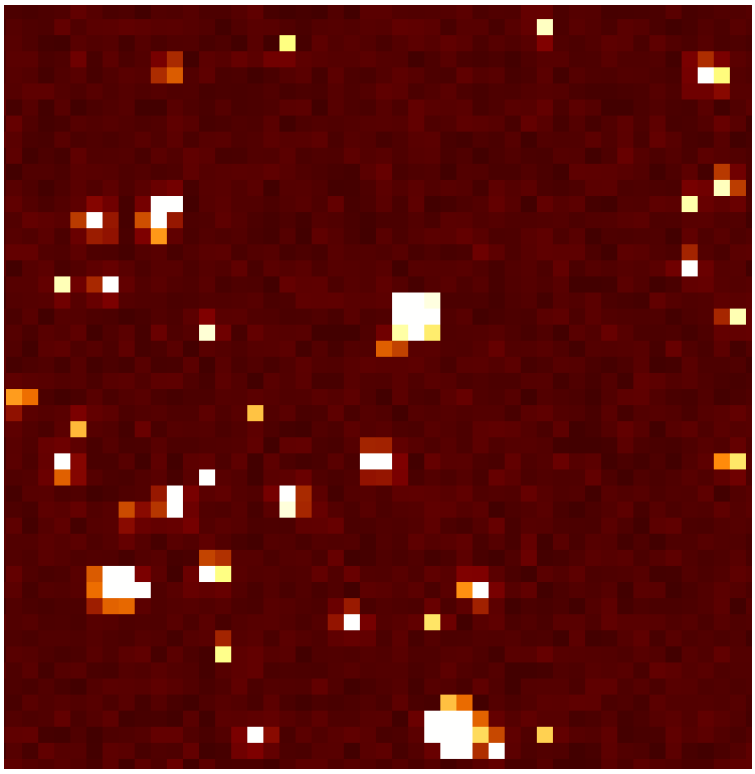

Figure 9.23: Slice of a 3D data cube at λ = 2.164 μm (Brγ line) simulated for an IFU with FoV = 1.6″×1.6″ and spaxel size of 40 mas. This data cube contains a sample of 37 cluster stars. Star fluxes were derived assuming a distance modulus for the GC of μ_0 = 14 mag and a uniform reddening distribution of A_V = 100 mag over the whole cluster.

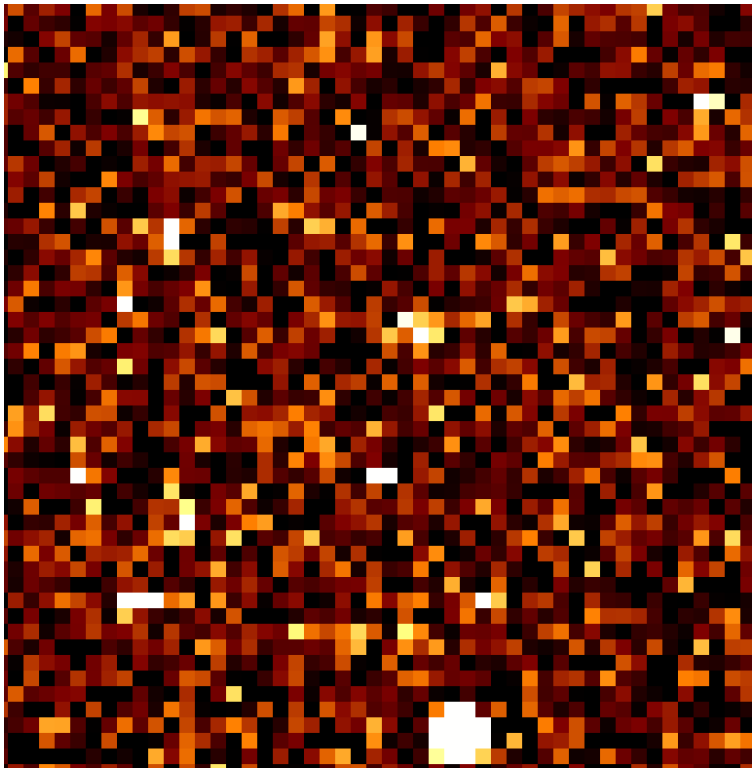

Figure 9.24: Same as Fig. 9.23 but for a uniform reddening distribution of A_V = 150 mag.

we will be able to observe deeply embedded (A_V = 200 mag) massive dense clusters at the distance of the GC. The completeness of the $L, L-M$ CMD is 25% but we would be able to determine the stellar number density of the cluster adopting M-band star counts. On the other hand, we will be able to observe massive dense clusters in the GC in the case of a uniform extinction of A_V = 150 mag, in the K, L, M bands, with a total exposure time of $\sim$25 h + overheads. The completeness of the three colour CMDs is 100% and we would be able both to study the stellar number density of the cluster and to probe the presence of hot and warm circumstellar matter (disks/envelopes) by means of the detection of IR excess ($K-L$, $L-M$). If a clumpy reddening is present in the cluster, the completeness will decrease according to the scale length of the extinction clumps. For the assumed scale of 0.25″ at the GC distance, we will be able to recover 34% of stars in the K-band, 64% in the L-band and 97% in the M-band with a $S/N >$ 5. In the case of a Gaussian reddening distribution, the recovered fractions will be $>$ 90% for the three bands. We will also be able to detect the presence of mass segregation in a dense massive cluster in the GC at least if the extinction is less than A_V = 200 mag.

Spectroscopy: The simulations showed that 3D spectroscopy of a dense massive cluster in the GC in the K-band region is feasible with the E-ELT, an IFU of 1.6″$\times$1.6″, a spaxel size of 40 mas, if the extinction does not exceed A_V = 100 mag. For instance, with a reasonable exposure time of 1200 s, the S/N at the Brγ line at 2.17 μm of stars with masses $>$ 8 $M_\odot$, i.e. $K \sim$ 24 mag, would be higher than 10. In addition, with a uniform extinction of A_V = 150 mag over the whole cluster, we will still be able to measure the Brγ line with a $S/N \sim$ 10 for the most massive stars, i.e. M $\sim$ 100 $M_\odot$ and $K \sim$ 25 mag. Note that in these simulations we did not take into account the spectrum of the H II region in which these stars are embedded.

Acknowledgements

We kindly acknowledge all members of the E-ELT Science Office, in particular G. Battaglia and J. Liske, for their help and support. We also acknowledge F. Comeron and I. Hook of the Science Working Group. We further thank A. Micol for his help with RGB image creation.

References

Alves J., Lombardi M., Lada C.J., 2007, A&A, 462, L17
Bally J., Zinnecker H., 2005, AJ, 129, 2281
Bonnell I.A., Bate M.R., Zinnecker H., 1998, MNRAS, 298, 93
Bonnell I.A., Bate M.R., Vine S.G., 2003, MNRAS, 343, 413
de Wit W.J., Testi L., Palla F., Zinnecker H., 2005, A&A, 437, 247
Dobbs C.L., Bonnell I.A., 2005, in "Protostars and Planets V", 8170
Grazian A., Fontana A., De Santis C., Gallozzi S., Giallongo E., Di Pangrazio F., 2004, PASP, 116, 750
Krumholz M.R., McKee C.F., 2005, ApJ, 630, 250
Lejeune T., Schaerer D., 2001, A&A, 366, 538
Meynet G., Maeder A., Schaller G., Schaerer D., Charbonnel C., 1994, A&AS, 103, 97
Moeckel N., Bally J., 2007, ApJ, 656, 275
Rieke G.H., Lebofsky M.J., 1985, ApJ, 288, 618
Schaller G., Schaerer D., Meynet G., Maeder A., 1992, A&AS, 96, 269
Stetson P.B., 1987, PASP, 99, 191
Stetson P.B., 1991, AJ, 102, 589
Stetson P.B., 1994, PASP, 106, 250
Stetson P.B., Bruntt H., Grundahl F., 2003, PASP, 115, 413
Trippe S., Davies R., Eisenhauer F., Förster Schreiber N.M., et al., 2010, MNRAS, 402, 1126
Zinnecker H., Yorke H.W., 2007, ARA&A, 45, 481

10 S5-3: Giant planet-mass objects in the Large Magellanic Cloud

Authors: A. Calamida, F. Comerón, H. Zinnecker

10.1 The science case

The goal of this science case is to probe the complete sub-stellar mass regime of a young star forming region in the Large Magellanic Cloud (LMC) down to giant planet masses ($M \leq 10\ M_{\rm Jup}$). This mass may be below the opacity limit setting the minimum mass of objects formed by fragmentation at the metallicity of the LMC. Therefore, the observations proposed here have the potential of revealing the opacity limit in a low-metallicity environment such as that of the early Milky Way, thus providing a data point that cannot be obtained from observations in our own Galaxy. The determination of the lowest-mass Initial Mass Function (IMF), and eventually of the location of the opacity limit at low metallicity, will be helpful to constrain the volume density of evolved giant planet-mass objects lurking in our own galactic disk, which have faded into invisibility since a long time ago.

Theoretical estimates for the opacity limit (Rees 1976; Bate et al. 2005; and references therein) lie around 10 $M_{\rm Jup}$, but simulations including realistic environment conditions of cores have shown that the opacity limit can be overcome to form objects of even smaller masses (Boyd et al. 2005). Indeed, observations of very young aggregates have revealed members with masses probably in the range of a few Jupiter masses only (e.g. Zapatero Osorio et al. 2002). As wide-field, deep infrared surveys of star forming regions are carried out in the next years, it may be expected that the existence and location of the lower limit to the sub-stellar mass function will be firmly established.

Nevertheless, even if the lower limit to the IMF can be observationally derived soon, it will be so only for the young clusters of the solar neighbourhood, all of which have virtually solar metallicity. Since the limit depends on the cooling curve of dense, cold molecular gas and the cloud opacity, metallicity is expected to play a major role in determining it. This has rather far-reaching consequences concerning the evolution of the mass function of galactic disks, since during most of its history the disk our Galaxy has been forming stars at sub-solar metallicity. Therefore, knowing the shape of the lower end of the IMF at solar metallicity, and knowing the location of its lower mass cutoff, is of limited value for the purpose of estimating the number of extremely low mass objects that may exist nowadays in the galactic disk, since an extrapolation of their present-day formation rates would not be possible towards the low-metallicity ages of our Galaxy.

Can this problem be addressed observationally? It seems unlikely at present. Available models at solar metallicity (Baraffe et al. 2003) indicate that a 5 $M_{\rm Jup}$ object (similar to the lowest-mass ones detected in the σ Orionis cluster; Caballero et al. 2007) coeval with the Sun would have cooled down to 220 K of surface temperature, having $M_J = 29.8$, $M_H = 28.8$, $M_K = 41$ mag. These faint absolute magnitudes mean that next-generation surveys reaching down to $J \sim 23$ mag could not detect them beyond 0.4 pc or about 100 000 AU. Objects of the same mass formed at the beginning of the life of the Galaxy would have faded to $M_J = 34$ mag and would not be detected beyond 13 000 AU. Somewhat younger objects (ages less than $\sim$3 Gyr) of the same mass could be detectable if present at a distance of a few parsecs, but their metallicities would be practically solar and thus not relevant to the determination of the evolution of the low-mass cutoff with metallicity.

Unless we are extremely lucky and one of those objects happens to be in the close vicinity of the Solar System, there is no possibility of detecting a real low-metallicity isolated planetary-mass objects (PMOs) in our Galaxy. Yet these objects might be extremely abundant if the early, low-metallicity Galaxy was at all able to form them.

Interestingly enough, the E-ELT will make possible the observation of low metallicity brown dwarfs and may answer the question of whether or not they formed, and in which amounts, in the early Galaxy.

In principle, old Galactic PMOs may be detectable with the E-ELT by carrying out deep surveys reaching up to a few parsecs from the Sun, but the chances to serendipitously detect one such object would be vanishingly small. However, the Magellanic Clouds are places where low-metallicity brown dwarfs

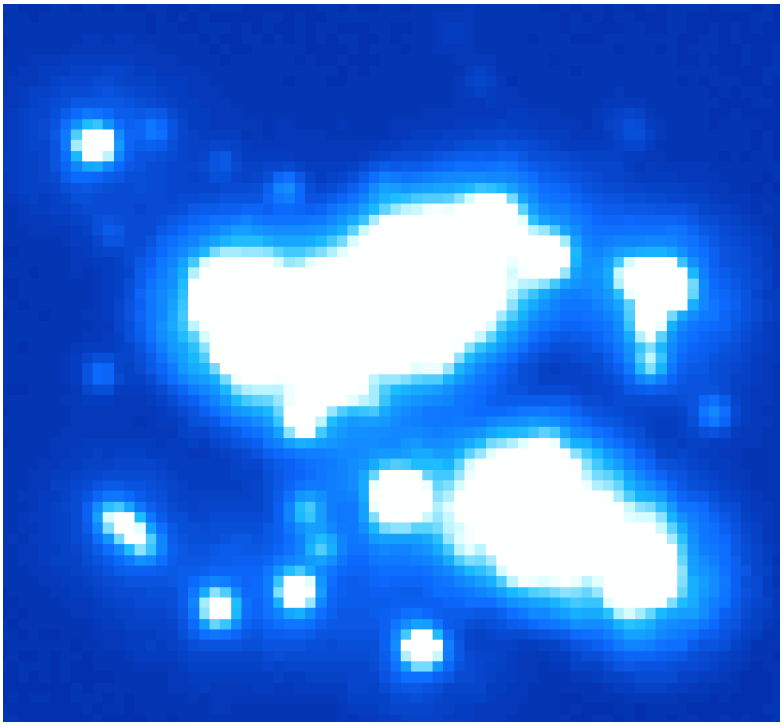

Figure 10.1: K-band image ($t_{\rm exp}$ = 1 h) of a simulated young star forming region (100 objects) observed with MAD/VLT.

are being formed *today* and are thus many orders of magnitude brighter than their old counterparts of the same metallicity in the disk of our Galaxy. In a 1 Myr old star forming region of the LMC, our 5 $M_{\rm Jup}$ planetary-mass object would still have a temperature of 1900 K and magnitudes $M_J \simeq 10.6$, $M_H \simeq 10.2$, $M_K \simeq 9.7$ mag, with a spectral type probably around mid-L. At the distance modulus of the LMC ($\mu_0 = 18.5$, Freedman et al. 2001), this implies $J \simeq 29.1$, $H \simeq 28.7$, $K \simeq 28.2$ mag. The numbers are approximate only for several reasons. First, large uncertainties plague evolutionary models as such early ages (Baraffe et al. 2003), especially in the giant planet-mass regime (Marley et al. 2007). Secondly, these models are calculated for solar metallicity, whereas the atmosphere characteristics, interior opacities and evolutionary tracks should be expected to vary sensibly for significantly sub-solar metallicity.

10.2 Goals of the DRM simulations

The goal of this proposal is to constrain the low-mass luminosity function down to the giant planet-mass regime in some star forming regions of the LMC, with the hope of determining the as yet uncertain location of the opacity limit at the metallicity (40% solar; Hunter et al. 2007) of the LMC. Typical low-mass star forming regions of the solar neighbourhood like Lupus 3 or ρ Ophiuchi would subtend an angle $\sim$2″ at the distance of the LMC, thus being appropriate for Laser Tomography Adaptive Optics (LTAO) observations which would provide diffraction-limited, deep imaging of the whole stellar and sub-stellar content of the cluster.

The regions to be observed would normally contain stars of much higher masses, up to a few solar masses for a low-mass aggregate containing Herbig Ae/Be stars at the top end of its mass function, as is the case for well-known nearby star forming regions like Chamaeleon I, Lupus 3 or ρ Ophiuchi. The object density is unknown, as the initial mass function is unknown down to the masses of interest in this proposal, but a rough guess based on the extrapolation of the IMF of galactic star forming region (Chabrier 2003) yields about 20 members of the aggregate per arcsec^2.

Fig. 10.1 shows a simulated low-mass star forming region at the distance of the LMC as it would be seen with the Very Large Telescope (VLT) and the Multi-Conjugate Adaptive Optics Demonstrator (MAD), with

a pixel scale of 28 mas. This simulation shows that, with this combination of telescope and instrument, we would be able to detect only the brightest stars of the region. The spatial resolution, indeed, and the photon collecting power of the VLT do not allow us to identify the faint giant planet-mass objects.

The challenge of these observations is the high density of objects and the fact that the members of the star forming region will cover a wide range in magnitudes, from $K \sim 18$ to $K \lesssim 30$ mag. The simulations should determine whether meaningful photometry of the faintest members can be achieved given the anticipated degree of crowding, or whether the faintest members may be rendered completely undetectable because of the combined wings of the PSFs of (much) brighter members of the star forming region. The simulations can be made more realistic by assuming the presence of a variable background due to reflection nebulosity, especially in the J-band, and variable extinction along the line of sight to each aggregate member. The photometric accuracy should be at the 0.2 mag level in order to confirm the low masses of the faintest detected objects on the basis of their colours. Another aspect that may be considered in the simulations is the fact that the LMC, at $\delta = -69°$, is observable at most only at high zenith distance ($ZD > 45°$) from any of the potential E-ELT sites.

10.3 Metrics / figures of merit

The input for the simulations will be a catalogue of stellar positions and magnitudes. The simulations will need to perform source detection and photometry, producing an output catalogue of recovered stellar positions and magnitudes.

The main tool to be used in deciding whether or not the proposal achieves its goals will be the comparison between the input and output luminosity function at each of the J, H and K bands and the magnitude scatter ($\sqrt{\frac{1}{N}\sum(\mathrm{input} - \mathrm{recovered})^2}$). The simulations should produce as outputs (i) the magnitude in each filter at which the number of recovered objects is 90% and 50% of the input objects; (ii) the magnitude at which the magnitude scatter of recovered objects in each filter reaches 0.1 and 0.2 mag.

10.4 DRM simulations

10.4.1 Methodology

In order to take into account the points of Section 10.2 and 10.3 we split the simulations into two main cases: in the first one we consider a uniform background for the star forming region, while in the second one we assume the presence of variable background emission. Moreover, we perform the simulations for different zenith distance (ZD) values, in order to asses how much the location of the telescope will influence our results. This is particularly important because of the low declination of our targets, the star forming regions in the LMC ($\delta = -69°$). We will also analyse the influence of the adopted pixel scale and the effect of the stellar density on the simulation results.

After producing images in the J, H, K near-infrared (NIR) bands for a simulated young star cluster, we will perform crowded field profile fitting photometry on the data. We will then compare the obtained luminosity functions with the input one, to check if giant planet-mass objects can be detected and measured with the sufficient accuracy (see Section 10.2).

10.4.2 Pipeline

We simulate images of a young star forming region assuming two background conditions and for different input parameters. First, we need to generate a catalogue with positions and magnitudes in the J, H, K bands. For this purpose we adopt a distribution of stellar masses drawn from the Chabrier (2005) IMF. Since we consider only low-mass star forming regions analogous to the nearby ones, the mass function is truncated at the high mass end at 2 $M_\odot$. The normalisation factor is chosen such that 100 objects are simulated and scattered at random in a circle of 1″ radius. Masses are transformed to J, H, K absolute magnitudes by adopting the evolutionary tracks of Baraffe et al. (2003), with solar chemical composition

Table 10.1: Some input parameters adopted to simulate NIR images of a young star forming region.

Band	$\lambda_{\rm eff}$ μm	Half-width μm	Pixel scale mas/pix	Zenith distance $^\circ$	Background emission mag arcsec^{-2}
J	1.22	0.213	2, 5	0	23.9
H	1.63	0.307	2.6, 5	0	24.8
K	2.19	0.39	3.5, 5	0, 30, 60	25.4

and for an age of 5 Myr. A distance modulus for the LMC of $\mu_0 = 18.5$ is added to the absolute magnitudes, plus a random extinction in K, A_K, for each star drawn from a uniform probability distribution ranging from 0 to 1. The corresponding extinctions in J and H are, respectively, $A_J = 2.52\,A_K$, $A_H = 1.61\,A_K$, adopting the Cardelli et al. (1989) reddening law.

In order to increase the statistics and to estimate with more accuracy the completeness function we also produce a catalogue with 10 000 stars distributed in a circle of 10″ radius, so that the average stellar density is preserved. We generate this catalogue both for a uniform stellar density distribution and for a density profile falling as $1/r$ from the centre of the star forming region, in order to test the effect of this parameter on the limiting magnitude and on the recovered fraction of stars.

We neglect field star and galaxy contamination given the small field of view (FoV) we are interested in (1″ radius) and the high stellar density in star-forming aggregates and young clusters.

After selecting the filter we rescale the star positions according to the assumed pixel scale. We simulate images with two pixel scales for each filter, 5 mas/pixel and the scale of the corresponding PSF for that band (see Table 10.1). The choice of the latter pixel scale is then related to the spatial sampling of the simulated PSF for the adopted adaptive optics mode (LTAO). This mode is used due to the need for excellent on-axis AO performance and the very small angular extent of star forming aggregates at the LMC distance. We perform the simulations assuming different values for the zenith distance, namely $ZD = 0, 30, 60^\circ$, for the K-band. LTAO PSFs at zenith distances other than 0° are only available for the I_c and the K-band (see [RD1]).

In the case of the variable background we assume that the emission is unresolved. We thus place a source in each pixel of the image, whose flux is randomly derived from a uniform probability distribution ranging between zero background and a brightness that is twice the value corresponding to the uniform background case (see values listed in Table 10.1).

In order to generate the images we use standard IRAF packages and the Large Binocular Camera Image Simulator (LBCImSim, Grazian et al. 2004). This simulator (see the LBC web page) has been developed to test the performance of the LBC, but it can be used to simulate optical and NIR images for any other telescope, such as the E-ELT. We proceed in two steps:

1. We first generate an artificial image, in the standard FITS format, using the IRAF MKOBJECTS task, with one second exposure time and with no noise. The image will contain the sources (100/10 000 objects plus the random brightness peaks in the case of the variable background) convolved with the LTAO PSF.
2. We then feed this image to the LBCImSim simulator that produces an output image adding the noise and other observational artifacts, such as system efficiencies, flat-field inaccuracies, sky background. It then simulates the observational strategy, i.e. the detector integration time (DIT) and the number of exposures.

Images of the sky background are also generated following the same steps and then subtracted from each single scientific image. The resulting FITS format images are then ready to be analysed.

The data analysis is carried out using DAOPHOT IV/ALLSTAR (Stetson 1987, 1991) and ALLFRAME (Stetson 1994), and personal IDL routines. DAOPHOT is one of the most popular software package for stellar photometry. It was developed by Peter B. Stetson at the Dominion Astronomical Observatory, more than twenty years ago, and has undergone a continuous evolution and sophistication. DAOPHOT is used

to identify the brightness peaks on the image and to perform synthetic aperture photometry. We select a few bright, isolated and non-saturated stars to reconstruct an empirical PSF and use ALLSTAR to perform profile fitting photometry on the single images. All the images are at last analysed together with the task ALLFRAME (for more details see Section 10.5.2). We end up with a catalogue with positions, apparent magnitudes and relative errors, for each image. In order to create a unique star catalogue and estimate the weighted mean magnitudes we use DAOMATCH/DAOMASTER (Stetson 1994).

We are now able to estimate the recovered fraction of stars, to compare the obtained luminosity function with the input one and to estimate the magnitude scatter (see Section 10.5.2).

The outline of the simulation pipeline is given in Fig. 10.2.

10.4.3 Inputs

10.4.3.1 Scientific data

A specific recipe of the scientific input for the simulations is as follows:

- A distribution of stellar masses drawn from the Chabrier (2003) initial mass function, whose most up-to-date analytical form (log-normal) can be found at Chabrier (2005). Since we will consider only low-mass star forming regions analogous to the nearby ones, the mass function should be truncated in the high mass end to 2 $M_{\odot}$. A realistic region would have $\sim$100 objects. However, for statistical sampling purposes some simulations have been carried out for 10 000 objects in a larger aggregate, preserving the surface density.
- The mass is transformed to J, H, K absolute magnitudes using the evolutionary tracks of Baraffe et al. (2003) for an age of 5 Myr (see Fig. 10.3).
- A distance modulus of 18.5 (Freedman et al. 2001) is added to the absolute magnitudes, plus a random extinction in K, A_K, for each star drawn from a uniform probability distribution ranging from 0 to 1. The corresponding extinction in J and H are respectively taken as $A_J = 2.52\,A_K$, $A_H = 1.61\,A_K$.
- The simulated stars are scattered at random in a circle of 1″ radius, for the 100 star aggregate and in a circle of 10″ radius for the 10 000 star aggregate. The distribution of stars over the area covered by the aggregate is assumed to be homogeneous.
- Stellar light reflected by dust is introduced as an increase in the background of the region. Assuming reflection nebulosity with a total reprocessed emission in the K-band equivalent to $1/10$ of the solar luminosity in that band, and typical emission nebulosity blue colours $(J-H) = -0.9$, $(H-K) = -0.6$, we obtain a background level of $\sim$23.9, $\sim$24.8, and $\sim$25.4 mag arcsec^{-2} in the J, H and K bands, respectively. It is proposed to simulate this contribution as (i) a uniform background level over the 1″ radius of the star forming region, and (ii) a variable background in which the emission per pixel takes a random value drawn from a uniform probability distribution going from zero background to a brightness that is twice the one corresponding to the values given above.

Fig. 10.3 shows an isochrone from the set of Baraffe et al., with an age of t = 5 Myr and a solar chemical composition, plotted in the $H, H-K$ colour-magnitude diagram (CMD) and shifted to the LMC distance.

Fig. 10.4 shows the synthetic $H, H-K$ CMD drawn from the Chabrier IMF for 100 objects with $M \leq 2\,M_{\odot}$, that we adopt to simulate our star forming region. Masses are transformed to J, H, K absolute magnitudes by adopting the theoretical isochrone (t = 5 Myr) of Baraffe et al. plotted in Fig. 10.3. The LMC distance modulus and a random extinction A_K ranging from 0 to 1 are also added.

10.4.3.2 Technical data

The technical data used for the simulations were mainly taken from [RD1].

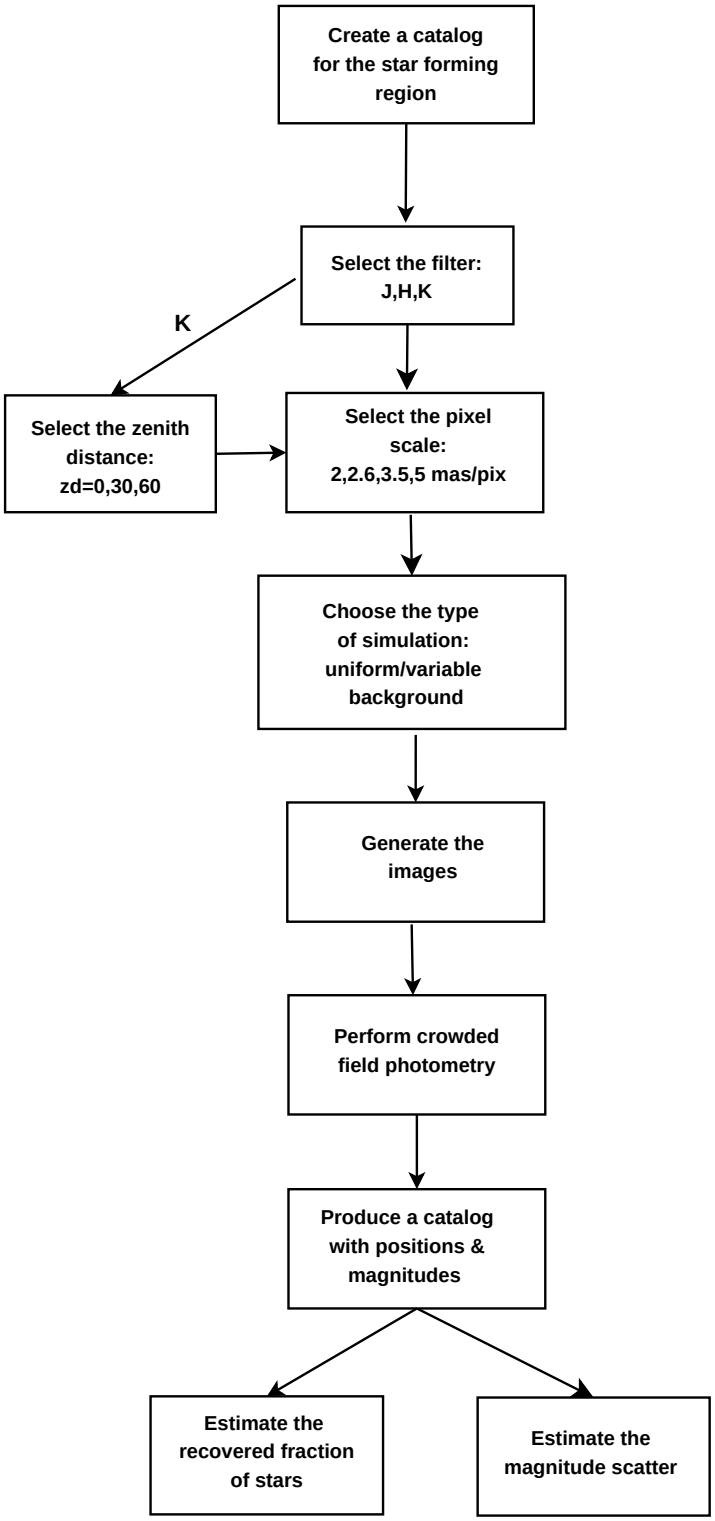

Figure 10.2: Flowchart of one simulation run.

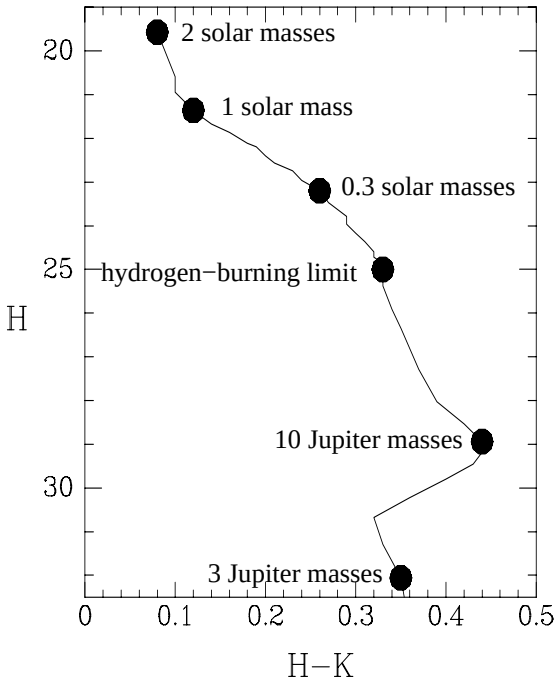

Figure 10.3: An isochrone from the set of Baraffe et al. (2003), with an age of t = 5 Myr and a solar chemical composition, is plotted on the $H, H-K$ colour-magnitude diagram (CMD). Theory has been transformed to the observational plane by adopting the colour-temperature relations by Baraffe et al. and a distance modulus for the LMC of $\mu_0 = 18.5$ mag. The positions of objects of selected masses are shown.

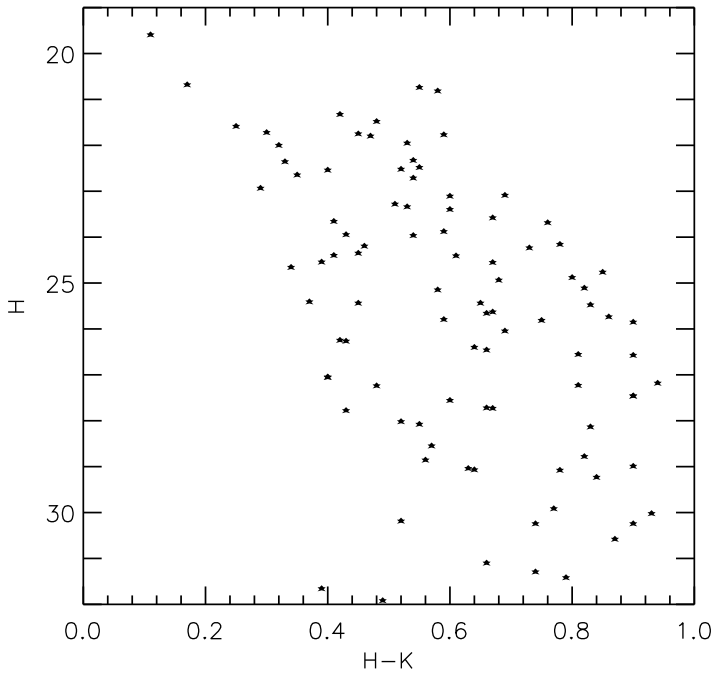

Figure 10.4: Synthetic $H, H-K$ CMD drawn from the Chabrier (2005) IMF for 100 objects with $M \leq 2\,M_{\odot}$. Masses are transformed to J, H, K absolute magnitudes by adopting a Baraffe et al. isochrone with an age of t = 5 Myr and solar metallicity (see Fig. 10.3). A distance modulus for the LMC of $\mu_0 = 18.5$ mag is added, plus a random extinction in the range A_K = 0–1 mag (see text for more details).

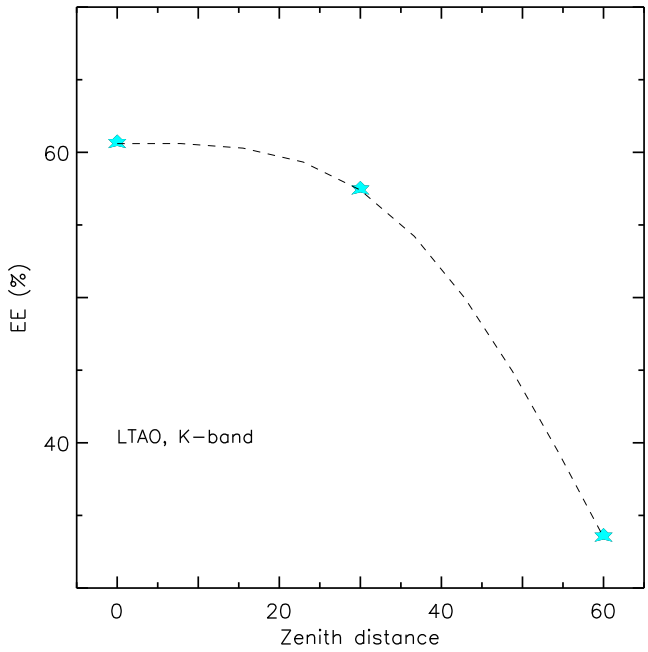

Figure 10.5: Encircled energy (EE) of the LTAO K-band PSF at a radius of 70 mas plotted as a function of zenith distance.

Site and background: We use the background model for a Paranal-like site as described in [RD1]. Since this is an imaging case we are only concerned with broad-band background surface brightness values. We adopt the values of J = 16, H = 14 and K = 13 mag arcsec^{-2} from [RD1].

Telescope: As described in [RD1]. We assume the protected silver/aluminium coating.

Instrument: We assume a diffraction limited imager, operating in the wavelength range 1–2.5 μm. We do not have any specific requirement on the instrument's FoV since our observed targets, at the distance of the LMC, subtend an angle of only $\sim$2″. We simulate different pixel scales for each band, ranging from 2 to 5 mas/pixel.

An instrumental efficiency was not available at the time this work was carried out. We choose to use a constant value of 0.63 (including the detector QE) so that the total telescope and instrument efficiency is $\sim$0.6 (assuming the Ag/Al coating).

The read-out noise (RON) of the detector was assumed to be 3 e$^-$/pix as in [RD1].

Adaptive Optics and PSFs: Since this science case requires high spatial resolution we have assumed LTAO for the simulations. In each band we only use the on-axis PSF, assuming that it does not vary significantly across the apparent target diameter of $\sim$2″. Note that the same assumption is made for the case of 10 000 stars distributed across a $\sim$10″ FoV, since we are not simulating a real cluster, but we are only increasing the number of stars to improve the statistics.

We assume a zenith distance of $ZD = 0°$, except in the K-band we will also consider $ZD = 30°$ and $60°$. Fig. 10.5 shows how the encircled energy of the LTAO K-band PSF decreases, for a fixed radius, when moving from $ZD = 0°$ to $ZD = 60°$.

To improve the speed of the simulations we have not used the original PSF images but rather their analytical fits described in [RD1].

10.4.4 Outputs

The simulation data products are images in the standard FITS format and tables with the stellar positions and recovered magnitudes in the different NIR bands.

For each simulation run we generate a set of images, according to the signal-to-noise (S/N) ratio that we want to achieve. These images are dithered of a random number of pixels ($\leq$ 40 pix) in the X and/or Y direction. An image of the sky background is subtracted from each simulated scientific image. We then stack all the images of a set in a median in order to increase the S/N of the faint objects which we are interested in, making their detection more straightforward. Fig. 10.6 shows the median of 25 K-band images ($t_{\rm exp}$ = 1 h) of the simulated young star forming region for the case of uniform background and $ZD = 0°$. Fig. 10.7 shows the median of the subtracted images after one or two ALLFRAME run (see Section 10.5.2 for more details). The image sky background, as estimated by DAOPHOT, has not been added in the calculation of the median. Fig. 10.8 shows a zoom of an H-band median image (left panel) and the relative subtracted image (right): after the subtraction of the brightest stars, the detection of the faint neighbours ($H \geq 29$ mag, $M \lesssim 10\ M_{\rm Jup}$) is more straightforward.

For each simulation run we then produce a catalogue with star positions and magnitudes to be compared with the input one. Some simulation runs for the 100 star cases are repeated five times, changing the dithering and the sky background images, and the output catalogues are then added together.

Fig. 10.9 shows the comparison between the K-band images of a 10 000 stars star forming region, simulated according to a stellar density distribution scaling as $1/r$ (left panel) and to a uniform distribution (right). Note that the crowding increases in the central 1 arcsec region because of the increased stellar density but it decreases in the external regions of the image. Furthermore, Fig. 10.10 shows the comparison between a section of the K-band images simulated for two different zenith distance values, $ZD = 60°$ (left panel) and $ZD = 30°$ (right). It is noteworthy how, for $ZD = 60°$, the faintest objects start to disappear in the increased noisy background due to the wings of the PSFs of the brightest star.

Some example data are available at http://www.eso.org/sci/facilities/eelt/science/drm/S5/.

10.5 Results of simulations

10.5.1 Simulation runs

The different simulation runs and their input parameters are summarised in Table 10.2.

We use the E-ELT Imaging Exposure Time Calculator [RD1] to estimate the observing time required to reach J=29.1, H=28.7, K=28.2 mag with a $S/N = 10$. These are the magnitudes that a young ($\sim$1 Myr) 5 $M_{\rm Jup}$ object would have at the distance of the LMC. We then need to reach these magnitudes with an accuracy of $\sim$0.1 mag in order to accomplish the goals of the proposal (see Section 10.2). The total exposure times adopted for each simulation run are listed in Table 10.2. We perform some tests and we find that the best strategy is to divide the observing time in 60 DIT of 60 second each (DIT=60s, NDIT=60), and in 10 exposures for the J-band and 25 for the H and the K-band. This combination allows us not to saturate the brightest stars in the field (short DIT) and to reach a good S/N on every image. The total observing time for a young star forming region will then be: $T_{\rm tot} = 10(J) + 25(H) + 25(K) \approx 60h + {\rm sky} + {\rm overheads}$.

10.5.2 Analysis

The data reduction is performed with DAOPHOT IV/ALLSTAR/ALLFRAME and personal IDL routines following these steps:

- We first identify the brightness peaks on each image: a threshold must be defined in order to choose a level of significance of the brightness enhancement. This threshold depends on the RON

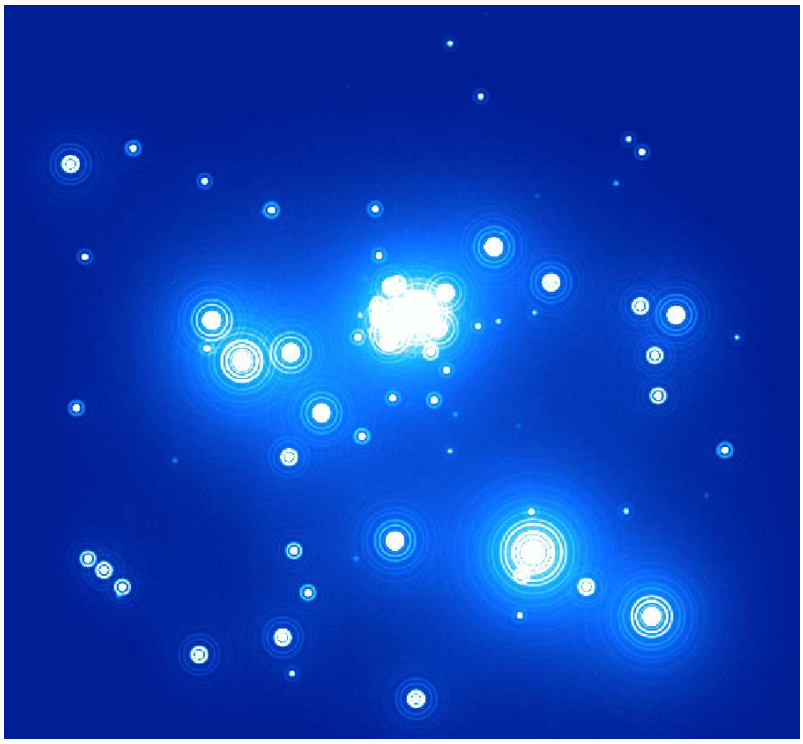

Figure 10.6: K-band median image of the simulated young star forming region for a total exposure time of 25 h, a pixel scale of 3.5 mas/pixel, $ZD = 0°$ and uniform background.

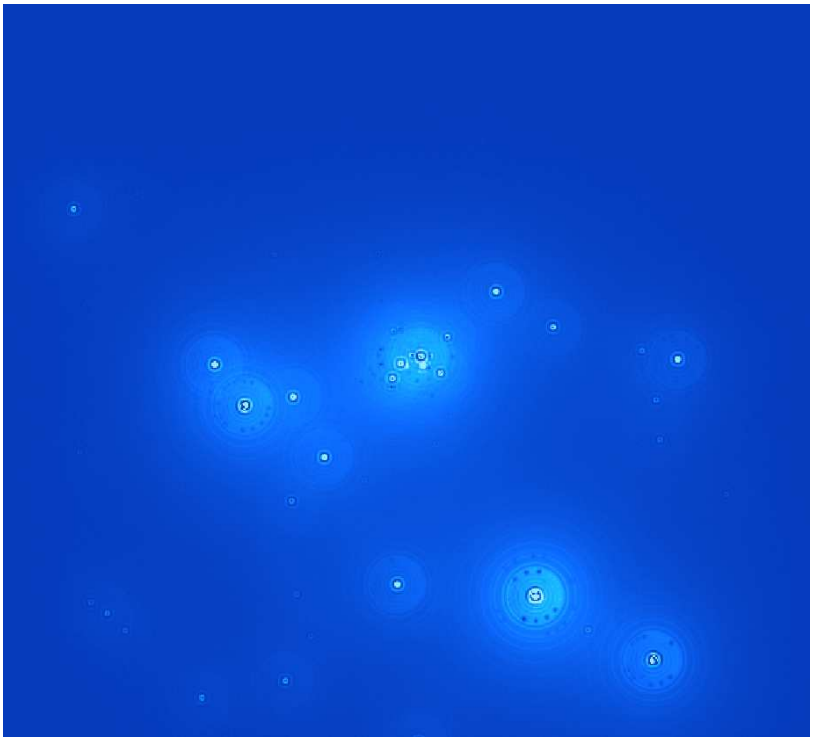

Figure 10.7: Median of the subtracted K-band images, after the ALLFRAME run, of the simulated young star forming region for a total exposure time of 25 h, a pixel scale of 3.5 mas/pixel, $ZD = 0°$ and a uniform background.

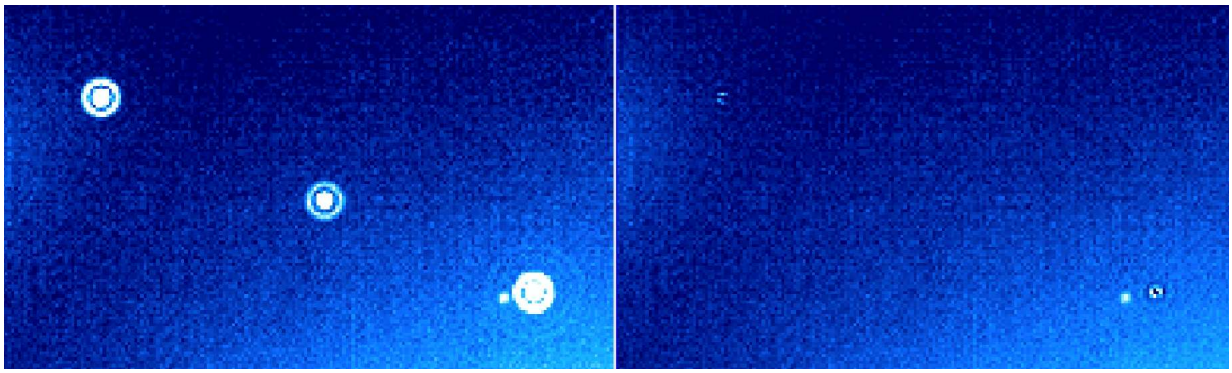

Figure 10.8: Zoom of an H-band median image (left panel) and the corresponding subtracted image (right), of the simulated young star forming region for a total exposure time of 25 h, a pixel scale of 2.6 mas/pixel and a uniform background.

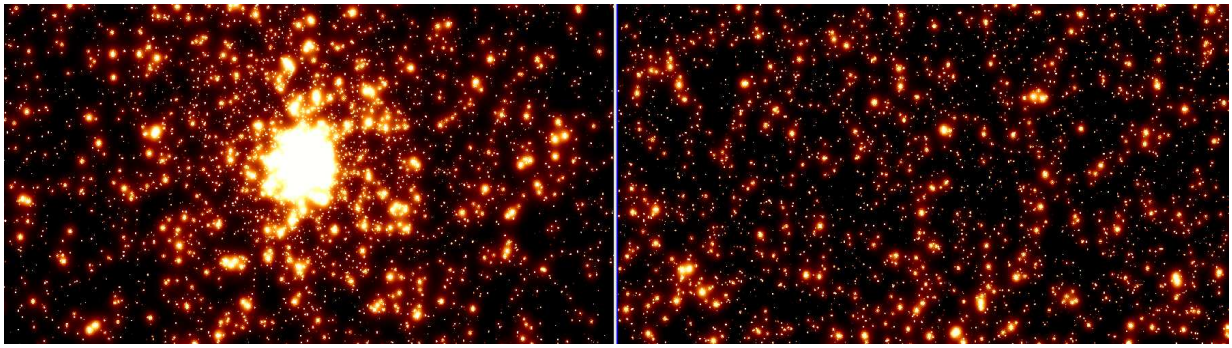

Figure 10.9: Left: K-band median image of the simulated 10 000 star forming region for a total exposure time of 25 h, a pixel scale of 3.5 mas/pixel, $ZD = 0°$, a uniform background and stellar density distribution scaling as $1/r$. Right: Same as before but for a uniform stellar density distribution.

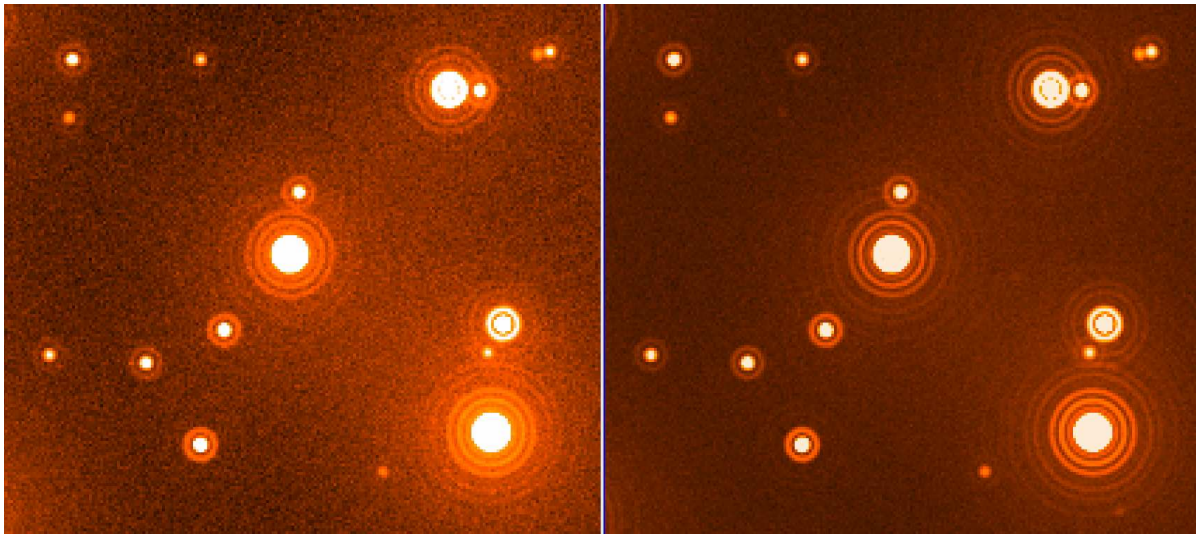

Figure 10.10: Left: Section of the K-band image of the simulated $10\,000$ star forming region for an exposure time of 1 h, a pixel scale of 3.5 mas/pixel, $ZD = 30°$. Right: Same as before but for $ZD = 60°$.

Table 10.2: Input parameters for the different simulation runs.

Band	NS[a]	Pixel scale mas/pix	ZD ∘	Sky[b] mag arcsec^{-2}	Back.[c]	DIT sec	NDIT	$T_{\rm tot}$ hours	N[d]
J	100	2	0	16	Uniform	60	60	10	5
J	100	5	0	16	Uniform	60	60	10	1
J	100	2	0	16	Variable	60	60	10	1
J	100	5	0	16	Variable	60	60	10	1
J	10 000	2	0	16	Uniform	60	60	10	2[e]
H	100	2.6	0	14	Uniform	60	60	25	5
H	100	5	0	14	Uniform	60	60	25	1
H	100	2.6	0	14	Variable	60	60	25	1
H	100	5	0	14	Variable	60	60	25	1
H	10 000	2.6	0	14	Uniform	60	60	25	2[e]
K	100	3.5	0	13	Uniform	60	60	25	5
K	100	3.5	30	13	Uniform	60	60	25	1
K	100	3.5	60	13	Uniform	60	60	25	1
K	100	3.5	0	13	Variable	60	60	25	1
K	100	5	0	13	Uniform	60	60	25	1
K	100	5	0	13	Variable	60	60	25	1
K	10 000	3.5	0	13	Uniform	60	60	25	2[e]
K	10 000	3.5	30	13	Uniform	60	60	25	1
K	10 000	3.5	60	13	Uniform	60	60	25	1

[a]Number of simulated objects in the image.
[b]Assumed sky brightness.
[c]Assumed star forming region background emission.
[d]Number of simulation runs.
[e]One simulation is performed for a uniform stellar density distribution and the other one for a density profile falling as $1/r$ from the centre of the star forming region.

(3 e$^-$/pix) of the detector and the mean sky value of the image. These quantities are used to define the random error per pixel, σ:

$$\sigma = \sqrt{\overline{\rm sky} + 2\,{\rm RON}} \tag{10.1}$$

A scaling factor set by the user fixes the threshold for the star search. In our case the threshold is fixed at $4, 12, 17$ for the J, H, K-band images. This is done by the routine FIND, which performs a rough estimate of the centroid and flux of the stars. To avoid the inclusion of non-stellar objects (extended sources, cosmic rays, cosmetic defects) in the catalogue, DAOPHOT defines two shape parameters, *sharpness* and *roundness*, that can be used for a first selection of the sources. This selection is particularly useful to reject fake detections on the images, such as the diffraction rings of the brightest stars (see Fig. 10.6).

- We then perform aperture photometry with the routine PHOT on all the objects detected, estimating the flux in an aperture of radius 5 pixels. This radius must be small enough to contain negligible contamination by the light from neighbouring stars. The contribution of the local sky background is estimated in a circular annulus around the object (r_1=7, r_2=20 pixel) under the assumption that the spatial scale variation of the background is smaller than the aperture, and then subtracted. There are two obvious drawbacks affecting this approach: (i) the crowding of the observed star forming region means that the star profiles overlap, so that it is very difficult to take into account the contribution of nearby objects; (ii) the measurement of the very faint giant planet-mass objects, characterised by a very limited number of counts, is dominated by the background error, fluctuations and statistics.

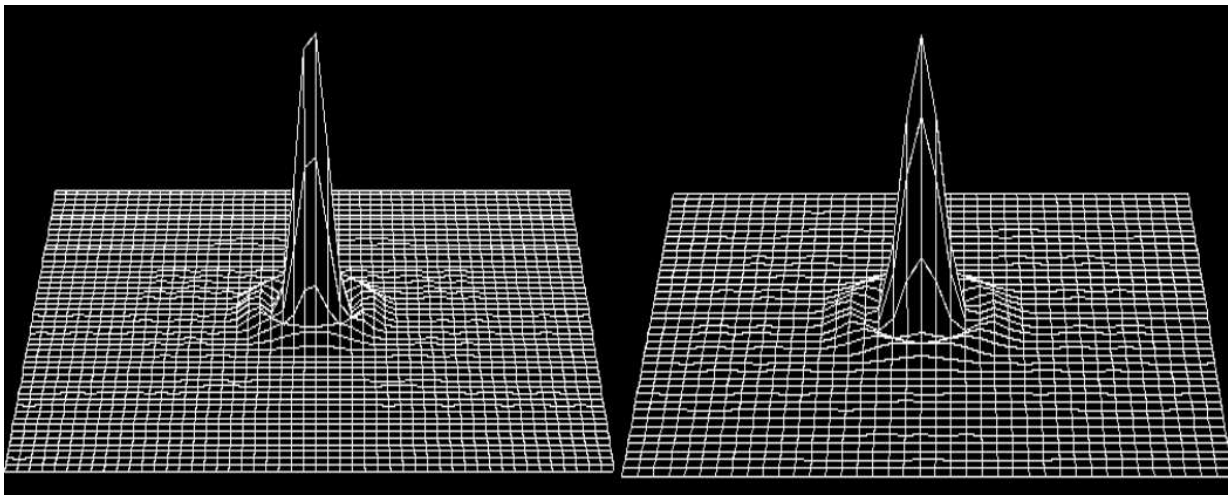

Figure 10.11: Left: 3D-profile of a star in a K-band image. Right: 3D-profile of the empirical K-band PSF reconstructed by DAOPHOT.

- We thus need to perform profile fitting photometry. We select a few bright, isolated and non-saturated stars in order to reconstruct an empirical PSF. DAOPHOT models the PSF with a mixed technique, in the sense that it uses the analytic formula plus a numerical map of the residuals. A Moffat function with β=2.5 is used for the analytical part, using a fitting radius of 2–3 pixels depending on the image pixel scale. We use this value in order to decrease the effect of the crowding, especially when fitting the very faint objects. The PSF is defined in a region with an area of radius 20 pixels (PSF radius, see Table 10.3). Given the list of selected PSF stars, the routine PSF fits a simple analytic formula with non-linear least squares. Then, all the observed residuals from the fit are interpolated on a common grid and averaged into a numerical look-up table of residuals. We assume that the PSF is constant all over our field of view ($\sim$2–10 arcsec^2). The comparison between the image of the empirical PSF and of a star for the case of the K-band with uniform background is shown in Fig. 10.11. It is worth noticing that the technique of the analytic function plus the numerical matrix works very nicely in reconstructing the PSF. On the other hand, Fig. 10.7 shows that our PSF does not perfectly reproduce the shape of the stars: a central brightness peak is left in the subtracted image. This residual flux is of the order of 2–3% and we take it into account when estimating an aperture magnitude correction (see discussion below). The result could be improved, for example, by feeding directly to DAOPHOT the analytical fit of the simulated PSF.

- We use this PSF and the aperture magnitude catalogue to perform profile fitting photometry on the single images with the task ALLSTAR. This program tries to exploit all the information available, performing a simultaneous fit of all the sources. During each iteration, it subtracts all the stars from a working copy of the image, according to the best estimates of the position and the magnitude of each object, computes the increments to the positions and magnitudes of the subtraction residuals around each position, and then checks each star to see whether it has converged or not. When a star has converged, it is permanently subtracted from the working image and the results are saved. A non-linear least square fitting technique is applied to derive the parameters of each object. ALLSTAR also writes out the final subtracted image, which can be visually inspected in order to check the quality of the fit and look for any faint lost or discarded objects.

- Once all the images have been reduced with ALLSTAR, a single catalogue containing the position and the magnitude for each detected star is available for each frame. In order to produce a unique catalogue with common positions and weighted mean magnitudes we use the tasks DAOMATCH and DAOMASTER. These are two programs that estimate the coordinate transformations between different star catalogues and allow to rescale them to a single reference coordinate system. Once that all the star lists are rescaled, it is possible to create a catalogue with weighted mean magnitudes. We use this as an input catalogue to perform profile fitting photometry simultaneously on all

the images with the task ALLFRAME. ALLFRAME starts with the master list of stars, which are in the x, y coordinate system of a reference image, inverts the coordinate of each star in each frame, and performs a local fit using the individual PSF. By applying a non-linear least squares fitting technique, during every iteration a new estimate of the position and magnitude of each star is provided. The position is rescaled in the reference frame coordinate system, and then averaged. This ensures that, in case the transformations are accurate enough, the centroid of the stars is known with great precision. This is probably the key point of this package: the accuracy in the centroid determination allows an optimal fit of stellar profiles, which ensures precise estimates of the magnitudes. Moreover, by using the same list of stars for all the frames, it is possible to reach a deeper magnitude limit, in the sense that the a priori knowledge of the existence of a star in a specific position makes it possible to fit it even if the S/N is very poor and, therefore, the fit on the individual image has failed. ALLFRAME writes out the final subtracted images too.

- After the ALLFRAME run, we stack all the subtracted images in a median image, and perform a search for any remaining stars in it. The S/N ratio of faint objects is raised, making their detection much easier (see Fig. 10.8). Unfortunately, in the search for these very faint stars, we also identify many spurious objects. In order to reject them, we run a program, SEPARATION (Stetson et al. 2003), that estimates the degree of crowding of each star of the catalogue. In most cases, we select all stars with *sep* > 5, i.e. stars that have required a correction of less than 1% for light contributed by known neighbours. The newly identified objects are then added to the input catalogue and ALLFRAME is run again. This procedure can be repeated until no more detectable objects, according to the assumed FIND threshold, are found in the subtracted median image.

In order to compare the final photometric catalogue with the input one, we have to transform the recovered PSF instrumental magnitudes to absolute magnitudes. We first estimate the aperture correction selecting a few bright and isolated stars on the frame. The corrections used in our simulation runs are listed in the fourth column of Table 10.3.

After applying the aperture correction we have to correct the flux for the fraction of energy outside of the PSF area. The encircled energy (EE) values adopted in the different simulation runs are listed in Table 10.3. We interpolated the EE values given in [RD1] for the LTAO PSFs to obtain the EE corresponding to our PSF area.

The result of the process described above is a catalogue with stellar positions, input and recovered magnitudes for each simulation run.

Table 10.3: Encircled energy (EE) and aperture corrections adopted for different pixel scales and PSF radii.

Band	Pixel scale mas/pix	PSF radius mas	Ap. corr.[a] mag	EE[b] %	EE corr.[c] mag
J	2	40	-0.37	21.8	-1.65
J	5	100	-0.49	36.2	-1.10
H	2.6	52	-0.17	39.7	-1.00
H	5	100	-0.23	48.2	-0.79
K	3.5	70	-0.10	57.6	-0.60
K	5	100	-0.08	60.6	-0.54
K_{zd30}	3.5	70	-0.20	57.4	-0.60
K_{zd60}	3.5	70	-0.33	33.5	-1.19

[a] Aperture correction applied to the PSF magnitudes.
[b] Percentage of energy enclosed in the PSF area of radius equal to the PSF radius.
[c] EE correction applied to the PSF magnitudes.

10.5.3 Compliance with figures of merit

As a first step we compare the input and the recovered luminosity function and estimate the magnitude scatter, as defined in Section 10.3, $S = \sqrt{\frac{1}{N}\sum(\mathrm{input} - \mathrm{recovered})^2}$. We adopt a bin size of 0.8 mag in order to have more than 2 stars to estimate the scatter. Fig. 10.12 shows the comparison between the input and the recovered K-band magnitude for the simulation runs with uniform background, pixel scale of 3.5 mas, and three different zenith distance values, namely $ZD = 0°$ (top panel), $30°$ (middle), $60°$ (bottom). It is worth noticing how the completeness of the recovered sample for a given magnitude scatter decreases for $ZD = 60°$.

Table 10.4 lists the limiting magnitudes reached with a $S/N \approx 4$ in the different simulation runs and the magnitude at which the magnitude scatter is 0.1 and 0.2 mag.

We can summarise the results listed in Table 10.4 discussing the influence of the different input parameters:

Background: Both the limiting magnitude and the scatter are affected by the presence of the variable background, in particular in the case of the J-band. At fixed pixel scale, the limiting magnitude and the magnitude at which the scatter is $S \sim 0.2$ mag decreases, when going from a simulation with a uniform background to one with a variable background, by $\sim$1 mag in the J-band, by $\sim$0.5 mag in the H-band, and by $\lesssim 0.5$ mag in the K-band. The magnitude at which $S \sim 0.1$ mag decreases by about 1 mag in the case of the J-band, by $\sim$0.7 mag in the H-band and by $\sim$0.9 mag in the K-band.

Pixel scale The limiting magnitude decreases by $\sim$3, 2 and 1 mag for the J, H and K band, respectively, when increasing the pixel scale. The scatter increases, and the magnitude at which is $S \sim 0.1/0.2$ mag decreases by more than 3 mag in the J and H bands, and by $\sim$1 mag in the K-band. As a matter of fact, decreasing the spatial resolution affects the detection and the measurement of the fainter objects (see Fig. 10.19). These tend to disappear in the wings of the PSF of much brighter stars. Furthermore, the sky background estimate is less accurate due to the increased pixel size.

Zenith distance The limiting magnitude decreases by $\sim$1 mag by moving to $ZD = 30°$ or $ZD = 60°$. The magnitude at which the scatter is $S \sim 0.2$ mag decreases by $\sim$0.6 mag and does not really change when $S \sim 0.1$ mag in the case of $ZD = 30°$. On the other hand, for the case with $ZD = 60°$ the magnitudes at which the scatter is $S \sim 0.1$ and 0.2 mag decrease by more than 1 mag.

Table 10.4: Limiting magnitudes for different simulation runs.

Band	Pixel scale mas/pix	Lim. mag[a] mag	Scatter[b] mag	Scatter[c] mag
J_{Uni}	2	30.2	29.5	29.0
J_{Var}	2	29.0	28.7	28.0
J_{Var}	5	26.0	25.5	24.0
H_{Uni}	2.6	29.6	29.2	28.0
H_{Var}	2.6	29.2	28.6	27.3
H_{Var}	5	27.2	25.8	24.0
K_{Uni}	3.5	29.8	29.3	28.5
K_{Var}	3.5	29.5	28.8	27.5
K_{Var}	5	28.4	27.8	26.2
K_{zd30}	3.5	29.0	28.7	28.5
K_{zd60}	3.5	28.8	28.0	27.4

[a] Limiting magnitude at $S/N \approx 4$.
[b] Magnitude at which S = 0.2 mag.
[c] Magnitude at which S = 0.1 mag.

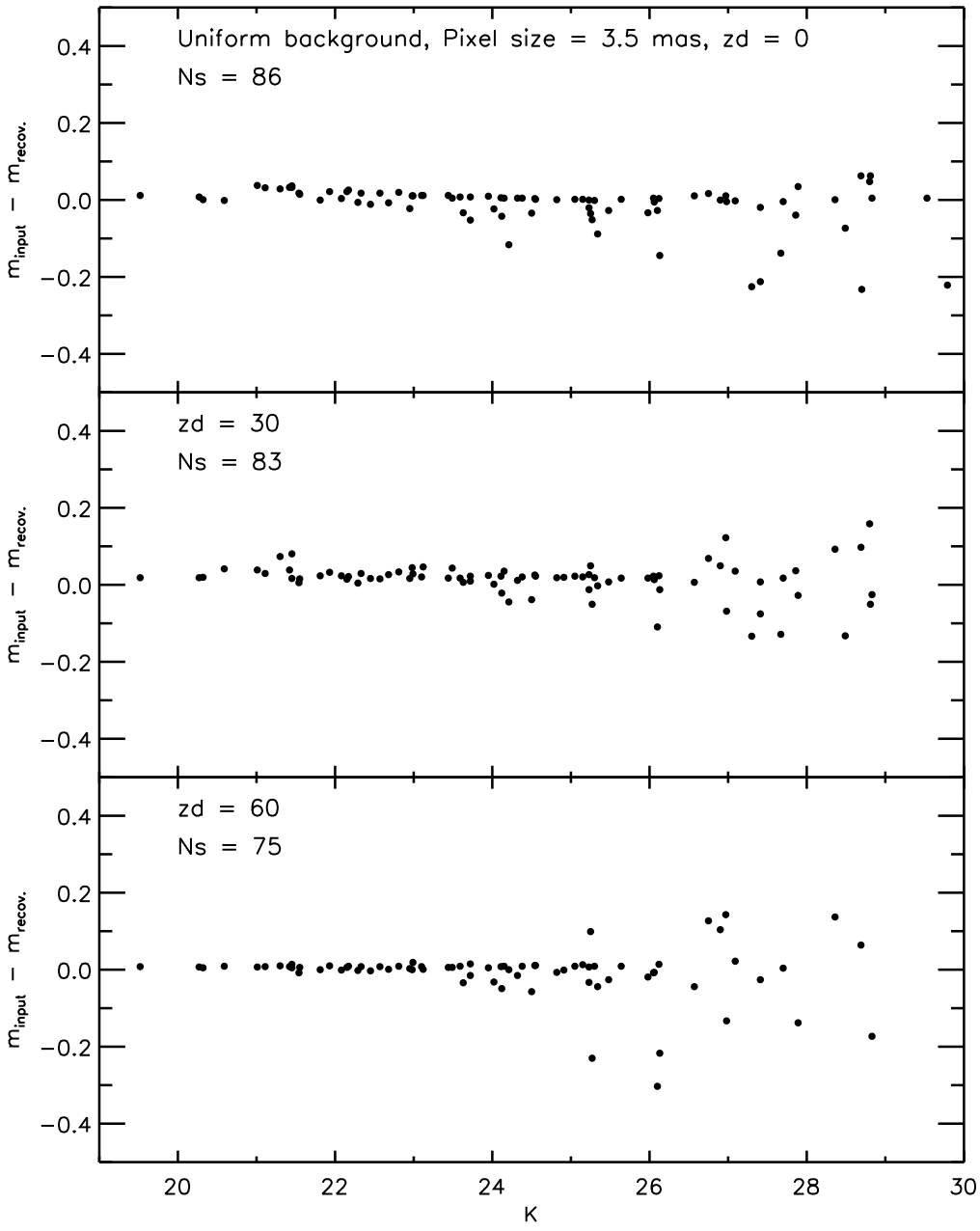

Figure 10.12: Top: Difference between the input and the recovered magnitude, $m_{\rm input} - m_{\rm recovered}$, plotted versus the input magnitude K for a simulation with uniform background, pixel scale of 3.5 mas and zenith distance, $ZD = 0°$. Middle: Same as top but for $ZD = 30°$. Bottom: Same as top but for $ZD = 60°$.

10.5.3.1 Completeness analysis

The completeness analysis is performed only for the 10 000 star simulation case, where we have enough statistics in order to sample the completeness function well. We plot in Figs. 10.13, 10.15 and 10.17 the recovered fraction of stars versus the input K, H and J magnitudes. The background and the stellar density distribution are assumed to be uniform, and the pixel scale is 3.5, 2.6, and 2 mas, respectively. Fig. 10.13 also shows the recovered fraction of stars versus K-band magnitude for the case in which stars are distributed according to the $1/r$ density profile (stars). The bin size used in this plot is 0.4 mag in order to have a significant number of stars ($\geq$ 2) in each magnitude interval.

To these data we fit the following completeness function:

$$C = [e^{(\mathrm{mag}-\mathrm{mag}_{50})/\alpha} + 1]^{-1} \tag{10.2}$$

where mag_{50} is the magnitude at which the completeness is 50%, and α is a shape parameter. Table 10.5 lists the magnitude at which the recovered fraction of stars is 90 and 50% for eight different simulation runs. The data plotted in Fig. 10.13 and listed in Table 10.5 indicate that the stellar density affects the result of the simulations, in particular in the case of the K-band. The K magnitude at which the recovered fraction of stars is 90% decreases by $\sim$0.3 mag ($K \sim$ 28.6–28.3 mag). This means that the mass of the objects at which the completeness is 90% increases from 12 to 17 M_{Jup}. A similar trend is also present in the case of the J and the H bands. The zenith distance influences our results only in the case of $ZD = 60°$. The magnitudes at which the completeness is 90 and 50% decreases by $\sim$0.7 mag.

Figs. 10.14, 10.16 and 10.18 show the recovered fraction of stars versus the object mass. The completeness function we adopt in this case is:

$$C = [e^{(\mathrm{mass}_{50}-\mathrm{mass})/\beta} + 1]^{-1} \tag{10.3}$$

where mass_{50} is the mass at which the completeness is 50%, and β is a shape parameter. Table 10.5 also lists the mass at which the recovered fraction of stars is 90 and 50% for the eight different simulations.

Table 10.5: Recovered fraction of stars at different magnitudes and masses for the 10 000 star simulation runs.

Band	90%[a] mag	50%[b] mag	90%[c] M_{Jup}	50%[d] M_{Jup}
J	29.15	29.4	24	16
$J_{1/r}$	29.05	29.4	27	18
H	28.8	29.4	18	12
$H_{1/r}$	28.7	29.3	22	14
K	28.6	29.4	12	9
$K_{1/r}$	28.3	29.2	17	9
K_{zd30}	28.5	29.3	12	9
K_{zd60}	28.0	28.7	15	11

[a]Magnitude at which the recovered fraction of stars is 90%.
[b]Magnitude at which the recovered fraction of stars is 50%.
[c]Mass at which the recovered fraction of stars is 90%.
[d]Mass at which the recovered fraction of stars is 50%.

The data plotted in Figs. 10.13 to 10.18 and listed in Table 10.5 indicate that we can partly achieve the scientific goals of the proposal (see Section 10.2). In particular, we should take into account both the uniform and the variable background simulation runs to draw our conclusions. It is noteworthy that the influence of the non-uniform nebulosity on the results is to be expected. This is blue reflection nebulosity, therefore brighter at shorter wavelengths. Moreover, we should consider the effect of observing at zenith distances larger than $ZD = 30°$.

***J*-band:** The completeness is $\sim$90% at $J \sim$ 29.1 mag ($M \sim$ 24 M_{Jup}), with $S <$ 0.1 mag and it is $\sim$50% at $J \sim$ 29.4 mag ($M \sim$ 16 M_{Jup}), with $S <$ 0.1 mag, for a simulation with pixel scale 2 mas and a uniform

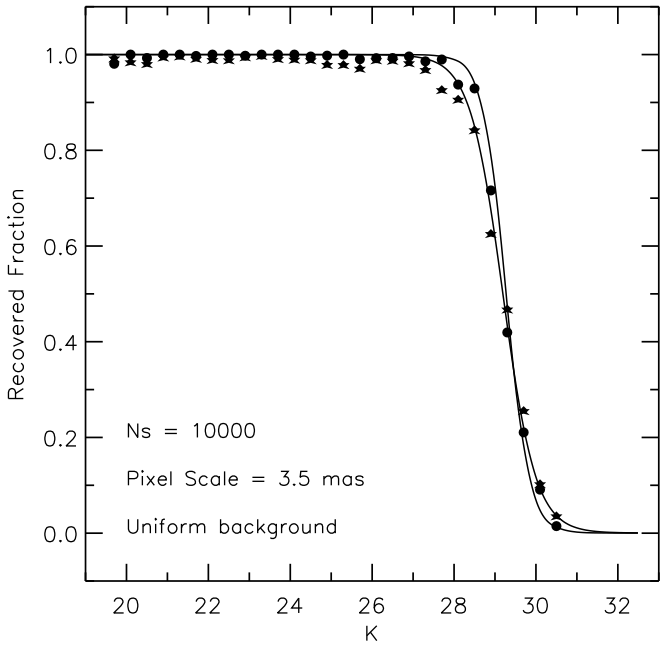

Figure 10.13: Recovered fraction of stars plotted versus the K-band input magnitude in bins of 0.4 mag for simulations with uniform background, a pixel scale of 3.5 mas, $ZD = 0°$ and a uniform stellar density distribution (filled circles) and a stellar density distribution scaling as $1/r$ (stars). The completeness function is overplotted as a solid line.

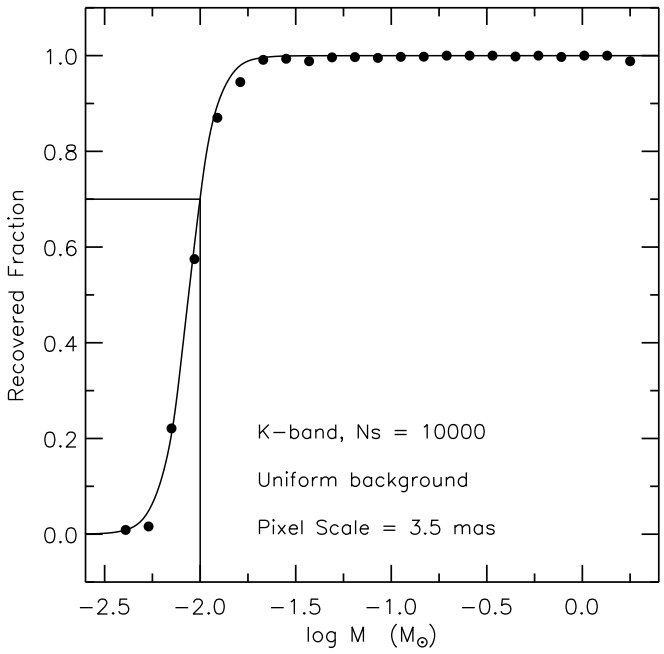

Figure 10.14: Recovered fraction of stars plotted versus the input mass is bins of 0.12 $M_{\odot}$. The simulation is performed with a uniform background, a pixel scale of 3.5 mas, $ZD = 0°$ and a uniform stellar density distribution. The completeness function is overplotted as a solid line. Note that $\log M(M_{\odot}) = -2$ corresponds to 10 $M_{\rm Jup}$, where the recovered fraction is ∼70%.

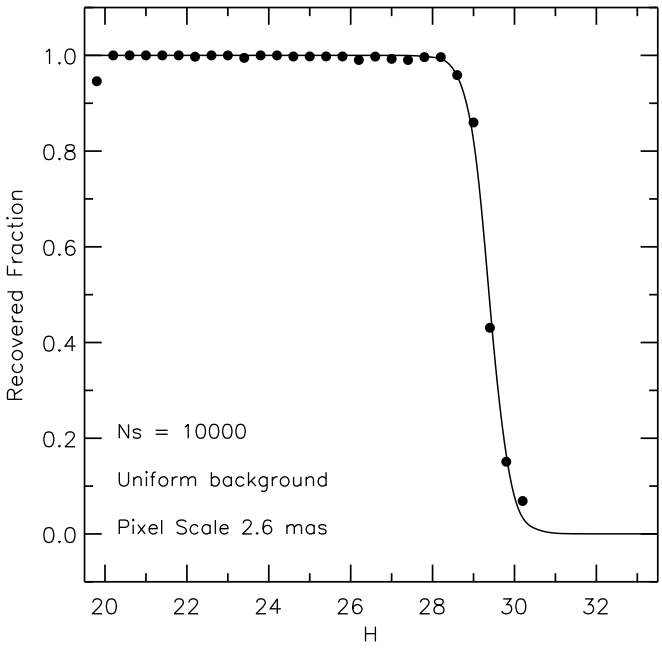

Figure 10.15: Same as Fig. 10.13 but for the H-band. The simulation is performed with a uniform background, a pixel scale of 2.6 mas and a uniform stellar density distribution.

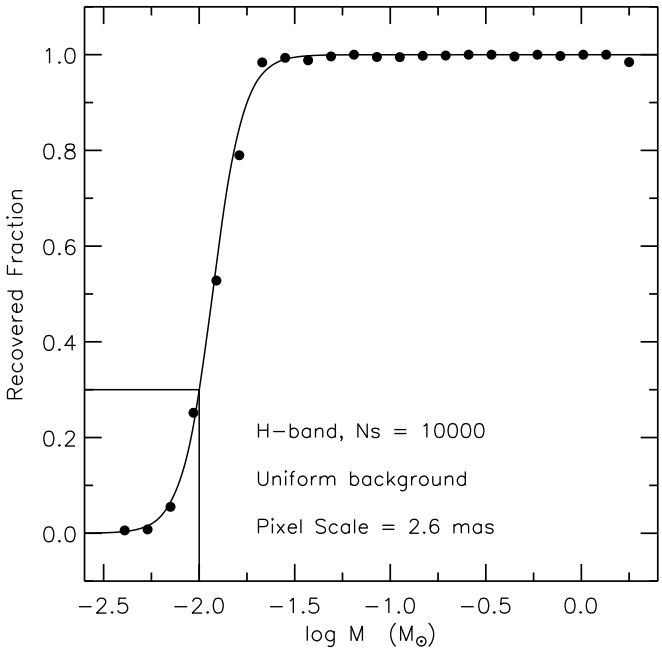

Figure 10.16: Same as Fig. 10.14 but for the H-band. The simulation is performed with a uniform background, a pixel scale of 2.6 mas and a uniform stellar density distribution. Note that $\log M(M_\odot) = -2$ corresponds to 10 $M_{\rm Jup}$, where the recovered fraction is ∼30%.

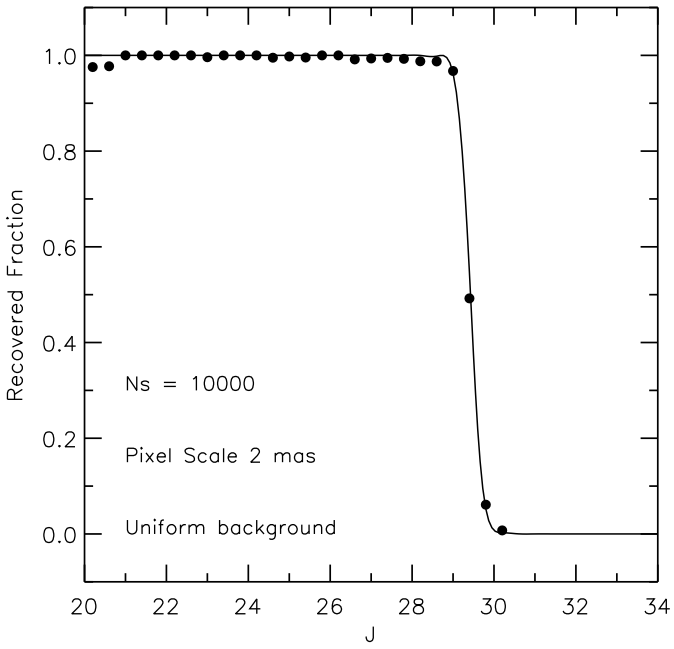

Figure 10.17: Same as Fig. 10.13 but for the J-band. The simulation is performed with a uniform background, a pixel scale of 2 mas, and a uniform stellar density distribution.

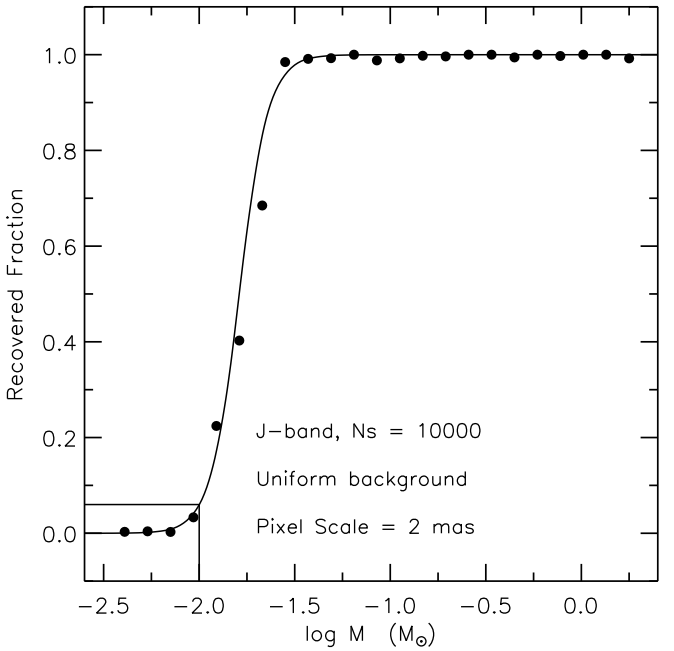

Figure 10.18: Same as Fig. 10.14 but for the J-band. The simulation is performed with a uniform background, a pixel scale of 2 mas, and a uniform stellar density distribution. Note that $\log M(M_{\odot}) = -2$ corresponds to 10 $M_{\rm Jup}$, where the recovered fraction is $< 10\%$.

background. If we account for a variable background, the scatter at $J \sim 29.1$ mag would be larger than 0.1 mag, and at $J \sim 29.4$ mag it would be larger than 0.2 mag, as we expect.

***H*-band:** The completeness is ∼90% at $H \sim 28.8$ mag ($M \sim 18\ M_{\rm Jup}$), with $S \sim 0.1$ mag and it is ∼50% at $H \sim 29.4$ mag ($M \sim 11\ M_{\rm Jup}$), with $S < 0.2$ mag, for a simulation with pixel scale 2.6 mas and a uniform background. If we account for a variable background, the scatter at $H \sim 28.8$ mag would be ∼0.2 mag, and at $H \sim 29.4$ mag it would be larger than 0.2 mag.

***K*-band:** The completeness is ∼90% at $K \sim 28.6$ mag ($M \sim 12\ M_{\rm Jup}$), with $S < 0.2$ mag and it is ∼50% at $K \sim 29.3$ mag ($M \sim 9\ M_{\rm Jup}$), with $S < 0.2$ mag, for a simulation with pixel scale 3.5 mas and a uniform background. The result does not significantly change when we assume the presence of a variable background, as we expect since the reflection nebulosity is brighter at shorter wavelengths. On the other hand, observing at $ZD = 60^{\circ}$ would lead to a completeness of ∼90% at $K \sim 28$ mag ($M \sim 15\ M_{\rm Jup}$) with $S \sim 0.2$ mag. The zenith distance then affects not only the detection limit but also the final photometric accuracy of the identified objects.

10.5.4 Sensitivity to input parameters

We performed different simulation runs changing various input parameters, as already mentioned in Section 10.4.1 and 10.5.1. The most critical input parameter is the pixel scale, which affects differently the three NIR bands. We adopted the pixel scale of the LTAO PSF sampling in each band, and a fixed 5 mas pixel scale, as requested in the proposal. The sky background estimate is less accurate when increasing the pixel scale (see Section 10.5.3). Furthermore, the decrease of the spatial resolution makes the faint object detection more difficult, since they disappear in the PSF wings of the brighter neighbours. This results in shallower limiting magnitudes and less accurate photometry. This effect is stronger in the J-band simulations since we increased the spatial sampling from 2 to 5 mas. Moreover, simulations at different pixel scale are performed for the case of variable background, for which the J-band is more affected (see Section 10.5.3 and Fig. 10.20). Fig. 10.19 shows a section of a J-band image with pixel scale 2 mas (left panel) and 5 mas (right), for a simulation with variable background. It is worth noticing how the stars are more defined in the 2 mas image so that the centroids can be estimated with more accuracy, and the increased sky background in the 5 mas image due to the bright stars of the region.

We plot in Fig. 10.20 the adopted pixel scale versus the J, H, K-band limiting magnitudes reached with $S/N \approx 4$ in the simulations with the variable background (see Table 10.4). The band which is most affected by this parameter is J, where the limiting magnitude decreases by ∼3 mag when moving from a 2 to a 5 mas pixel scale.

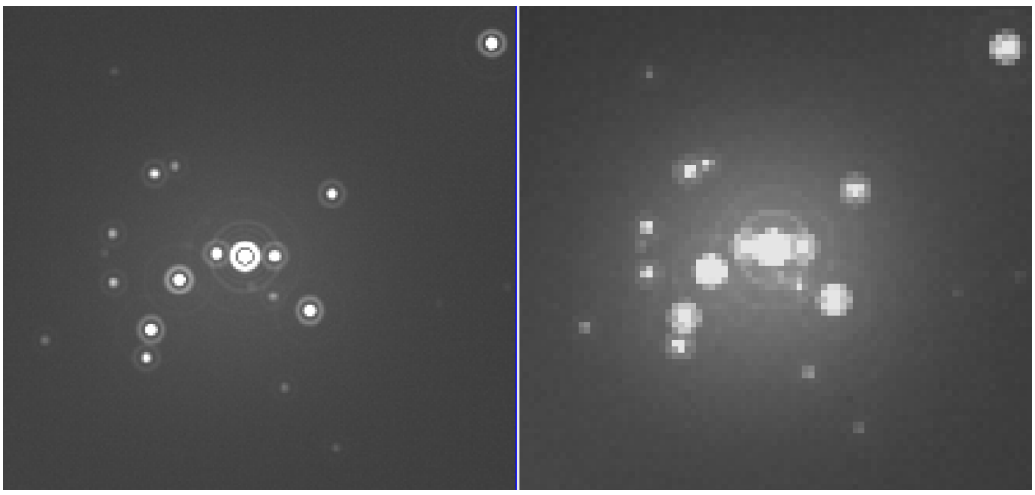

Figure 10.19: Left: J-band image for a simulation with variable background and pixel scale 2 mas. Right: Same but for a pixel scale of 5 mas.

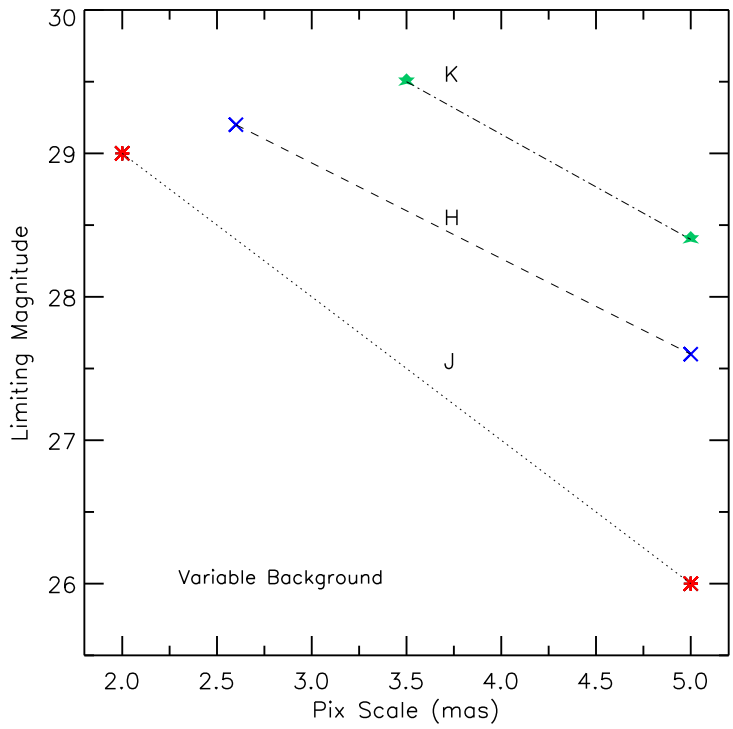

Figure 10.20: Limiting magnitudes at $S/N \approx 4$ plotted versus the pixel scale.

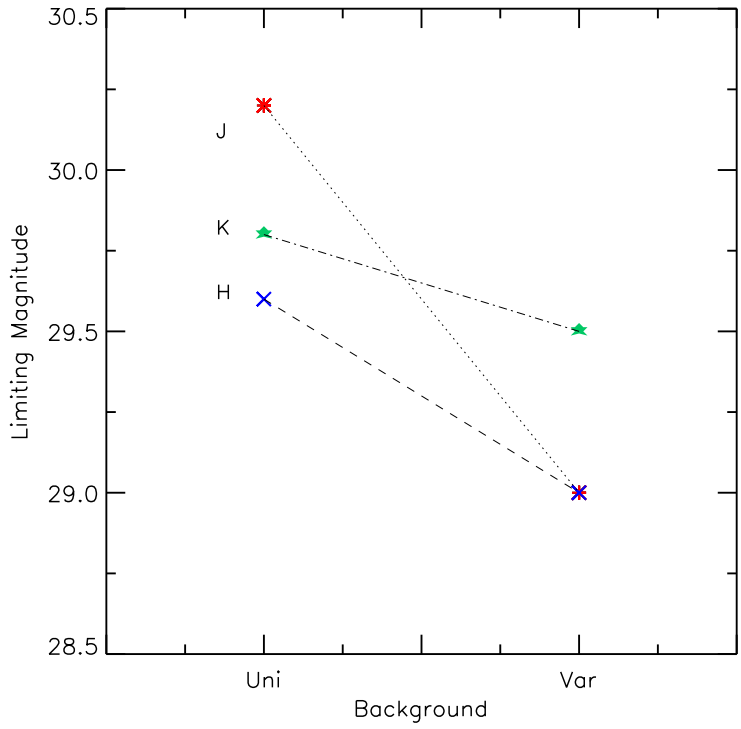

Figure 10.21: Limiting magnitudes at $S/N \approx 4$ plotted versus the assumed background (uniform or variable)

The magnitude scatter is also affected by the adopted pixel scale. The magnitude at which the scatter is 0.1 mag decreases by more than 3 mag in the cases of the J and H bands, and by more than 1 mag for the K-band.

The other critical parameter is the presence of a variable background. The variable background is introduced in the simulations as an unresolved background (i.e. the physical variation length scale is smaller than the PSF FWHM) and it then acts like an additional source of noise, except for the fact that it does not decrease when increasing the exposure time. Fig. 10.21 shows the limiting magnitudes reached with $S/N \approx 4$ versus the assumed type of background. As we already discussed in Section 10.5.3, the band which is most affected is J: the limiting magnitude decreases by more than 1 mag. This difference is smaller ($\sim$0.6 mag) in the case of the H-band, and it has an almost negligible effect on the K-band ($\sim$0.3 mag). The magnitude at which the scatter S is 0.1/0.2 mag also decreases by the same amount for the three bands.

The simulations also showed that the zenith distance affects not only the detection limit but also the photometric accuracy. Observing in the K-band at zenith distances from 30 to 60° decreases the limiting magnitude by $\sim$1 mag. Moreover, the magnitude at which the scatter is 0.1 and 0.2 mag decreases by $\sim$1 mag too.

10.5.5 Calibration requirements

We will adopt a standard observation strategy and the regions that we will observe in the LMC could be selected on the basis of their thermal emission detected by other facilities.

No special calibration requirements are needed. The target fields are expected to contain sufficiently bright stars near the centre to allow an on-target accurate determination of the PSF, which will be taken as constant due to the very small angular size of the target aggregates. These stars will also be included in wide-field surveys to be carried out in the coming years by survey telescopes like VISTA or in dedicated support observations by instruments such as HAWK-I at the VLT, thus providing built-in flux zero-points in the target fields.

10.5.6 Limitations

The completeness analysis has been performed only for the simulation runs with uniform background. Moreover, the simulations for different zenith distance values have been performed only for the K-band, so we cannot quantify this effect in the other bands. It is noteworthy that to accomplish the goal of the proposal, i.e. to identify the faintest giant planet-mass objects in the LMC star forming regions, we need to measure at least one colour. For these reasons the results quoted in the previous sections should be regarded as a best estimate.

Finally, we point out that although DAOPHOT is a well-tested package it is not usually used to analyse high-Strehl images as in this work. Hence it may be worth comparing the results derived here with DAOPHOT with those obtained with other analysis packages such as STARFINDER. However, this comparison is beyond the scope of this report.

10.6 Concluding remarks

As already discussed in Section 10.5.3, our simulations indicate that we can partly achieve the scientific goals of this science case. In particular, the simulations show that nearly complete samples of young brown dwarfs above the deuterium-burning limit ($M \geq 13\ M_{\rm Jup}$) can be obtained in LMC star forming regions with the E-ELT, and that even giant planet-mass objects below that limit can be detected under favourable conditions. By favourable conditions we mean star forming regions with an almost uniform background and observing at zenith distances less than 30°. Furthermore, it is noteworthy that these results can be accomplished only if the instrument pixel scale is $<$ 4 mas.

References

Baraffe I., et al., 2003, A&A, 382, 563
Bate M.R., 2005, MNRAS, 363, 363
Boyd D.F.A., Whitworth A.P., 2005, A&A, 430, 1059
Caballero J.A., et al., 2007, A&A, 470, 903
Cardelli J.A., Clayton G.C., Mathis J.S., 1989, ApJ, 345, 245
Chabrier G., 2003, PASP, 115, 763
Chabrier G., 2005, in “The initial mass function 50 years later”, eds. Corbelli, Palla, Zinnecker, Springer
Freedman W.L., et al., 2001, ApJ, 553, 47
Grazian A., Fontana A., De Santis C., Gallozzi S., Giallongo E., Di Pangrazio F., 2004, PASP, 116, 750
Hunter I., et al., 2007, A&A, 466,277
Marley M.S., Fortney J.J., Hubickyj O., Bodenheimer P., Lissauer J.J., 2007, ApJ, 655, 541
Rees M.J., 1976, MNRAS, 176, 483
Stetson P.B., 1987, PASP, 99, 191
Stetson P.B., 1991, AJ, 102, 589
Stetson P.B., 1994, PASP, 106, 250
Stetson P.B., Bruntt H., Grundahl F., 2003, PASP, 115, 413
Zapatero Osorio M.R., et al., 2002, ApJ, 578, 536

11 G4-1: The resolved stellar populations of elliptical galaxies

Authors: J. Liske, E. Tolstoy

11.1 The science case

One of the main unsolved questions in astronomy is how galaxies form and evolve. In 1936 Edwin Hubble presented his tuning fork diagram classifying the different types of galaxy to be found in the Universe. This was the first attempt to find a pattern in the properties of different systems and thus search for a common evolutionary link. Elliptical galaxies appear to be shaped predominantly by a single component composed of old stars, while spiral galaxies have several components with a range of stellar ages, gas disks, dust and stellar bars. This 'Hubble tuning fork' diagram survives until today as the standard manner of presenting the morphologies of galaxies, with the addition of small dwarf spheroidal and irregular type galaxies. Fitting all these galaxy types into a common evolutionary scenario still remains to be achieved. According to the most widely accepted current structure formation scenario (the Cold Dark Matter, or CDM paradigm), all galaxies are built up from smaller pieces, the fundamental building blocks of galactic evolution, coalescing through time starting in the early universe (at high redshift) to form the galaxies we see today. The present day systems therefore provide a unique insight into the galactic assembly process and the chemo-dynamical evolution of the galaxy distribution.

The directly observable components of any galaxy are gas, dust and stars, and there is an intimate link between them. Stars form from gas, and synthesise elements in their interiors and the stars that explode at death disperse these elements into the gas from which dust and subsequent generations of stars are formed. All galaxies are thus the integrated products of all the star formation during their entire lifetimes, and the chemical elements in the stellar populations of different ages provide the most detailed evidence for this past star formation. Because low mass stars can have lifetimes comparable to the age of the Universe, the low mass tail of the ancient star formation that occurred at the formation epoch of a galaxy remains visible today and provides unique clues to the earliest physical process in the Universe. Stars of all ages provide an accurate and detailed probe of changing galaxy properties. By observing large numbers of individual stars we can measure how the rate of star formation and chemical composition of a galaxy has varied from its formation to the present and thus how galaxies were built up over time. To unravel this formative epoch detailed spatial, kinematic and chemical surveys of resolved stellar populations are required; providing a unified picture between local near-field cosmology, predictions from high redshift surveys and theoretical simulations of galaxy formation and evolution.

Until now the sensitivity and resolution limitations have meant that detailed studies have only been possible within the Local Group and specifically around our own Galaxy. This means that only small dwarf type galaxies have had their ancient stellar populations accurately probed; massive galaxies still await this careful scrutiny. The Local Group (LG) contains only two massive galaxies (spiral systems M31 and the Milky Way) and around 40 smaller, mostly dwarf, galaxies. This is hardly representative of the range of galaxy types, and our LG is not necessarily representative of the high-density regions of the Universe where most galaxies live. Careful studies of dwarf galaxies have already shown inconsistencies between observations and the standard CDM picture, and these need to be extended to larger galaxies to make an accurate comparison with the properties of small galaxies. To make significant progress we need to study large numbers of resolved stars in a range of galaxy types and this requires us to look beyond the halo of the Milky Way.

The ultimate goal is to study the resolved stellar populations in giant elliptical galaxies, of which there is no example in the LG, and we have to look at Cen A (at 4 Mpc) to find the closest example, of a peculiar elliptical. However, the best place to look at the properties of a range of elliptical galaxy types is the Virgo cluster which contains thousands of large galaxies and tens of giant ellipticals of a range of size and position in the cluster. Here we explore the feasibility of obtaining accurate photometry of the resolved stellar populations of elliptical galaxies, ideally over their entire extent including the dense central regions, in order to construct their colour magnitude diagram (CMD), over a range of distances out to the Virgo cluster.

11.2 Goals of the DRM simulations

The aim of this science case is to construct accurate CMDs of the resolved stellar populations (RSPs) of elliptical galaxies. The key points of interest are the tip of the red giant branch (TRGB), the horizontal branch (HB) as well as the main sequence turn-off (MSTO). The goal of the simulations is to quantify under which circumstances appropriately accurate photometry can be obtained in order to probe these key points in the CMD. Specifically, the simulations should determine the limiting magnitude down to which accurate photometry of a galaxy's RSP is possible as a function of:

- the galaxy's distance;
- the surface brightness within the galaxy, equivalent to galactocentric radius for a given profile;
- the observing band;
- the performance of the assumed Adaptive Optics (AO) system;
- the assumed stellar population.

Obviously, at very low surface brightnesses (or large radii), where individual stars are well separated from one another, the limiting magnitude can be determined by a straightforward S/N calculation. The key task for the simulations is hence to quantify how the photometry is affected by stellar crowding as one moves to the denser inner regions of a galaxy.

The simulations should also address the question of which combination of bands is best suited for the construction of CMDs with the E-ELT. In the past, CMDs were usually constructed in the optical wavelength range (e.g. in the V and I bands). However, at present it seems unlikely that the E-ELT will be able to deliver diffraction-limited performance in the optical regime. Nevertheless, the need for the highest possible resolution argues for the shortest possible wavelengths. On the other hand, the quality of the AO correction increases with wavelength. In addition, one would like to use as widely separated bands as possible, because the longer the wavelength base the more widely the different stellar populations are spread across the CMD, thus improving the diagnostic power for a given photometric error. Given these conflicting drivers it is not at all clear a priori what the best combination of bands might be.

We would also like to know what combination of field of view (FoV) and total exposure time is required in order to populate the CMD with sufficiently large numbers of stars, and how this is affected by the telescope's site and mirror coating.

Finally, we would like to know how the results are affected by uncertainties in the determination of the PSF. In particular, what is the effect of uncalibrated anisoplanatism, i.e. of an undetected variation of the PSF across the FoV?

11.3 Metrics / figures of merit

The metric with which to compare different simulations with different input parameters is the limiting magnitude at which the photometric error is 0.05 mag. We will consider the main scientific goal of this science case to be met if we can reach at least 1 mag below the TRGB with this photometric accuracy at 1 effective radius of an elliptical galaxy at the distance of Virgo.

11.4 DRM simulations

11.4.1 Methodology

The main aim of this study is to quantify the effect of crowding on the accuracy of stellar photometry. It is in fact possible to solve this problem analytically (see e.g. Olsen et al. 2003). Given (i) the knowledge of how exactly the photometry is performed, (ii) the PSF, and (iii) the surface number density of the background stars as a function of their apparent magnitude one can calculate the expected photometric error as a function of any parameter of interest. However, in this study we have taken the 'brute force' approach of obtaining the answers to the questions posed in Section 11.2 from Monte Carlo simulations.

Broadly speaking, we proceed in the following way. We simulate an image of a (small part of a) galaxy at a given distance and at a given surface brightness level by populating the image with an appropriate number of stars drawn from a luminosity distribution appropriate for an assumed stellar population. This image is then analysed to detect stars and to measure their fluxes. Comparing these to the known input fluxes we then determine the accuracy of the photometry as a function of the input flux. This process is then repeated for different combinations of the various scientific and technical input parameters, and the results are compared in order to assess the effect of varying the parameters of interest.

One of the aims of the present work is to identify the combination of observing bands that is best suited to imaging studies of RSPs with the E-ELT. However, we will a priori restrict this investigation to the red-optical to near-IR wavelength regime (I_C–K bands). As can be seen from Fig. 8.2 of [RD1] the fraction of energy enclosed in the diffraction limited core of an LTAO PSF drops to below 1 % for wavelengths shorter than $\sim$0.8 μm. In other words, for shorter wavelengths the AO correction becomes so ineffective that the diffraction limited core essentially disappears, and with it the high spatial resolving power so essential for this science case. On the other hand, at wavelengths longer than K the PSF's diffraction limited core becomes so large that these bands are unlikely to be competitive. Note that in this work we distinguish between the Cousins I_C-band and the longer wavelength Johnson I-band (cf. Table 9.1 of [RD1]; see also Section 11.4.3.1). Both are considered here because, as we will see, they span the wavelength range in which the AO correction makes the transition from being ineffective to being effective.

11.4.2 Pipeline

We now describe in detail the simulation and analysis sequence that will be used to address the issues raised in Section 11.2. An illustration of the simulation and analysis pipeline is given in Fig. 11.1.

11.4.2.1 The stellar population

The stellar population is defined by the following parameters: the abundance of α elements, the chemical composition (mainly identified by its metallicity, [Fe/H]), the slope and mass range of the initial mass function (IMF), and the star formation history (SFH). For a given set of values for these parameters a purpose-built piece of code assembles the isochrones corresponding to the selected α element abundance, chemical composition and age range of the SFH into a single output file. The code then first assigns a weight to each isochrone according to the specified SFH, followed by the assignment of a weight to each initial mass point along the isochrone according to the selected IMF. Finally, the inter and intra-isochrone weights are multiplied (and normalised) to arrive at a final weight for each evolutionary stage of each initial mass present in the stellar population. These weights are written to the output file along with the contents of the selected original BaSTI isochrone files, which includes among other things the initial mass, the absolute V-band magnitude and a number of colour terms.

Note that we do not perform any interpolation between isochrones. Although important for building entirely realistic stellar populations, this level of detail is not required for our purposes.

The code to build the stellar population is available here.

11.4.2.2 The input catalogue

The simulation of an image requires an input catalogue containing the x,y positions (in pixels) and apparent magnitudes of the stars to be included in the image. This is achieved in two steps.

First we generate a multi-band star catalogue. This requires as input: the stellar population file generated in the previous section, a distance modulus (DM), a V-band surface brightness, the size of the FoV to be simulated, the bands to be included in the catalogue, and the pixel sizes of the images to be simulated in each of the bands included in the catalogue. Given this information a purpose-built piece of code first works out the total number of stars required to fill the requested FoV with the desired average V-band surface brightness. It then randomly draws the required number of stars from the stellar population according to the weights attached to each evolutionary stage of each initial mass point present. For each star the code then calculates the apparent magnitudes in the requested bands from the absolute V-band

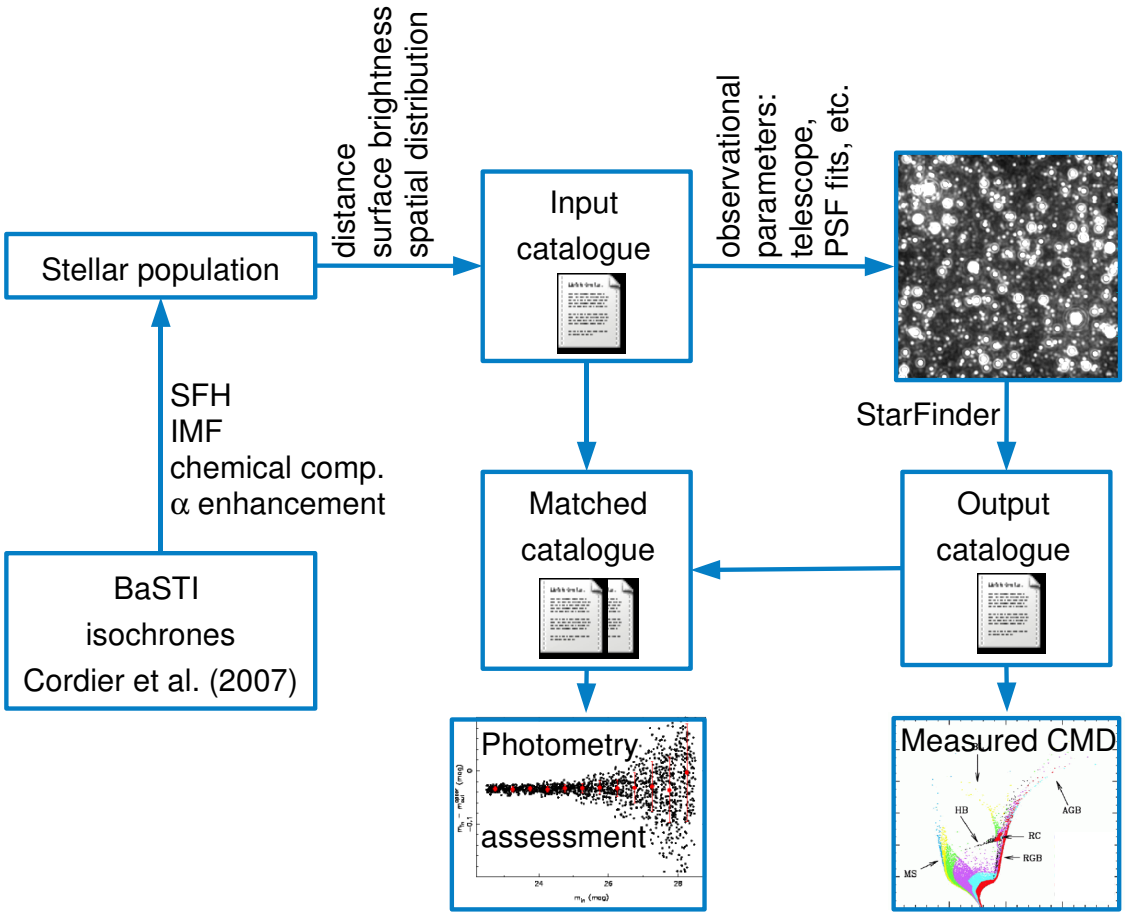

Figure 11.1: Flowchart of simulations.

magnitude, the appropriate colour terms and the DM. Finally, each star is allocated a random x,y image positions for each requested band according to the selected FoV size and pixel sizes.

From this multi-band catalogue we then extract the band-specific information (i.e. positions and magnitudes) into separate files which are subsequently used as input catalogues for the generation of the images.

Note that the above catalogue contains stars over the entire initial mass range that was specified for the stellar population. For some simulations this leads to a small complication: for the most distant galaxy and for highest surface brightnesses considered here the number of stars required can be so large as to make the image simulation impractical due to excessively large execution times. Mitigating this problem requires us to use a very small FoV. (Note that for the purpose of determining the photometric errors the size of the FoV is irrelevant. It does not matter whether we simulate one image of size, say, 10″ or four images of size 5″.) However, even for a FoV of just a few arcsec the required number of stars can be as large as a few $\times$ 10^7, whereas the number of stars included in the image simulations should not exceed $N_{\rm max} \approx 1.5 \times 10^6$ in order to keep execution times reasonable. Hence, whenever the required number of stars exceeds $N_{\rm max}$ we identify an initial mass threshold such that the number of stars above this threshold is $\approx N_{\rm max}$, and we exclude the stars below this threshold from the image simulation. This effectively lowers the surface brightness by up to a few $\times$ 0.1 mag arcsec^{-2} but we have verified that this does not significantly affect the photometric errors (because these are dominated by the brightest stars).

The code to build the multi-band stellar catalogue is available here.

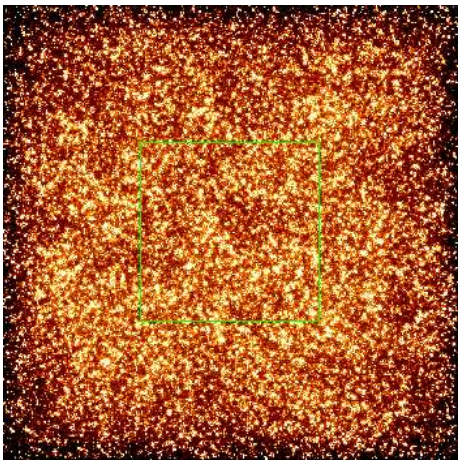

Figure 11.2: Low-resolution rendering of a simulated H-band image of a high-density field with μ_V = 19.5 mag arcsec^{-2} at distance modulus 31.2. The image was scaled to bring out the edge effect caused by the fact that MKOBJECTS does not include the light from stars beyond the image edge. The size of the image is 2.55″ but only the central 1″ region marked by the green box is entirely unaffected and may hence be used in the subsequent analysis.

11.4.2.3 Image simulation

Images are simulated using the IRAF task MKOBJECTS which requires as input the stellar catalogue generated in the previous section, a PSF (see Section 11.4.3.2), and a magnitude zero-point to convert the magnitudes into number of photons (see Table 9.1 of [RD1]). The latter is adjusted to take into account the size of the telescope's photon collecting area, the exposure time and the system throughput. For each star in the input catalogue MKOBJECTS adds an appropriately scaled copy of the PSF to the image at the requested position (see the MKOBJECTS documentation for the details of this process). We also include a spatially constant background (see Section 11.4.3.2) and add photon noise from the stars and the background. We do not need to consider read-out noise and Poisson noise from any dark current because these components are always sub-dominant compared to the photon noise from the background, even for short exposure times (which would possibly be required in reality to deal with the variability of the NIR sky) and for the small pixel sizes considered here. This also implies that we can simulate even long-duration observations of many hours as a single exposure without having to worry about the fact that in reality such an observation would consist of many hundreds of individual exposures.

The PSFs used for the image simulations do not just comprise the central parts around the diffraction limited cores but include significant fractions of the seeing halo, going down to contrast levels (with respect to the peak value) of 10^{-3} – 10^{-5} depending on the computational cost (see Section 11.4.3.2 for more details).

When creating the stellar catalogue above we already had to choose a FoV and a pixel size, and obviously the size of the simulated image in pixel is chosen to correspond to these values. When choosing the size of the FoV we take into account the fact that for high stellar densities (i.e. at high surface brightnesses) the images will suffer from a significant edge effect, as demonstrated in Fig. 11.2. Unlike a pixel near the centre of the image, one near the edge does not receive equal amounts of light (on average) from stars in all directions, because there are no more stars beyond the edge of the image. Even if we were to include stars beyond the image's edges in the catalogue, MKOBJECTS unfortunately ignores them. This affects all pixels that lie within one PSF radius (see Section 11.4.3.2 and Table 11.2) of an image edge. The photometry of objects lying in these areas will not have the noise properties it should, and hence these areas must be excluded from the analysis. Consequently, the size of the FoV used in this and the previous section must be larger than the desired one by two PSF radii. Since we need to keep the FoV size small (as discussed above) we choose an effective FoV of 1″, which requires actual FoV sizes of 1.9–2.7″ depending on the band (see Fig. 11.2).

11.4.2.4 Image analysis

The images simulated in the previous section are not processed in any way before we perform the photometry on them. This implies that we assume a perfect basic data reduction process (i.e. bias subtraction, flat-fielding, etc.). Note, however, that this does *not* include sky subtraction.

Given that we wish to perform accurate stellar photometry in crowded fields, and given that the shapes of AO PSFs can be quite complex (see e.g. Fig. 11.10), it is clear that we require PSF photometry as opposed to aperture photometry. Here we choose to use StarFinder (Diolaiti et al. 2000). StarFinder is an IDL code that was developed specifically for the purpose at hand, namely the analysis of crowded stellar fields with well-sampled, non-varying but complex PSFs. Briefly, the code proceeds in the following way: (i) detection of suspected objects by identifying peaks above some threshold; (ii) verification and detailed analysis of the suspected objects by decreasing intensity; (iii) re-fitting of detected objects. The detected objects are then subtracted from the image and the above procedure is repeated in order to detect and measure fainter stars that had previously gone undetected. In fact, the procedure is repeated a total of three times after which usually no more new stars are found.

Step (ii) of the above procedure consists of the following operations: re-detection of the suspected object after the subtraction of already confirmed stars in order to reject spurious detections due to PSF features of brighter stars; cross-correlation of the suspected object with the PSF in order to reject spurious detections that are not PSF-like such as noise spikes; fitting of the PSF to the object in order to determine its position and flux, taking into account the contributions of neighbouring brighter stars and of the local background.

StarFinder is also capable of measuring the PSF from individual suitable stars in the same image. However, this requires significant 'manual' interaction with StarFinder. But for our purposes we need to use StarFinder in a fully automated way. Hence we directly provide StarFinder with a PSF to use in the analysis. The assumption here is that this PSF would be the result of an unspecified PSF calibration procedure. By default we feed StarFinder a PSF that is essentially the same as the one used when creating the image, except that it is of much smaller extent. The underlying assumption in this case is of course that the PSF calibration procedure is essentially perfect. However, in order to gauge the effect of PSF uncertainties we will also investigate cases where we drop this assumption and where the analysis PSF is slightly different from the one used to create the image (see Section 11.5.5).

The result of the StarFinder analysis is a catalogue of detected stars, with x,y positions and fluxes which we convert to magnitudes using the same zero-point that was used in the image simulation. We also correct the magnitudes for the known flux difference between the smaller analysis PSF and the larger PSF used when simulating the image. We refer to this as an 'aperture correction'.

We then match the StarFinder output catalogue to the input catalogue by identifying those pairs of input/output objects where each member of the pair identified the other as the closest object in that catalogue, and where the distance between the two may not exceed 0.5 pixel.

Finally, the process of input catalogue generation, image simulation, image analysis and catalogue matching is repeated until we have enough statistics in order to reliably measure the magnitude errors. These are determined as a function of input magnitude in bins of 0.5 mag from the distribution of differences between the input and aperture corrected output magnitudes (cf. Fig. 11.3).

11.4.3 Inputs

11.4.3.1 Scientific data

Isochrones: Throughout this work we have used the BaSTI isochrones (Cordier et al. 2007; Pietrinferni et al. 2004) in order to create stellar populations and star catalogues with which to populate the simulated images. BaSTI offers isochrones for different heavy element mixtures and chemical compositions. We only use the scaled solar models with three different metallicities: [Fe/H] = -1.8 (metal-poor), -1 (intermediate) and -0.6 (metal-rich). We always assume a Salpeter IMF with slope -2.35 over the entire mass range offered by the BaSTI models ($0.5 \leq M/M_{\odot} \leq 11$). For the SFH our default assumption is

Table 11.1: Physical parameters of the simulated galaxies. All photometric properties were derived from V-band observations. Data from McConnachie et al. (2005), Mateo (1998) (NGC 205); Rejkuba (2004), van den Bergh (1976) (Cen A); GOLDMine database (M87).

	NGC 205 (LG)	Cen A (NGC 5128)	M87
DM	24.58	27.92	31.2
kpc/arcsec	0.00396	0.0186	0.084
kpc/arcmin	0.238	1.116	5.055
Profile	Exponential	de Vaucouleurs	de Vaucouleurs
Scale or effective radius (arcsec)	h = 102	R_e = 330	R_e = 105
Central or effective SB (mag arcsec^{-2})	μ_0 = 20.4	μ_e = 22.15	μ_e = 20.58

a constant star formation rate (SFR) from 14 to 12 Gyr ago, but we also explore the effect of younger populations.

Unfortunately, the BaSTI models do not offer stellar luminosities in the Johnson I-band but only in the shorter wavelength Cousins I_C-band. We were thus forced to use the I_C-band luminosities in all I-band simulations. The main difference between the I_C and I-band simulations is therefore the PSF (although the zero-points and system transmission are also slightly different).

Galaxy data: We will perform simulations for three different galaxies at three different distances: NGC 205 is a nucleated dSph belonging to the M31 system and hence the LG; Centaurus A (Cen A) is the nearest giant elliptical; and M87 is the central elliptical of the Virgo cluster. Table 11.1 lists their distances and photometric profile parameters. The latter will be used to convert surface brightnesses to galactocentric distances.

11.4.3.2 Technical data

The technical data used for the simulations were taken from [RD1]. All magnitudes in this study refer to Vega zero-points, not AB.

Site and background: We follow [RD1]. By default we use the background emission model for the Paranal-like site, but we also explore the effect of using the High & Dry site. However, this will only affect those simulations where the photometric errors are not dominated by crowding, and only in the K-band (due to the thermal background). We consider an airmass of 1.0 for all the simulations.

Telescope: As described in [RD1]. By default we use the protected silver/aluminium coating, but we also explore the effect of using the bare aluminium coating. The choice of coating is likely to significantly affect only the I_C and I-band simulations (due to the dip in transmission of the bare aluminium coating in this wavelength range), and only in those cases that are not too strongly affected by crowding. In addition, the K-band may be affected due to the different thermal properties of the two coatings.

Instrument: We assume a diffraction limited imager operating in the I_C to K bands (i.e. in the range 0.8–2.2 μm). In each band we assume a pixel size that optimally samples the diffraction limited core of the PSF. These are listed in Table 11.2. The instrument's FoV is treated as a free parameter.

An instrumental efficiency was not available at the time this work was carried out. We choose to use a constant value of 0.7, excluding the detector QE for which we use the values from [RD1]. However, these values will only have an impact on those simulations not too strongly affected by crowding. For the detector read-out noise and dark current values given in [RD1] the photometric errors in all simulated observations are completely dominated by either background photon noise or crowding, and hence they can be safely ignored.

Adaptive Optics and PSFs: Given the need for high spatial resolution in this science case we assume LTAO correction. By default we use the on-axis, 0° zenith distance PSFs.

Table 11.2: Some general properties of the PSFs used for the simulations and the analysis.

Band	Pixel size (mas)	$R_{\rm PSF}^{\rm sim}$ [a] (mas)	$R_{\rm PSF}^{\rm sim}$ [a] (pix)	Contrast[b]	$N_{\rm PSF}^{\rm SF}$ [c] (pix)	$R_{\rm PSF}^{\rm SF}$ [d] (mas)	Contrast[e]	Ap. corr.[f]
I_C	1.21	459	380	10^{-3}	300	182	7×10^{-3}	1.87
I	1.45	642	443	10^{-4}	300	218	2×10^{-3}	1.64
J	2.01	865	430	10^{-5}	210	211	4×10^{-4}	1.58
H	2.66	776	292	10^{-5}	124	165	4×10^{-4}	1.52
K	3.54	737	208	10^{-5}	82	145	4×10^{-4}	1.39

[a] Extent of the PSF profile used for simulating images.
[b] Ratio of the PSF profile value at $R_{\rm PSF}^{\rm sim}$ and its central value.
[c] Linear extent of the PSF image given to StarFinder for the analysis of the simulated images.
[d] Radius of the circle enclosed by the PSF image given to StarFinder.
[e] Ratio of the PSF profile value at $R_{\rm PSF}^{\rm SF}$ and its central value.
[f] 'Aperture' correction factor needed to account for the flux difference between the smaller, square PSF image used by StarFinder and the larger, circular PSF used in the image simulation.

We do not use the simulated PSF images of [RD1] themselves but rather their analytical fits, as described in [RD1]. Specifically, when creating the simulated images we use the 1D radial profile of the analytical PSF fit. This results in much shorter execution times of the simulations than when using a 2D PSF image. (It also implies that the edge of the PSF in the simulated image is circular rather than square as is the case when using a PSF image.) We do not just use the core of the PSF but preserve the seeing halo down to a contrast of 10^{-5} (with respect to the PSF's peak). Unfortunately, going down to this contrast level is only possible for the J, H and K bands. For the I_C and I bands this contrast would require us to go to such large radii (in terms of image pixels) as to seriously impact the execution time and to render the simulations impractical. This limitation forced us to use contrast levels of 10^{-3} and 10^{-4} for the I_C and I bands, respectively. Table 11.2 lists the contrast levels and the corresponding radii used for all bands considered.

Unfortunately, StarFinder cannot work with a 1D PSF profile but requires a 2D PSF image. Hence, for the image analysis we used 2D PSF images generated from the analytical PSF fits. However, for computational reasons the extent of these images had to be considerably smaller than that of the profiles used for the image generation, going down to contrasts of only $\sim10^{-2}$ – $\sim10^{-3}$, depending on the band (see Table 11.2). However, we do not consider this a shortcoming because in reality it would anyway be difficult to measure the PSF to contrast levels much better than this.

11.4.4 Outputs

The output of the simulation and analysis pipeline consists of the simulated images (stored as FITS files), corresponding matched input/output catalogues, and estimates of the photometric errors as a function of magnitude. Obviously, the output catalogue can also be used to plot CMDs.

11.5 Results of simulations

11.5.1 Simulation runs

The various simulation runs performed and their parameters are summarised in Table 11.3. These will be discussed in detail in the next section. For each run the simulations were carried out for a range of surface brightness values, covering the galaxy in question from its outskirts all the way to its central region. For runs 5 and later the final column of Table 11.3 lists the parameter whose effect on the simulation result was investigated with this run.

Table 11.3: Simulation runs and their parameters.

Run	Galaxy	Bands	$t_{\rm exp}$	Parameter investigated
1	M87	I_C, I, J, H, K	100 h	
2	Cen A	I_C, I, J, H, K	100 h	
3	NGC 205	I_C, I, J, H, K	100 h	
4	M87	J	100 h	Resolution
5	M87	J	100 h	AO performance
6	M87	I_C, I, J, H, K	10 h	
7	M87	H	100 h	[Fe/H] = -0.6, -1.8
8	M87	I_C, K	10 h	Al coating
9	M87	K	10 h	High & Dry site
10	M87	H	100 h	PSF calibration errors

11.5.2 Analysis

11.5.2.1 Pipeline validation

We begin by verifying our procedures and by illustrating the results of a simple simulation run in Fig. 11.3. In this simulation we have replaced the input catalogue generation procedure described in Sections 11.4.2.1 and 11.4.2.2 with a very simple scheme where the input catalogue contains only a single object drawn from a constant magnitude distribution. Hence there are no crowding effects in these simulations. Our purpose here is simply to verify that for this simple case we obtain sensible results.

In Fig. 11.3 we show the results of this simulation run (in which we have assumed $t_{\rm exp}$ = 10 h) for the I_C and K bands. The black points in the upper panels show the difference between the input and aperture-corrected output magnitude of each simulated star. The red points and errorbars show the median and 68 percentile ranges in bins of 0.5 mag. Note that despite the aperture correction the average magnitude difference is not precisely zero as it should be, in the sense that too little flux is recovered. The offset is constant as a function of magnitude but changes from $\sim$0.01 mag to $\sim$0.03 mag when going from the I_C to the K-band. Despite some investigation we have been unable to identify the source of this offset. However, for the purpose of deriving the (random) photometric errors, this small offset is irrelevant. The calibration of the absolute photometry is a source of systematic error that has to be addressed in real observations but this was not considered in the simulations.

The red points in the bottom panels of Fig. 11.3 show the photometric errors from the top panels. This is the key result of the simulations. From this plot we can read off the limiting magnitude at which we can achieve a desired accuracy. Note that at the bright end the errors do not quite tend towards zero like they should, but instead seem to flatten off towards a finite value (more so in the K than in the I_C-band). This has been traced to a numerical interpolation issue. However, this effect is at least a factor of 10 below the level of accuracy we are interested in, and so we do not consider this problematic. The green lines in the bottom panels of Fig. 11.3 show a model of the errors based on our understanding of how StarFinder performs the photometry. This matches the simulation results extremely well, and we conclude that overall the simulation and analysis pipeline works as expected.

As a benchmark, we find from these simulations that in the absence of any crowding the 10 h limiting magnitudes at which the photometric error is 0.05 mag are given by: I_C = 27.9, I = 29.4, J = 28.6, H = 28.0, and K = 27.6 mag.

As an additional check we compare in Fig. 11.4 the S/N computed by the E-ELT Imaging Exposure Time Calculator (ETC; version 2.14; see [RD1]) with the results of a simulation like the one described above but for $t_{\rm exp}$ = 1 h. In each panel the red, green and blue points show the ETC results for three different magnitudes as a function of the size of the S/N reference area, i.e. the area around the assumed star over which the ETC sums the object and background fluxes to be used in the S/N calculation. The results of our simulations are shown as orange stars. Recall that StarFinder is given a PSF image to perform its analysis. However, it does not in fact use the whole PSF image when fitting the stars in the image, but only the central region of size $\sim$10 pixel. This corresponds roughly to the diameter of the first dark ring of

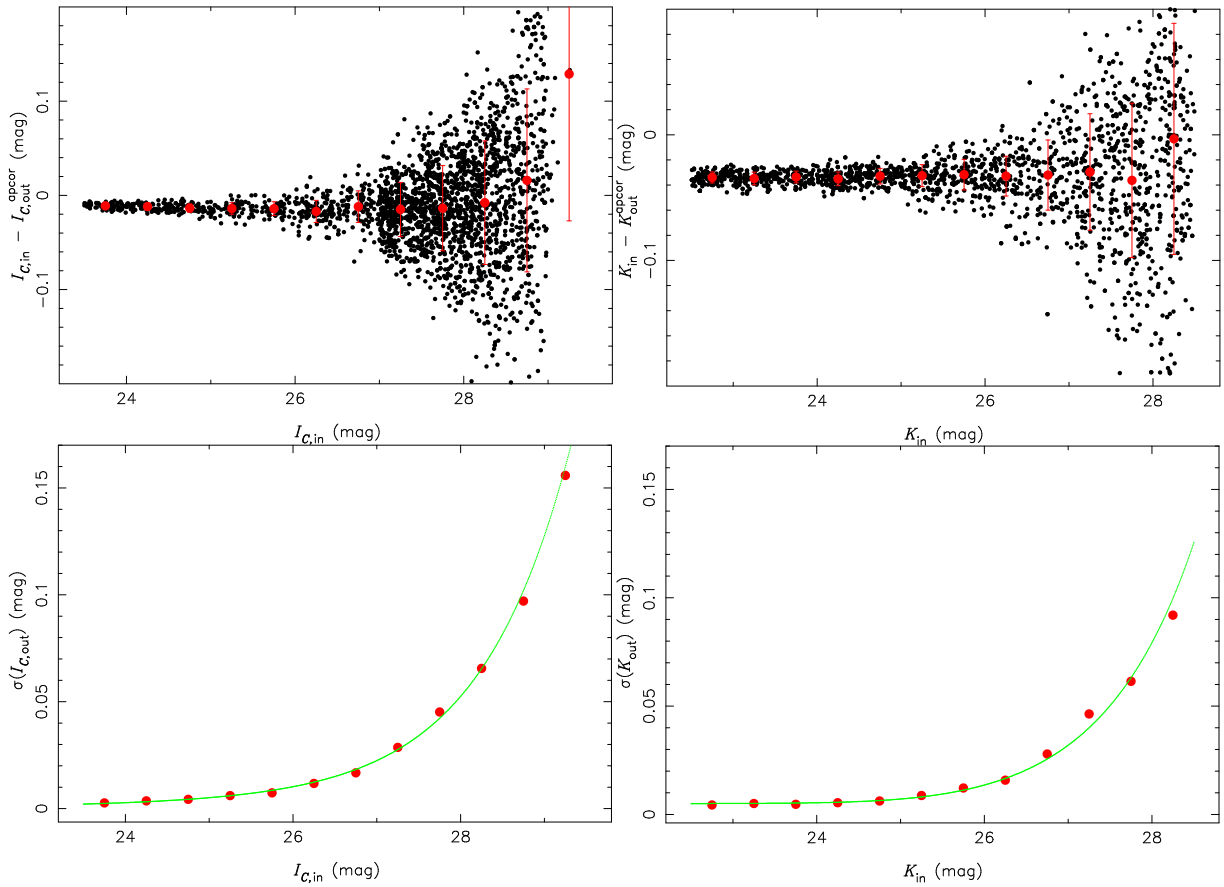

Figure 11.3: Photometric assessment in the I_C (left) and K (right) bands of a simple example simulation run. In this example, the simulated images contained only a single star, and we assumed $t_{\rm exp}$ = 10 h. The black points in the top panels show the difference between the input and aperture-corrected output magnitudes as a function of the input magnitude. The red points and errorbars show the median and 68 percentile ranges (i.e. 1σ errors) in bins of size 0.5 mag. In the bottom panels the red points explicitly show the magnitude errors from the top panels. The green line is a model of the errors based on the knowledge of how StarFinder performs the photometry.

an Airy disk. Hence, we have (somewhat arbitrarily) placed the simulation results along the x-axis at the size of the diffraction limited core at the wavelength in question.

We find that in all cases our simulation results lie very close to the S/N curves outlined by the ETC results. This means that (i) our simulations agree very well with the ETC, and that (ii) the S/N obtained from PSF photometry is similar to that obtained from aperture photometry if the aperture is chosen to be roughly of the same size as the diffraction limited core.

Since the shapes of the S/N curves in Fig. 11.4 depend on the shapes of the PSFs and hence on the AO correction, this figure also provides us with a first assessment of the effectiveness of the AO correction as a function of wavelength. As already alluded to in Section 11.4.1, and as will be discussed in detail in Section 11.5.2.3, this issue is of some importance to the present science case. Here we note that the structure of the S/N curves outlined by the ETC results in Fig. 11.4 in the I_C and I bands is qualitatively different from that in the J, H and K bands. In the former case we see two 'bumps', where the bump at larger sizes is the remnant of the seeing peak: in the absence of any AO correction the S/N should peak near the value of the seeing, as shown by the blue crosses in the I_C-band panel. (Note that the PSF simulations assumed a seeing of 0.78 arcsec at 0.5 μm [RD1], which corresponds to 0.71 arcsec in the I_C-band.) As one increases the fraction of the flux that is re-distributed from the seeing halo of the PSF to its diffraction limited core by the AO correction a second, inner peak develops in the S/N curve. In the I_C-band the AO correction is not effective enough to raise this inner peak above the outer seeing peak. In

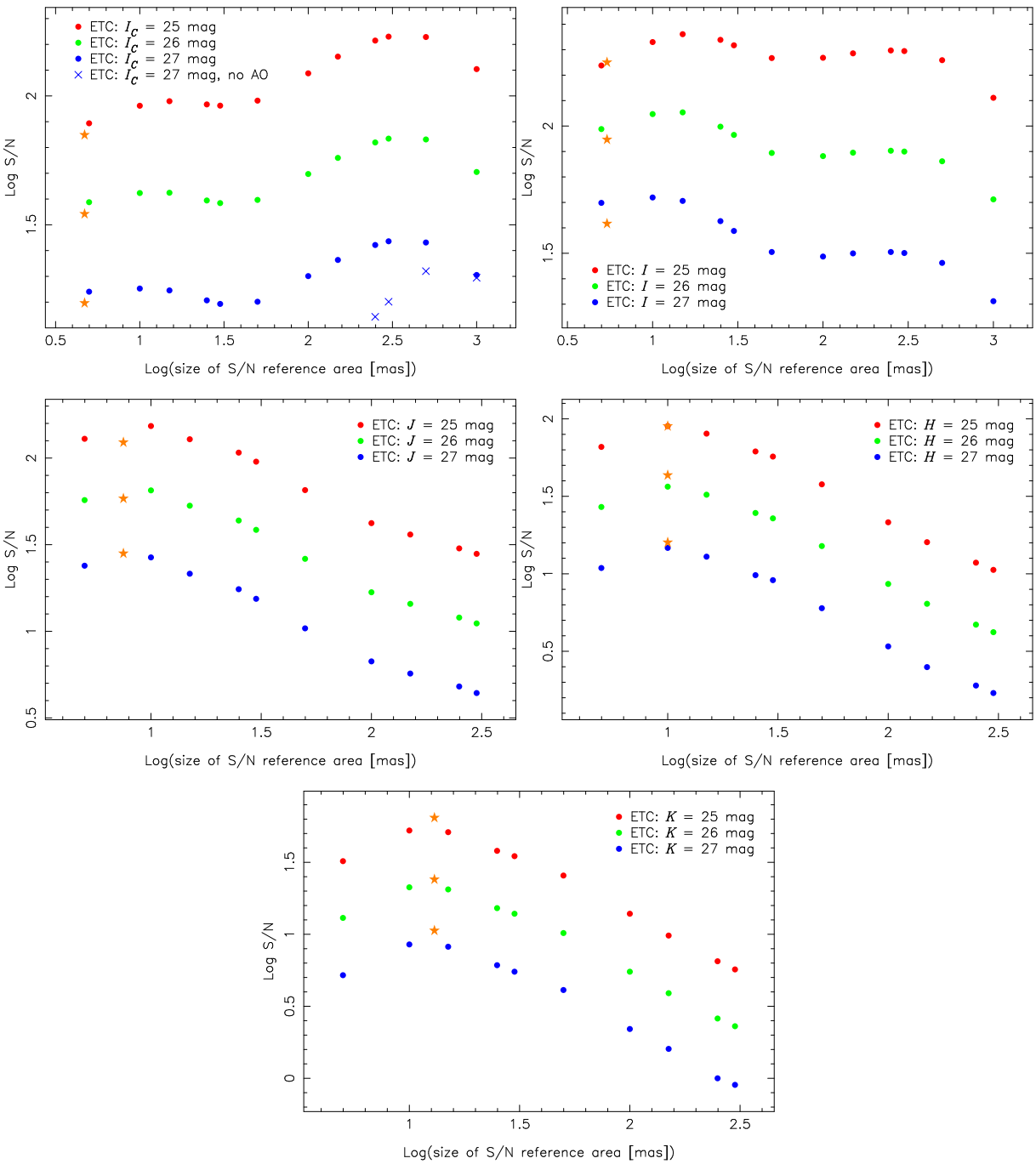

Figure 11.4: Comparison of the output of the E-ELT Imaging ETC with the results of the simulations for the simple case of a single star (i.e. without any crowding). In each panel the red, green and blue points show the ETC's S/N as a function of the size of the reference area used for the S/N calculation, for three different magnitudes as indicated. The different panels show the results for the I_C to K bands. The I_C-band results were derived from the I-band values by correcting the fraction of energy ensquared by the S/N reference area from the I-band PSF values used by the ETC to the I_C-band PSF. The blue crosses in the I_C-band panel are the ETC results for no AO correction (instead of the LTAO correction used in all other cases). We have assumed $t_{\rm exp}$ = 1 h in all cases. The orange stars indicate the results of the simulations for the same assumptions. Along the x-axis these results are placed at the theoretical size of the diffraction limited core of the PSF at the appropriate wavelength.

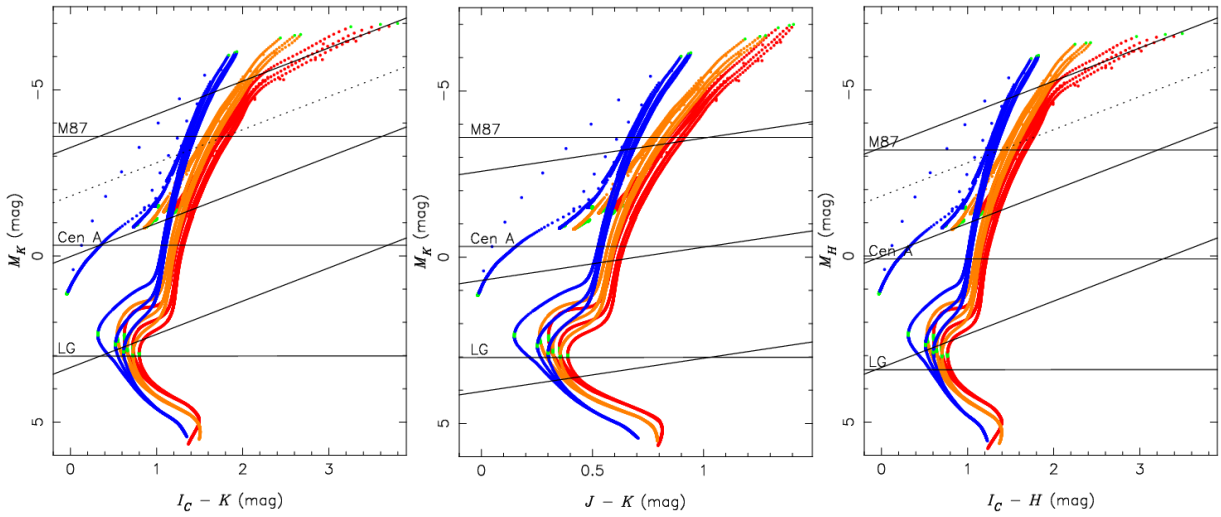

Figure 11.5: Left: Example isochrones from the BaSTI models in the M_K vs. $I_C - K$ plane. We show the scaled solar models with [Fe/H] = -1.8 (blue), -1 (orange) and -0.6 (red). In each case we show isochrones with $t_{\rm age}$ = 5, 9 and 13 Gyr. The TRGB, the faint end of the HB and the MSTO are marked by green points. The horizontal lines show the absolute magnitudes that correspond to the K-band magnitude limit (where σ_K = 0.05 mag) at the distances of M87, Cen A and NGC 205 (LG). The slanted lines are the I_C limits. The dashed slanted line shows the limit for the I-band at the distance of M87. All limits assume $t_{\rm exp}$ = 10 h and no crowding. Middle and right: Same for the M_K vs. $J - K$ and M_H vs. $I_C - H$ planes, respectively.

the I-band the inner peak exceeds the outer, but it is not yet fully developed and the maximum occurs not at the size of the diffraction limited core but at a somewhat larger size. In the J-band the inner AO peak is fully developed, although a small remnant of the seeing peak is still visible. In the H and K bands the seeing peak is completely lost. Thus Fig. 11.4 demonstrates nicely how the AO correction improves with increasing wavelength.

Finally, we note that a priori one might expect PSF photometry to be optimal in the sense that it should deliver the best possible S/N. From Fig. 11.4 we can see that this is indeed the case in the J, H and K bands: the simulation results lie at the peaks of the S/N curves outlined by the ETC results. However, this is clearly not the case in the I_C and I bands. The reason is of course the inability of the AO to provide a fully developed diffraction limited PSF core in these bands (as discussed above, cf. also Fig. 11.10) coupled with the fact that StarFinder only uses the core region of the PSF when fitting it to stars. This procedure is clearly sub-optimal in the case of a low-Strehl PSF. Significantly more accurate photometry could be achieved (for non-crowded stars) if StarFinder used a larger fraction of the PSF when performing the final fit of a star with the PSF, instead of just the core.

11.5.2.2 Initial considerations

Before we launch into the discussion of the full-scale simulations it is instructive to orient ourselves by briefly considering what could be achieved in this science case in the best possible case, i.e. in the absence of any crowding. In Fig. 11.5 we show a number of isochrones of various ages and metallicities in the luminosity-colour plane of three different filter combinations, along with the 10 h magnitude limits from the previous section (i.e. for σ = 0.05 mag and not considering any crowding) at the distances of the three galaxies to be investigated here (see Table 11.1). From Fig. 11.5 we can immediately draw the following conclusions:

- Given the stellar luminosities in the various bands considered here, the I, J and H bands clearly provide the deepest photometry, in that order. The I_C-band suffers from its relatively low-Strehl PSF while the K-band is affected by its relatively high background (due to the thermal component coming in). A priori one would like to use the longest possible wavelength baseline for the construction of

the CMD, i.e. the I_C and K bands, but from the left panel of Fig. 11.5 it is clear that this would not be the optimal choice (at least in the absence of any crowding), in particular because of the relatively poor I_C-band limit. Using the I-band instead would provide a substantial improvement (assuming that the I-band luminosities are similar to the I_C-band luminosities), while the combination of J and K encompasses too short a wavelength range (note the different x-axis scale of the middle panel of Fig. 11.5). However, the best combination is clearly I and H.

- At the distance of the LG the depth of the MSTO can easily be achieved at the required photometric accuracy (σ = 0.05 mag) in less than 10 h, at least in the I, J and H bands, for all isochrones considered here.
- At the distance of Cen A the depth of the HB can easily be achieved at the required photometric accuracy in less than 5 h, at least in the I, J and H bands, for all isochrones considered here with the exception of the youngest and metal poorest. However, the MSTO is essentially out of reach, except perhaps in the I-band. For $t_{\rm exp}$ = 100 h we can reach the brightest MSTO shown in Fig. 11.5 with σ_I = 0.05 mag, while the faintest can be reached with σ_I = 0.1 mag. In the J and H bands only the brightest MSTO can be reached in 100 h, and only with σ = 0.1 mag.
- At the distance of M87 we can easily reach well below the TRGB at the required photometric accuracy in 10 h, at least in the I, J and H bands, for all isochrones considered here. Reaching to the depth of the HB is challenging. In $t_{\rm exp}$ = 100 h the full extent of the HB can be probed in the I, J and H bands for most isochrones (not considering the youngest and metal poorest) with an accuracy of σ = 0.1 mag only.

11.5.2.3 Main results

We now go on to consider the results of the full simulations shown in Figs. 11.6–11.8 for M87, Cen A and NGC 205 (LG), respectively. In these simulations we have assumed a stellar population of intermediate metallicity (i.e. [Fe/H] = -1) and with a constant SFR from 14 to 12 Gyr ago. We also used a very long exposure time of 100 h because we wanted to ensure that the photometric errors in the inner parts of the galaxies would be dominated by the stellar crowding (as opposed to the background) in order to demonstrate what can be achieved in principle.

Each panel in these figures summarises the results for one of the I_C to K bands. The colour image in each panel shows the (logarithm of) the magnitude error in the panel's band as a function of magnitude and V-band surface brightness (bottom axis), or, equivalently, galactocentric distance (top axis). Note that a cut through one of these images at a constant surface brightness produces a curve equivalent to those shown in the bottom panels of Fig. 11.3. The two contours overlaid on the images mark the locations where the magnitude errors are 0.05 and 0.1 mag, respectively. These contours therefore show the respective limiting magnitudes as a function of distance from the galaxy's centre. The grey hashed regions show the magnitude ranges of the TRGB, of the faint end of the HB and of the MSTO for the various isochrones shown in Fig. 11.5, while the dashed lines mark the exact locations of these features for the specific stellar population used in these simulations.

Fig. 11.6 shows our results for the most challenging case investigated here, M87. We have simulated observations at surface brightness values corresponding to the range 0.5–10 R_e. At the largest radii there are no crowding effects and the magnitude limits are determined purely by the background. Our results here agree very well with those derived in Section 11.5.2.1 from single object simulations, as they should. Crowding effects begin to dominate the magnitude errors from 3–4 R_e, depending on the band, and then quickly become quite severe. For the parameters used in this simulation the limiting magnitudes increase by up to $\sim$4 mag when moving from the outskirts of the galaxy to its central parts. Nevertheless, in all bands studied here we can probe to the depth of the TRGB with a photometric accuracy of 0.05 mag to well below 1 R_e. The best performance is obtained in the J-band where we can probe to $\sim$0.6 mag below the TRGB all the way down to 0.5 R_e. However, the best *combination* of bands is probably I and H. While J and K yield similar (or even slightly better) photometric performance, the middle panel of Fig. 11.5 shows that in $J-K$ the different isochrones are much less well separated than in $I-H$ (assuming I_C and I-band luminosities do not differ by much), requiring a considerably higher photometric accuracy in order to be able to distinguish between isochrones of different ages and/or metallicities.

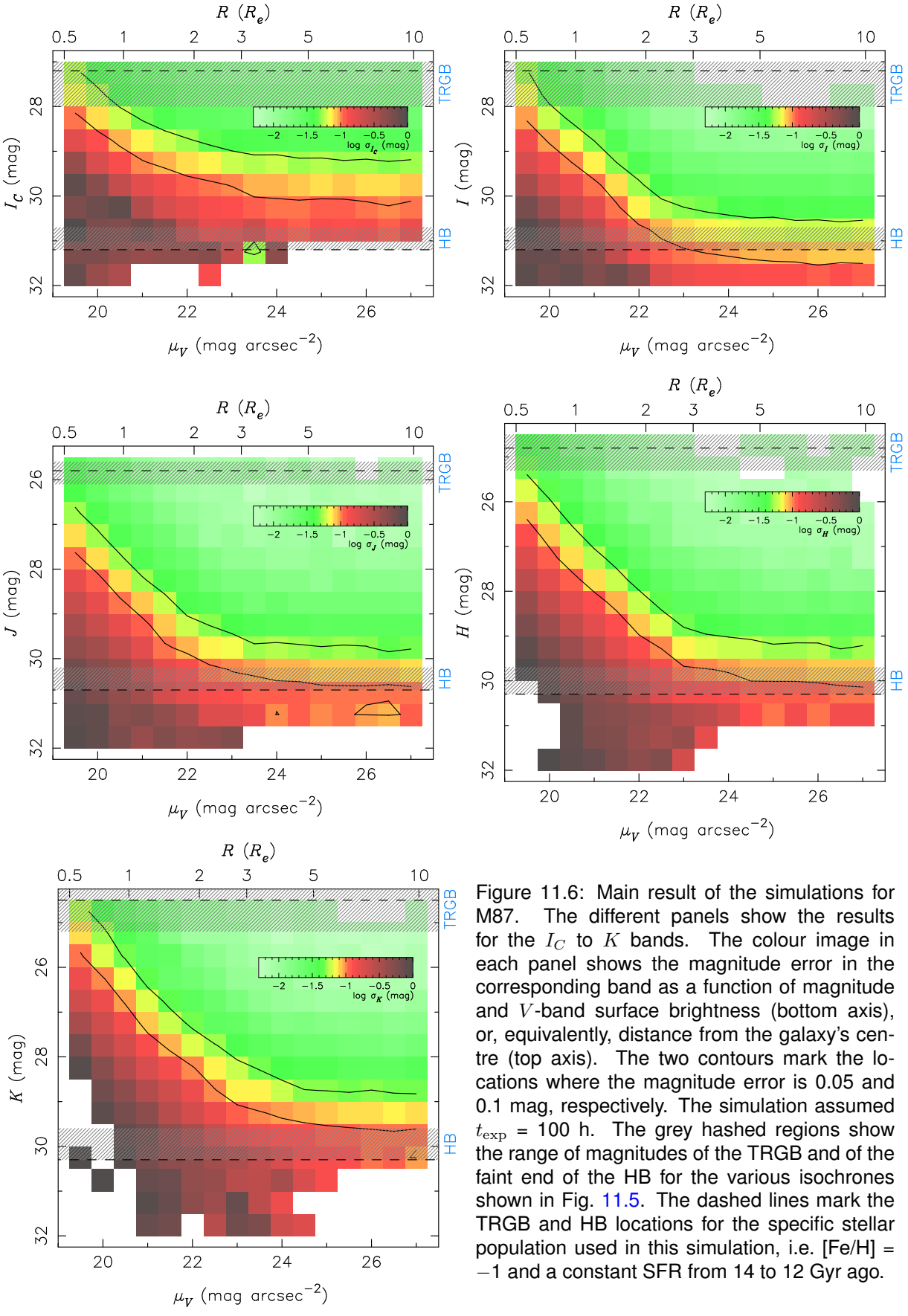

Figure 11.6: Main result of the simulations for M87. The different panels show the results for the I_C to K bands. The colour image in each panel shows the magnitude error in the corresponding band as a function of magnitude and V-band surface brightness (bottom axis), or, equivalently, distance from the galaxy's centre (top axis). The two contours mark the locations where the magnitude error is 0.05 and 0.1 mag, respectively. The simulation assumed $t_{\rm exp}$ = 100 h. The grey hashed regions show the range of magnitudes of the TRGB and of the faint end of the HB for the various isochrones shown in Fig. 11.5. The dashed lines mark the TRGB and HB locations for the specific stellar population used in this simulation, i.e. [Fe/H] = -1 and a constant SFR from 14 to 12 Gyr ago.

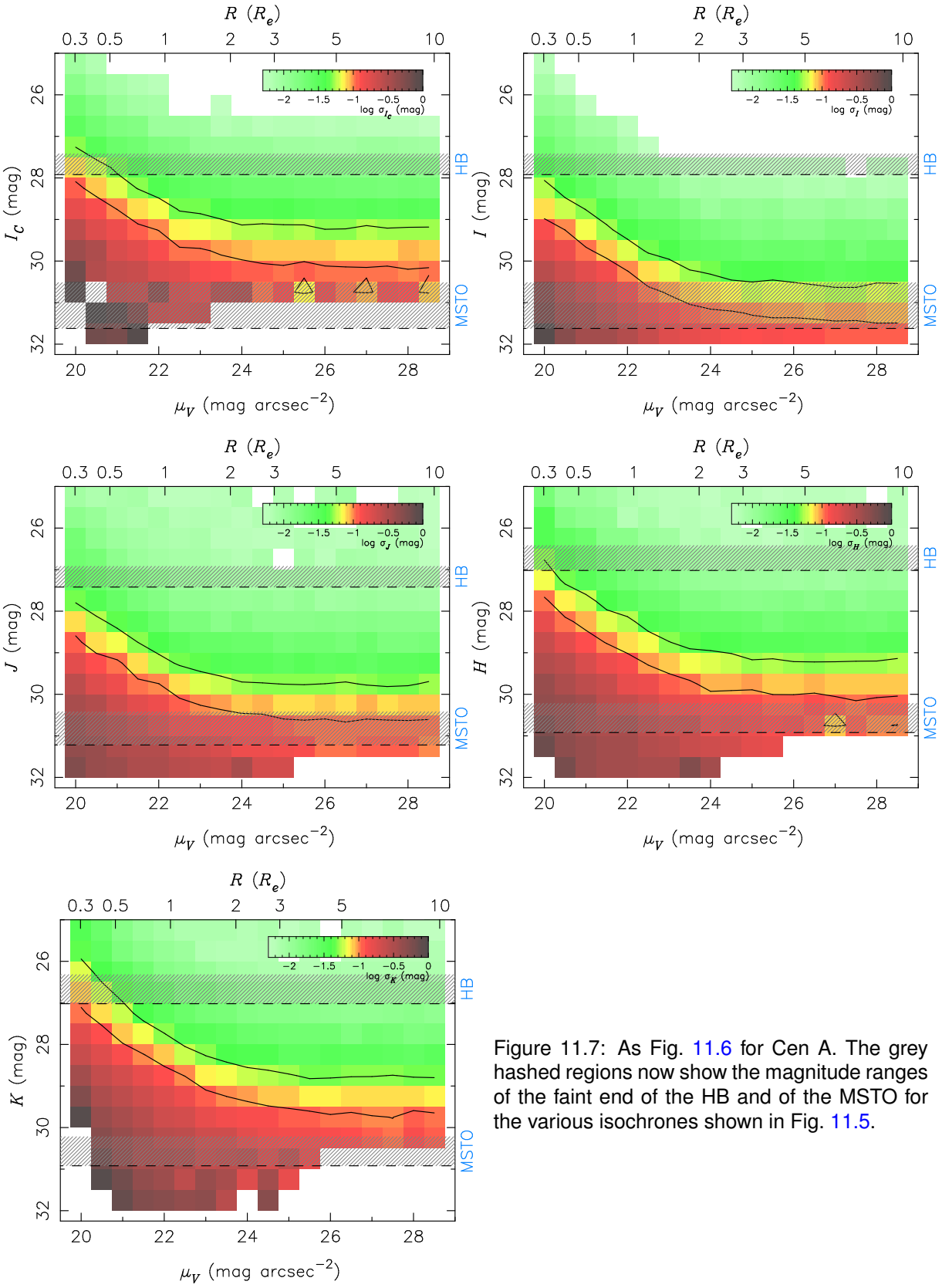

Figure 11.7: As Fig. 11.6 for Cen A. The grey hashed regions now show the magnitude ranges of the faint end of the HB and of the MSTO for the various isochrones shown in Fig. 11.5.

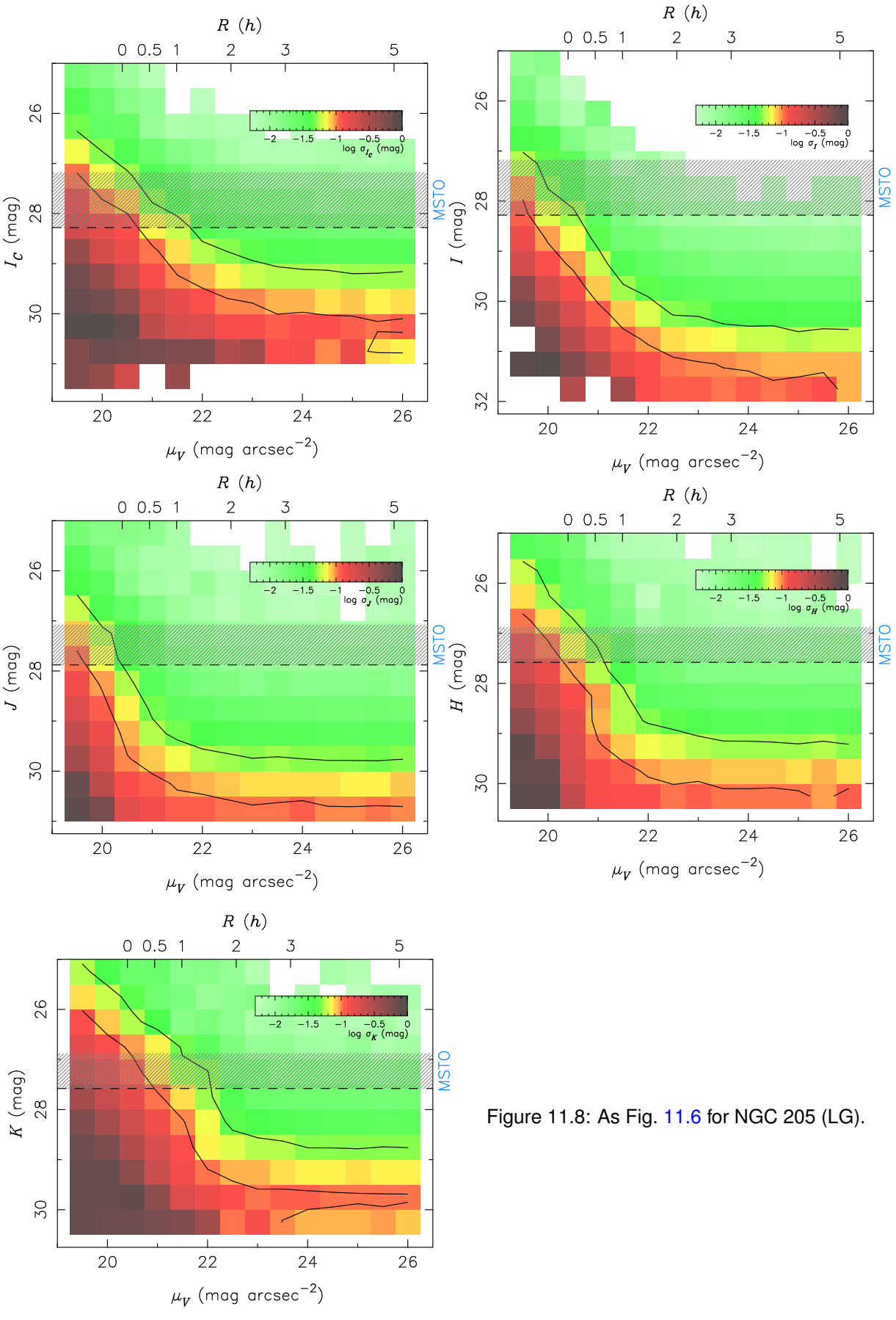

Figure 11.8: As Fig. 11.6 for NGC 205 (LG).

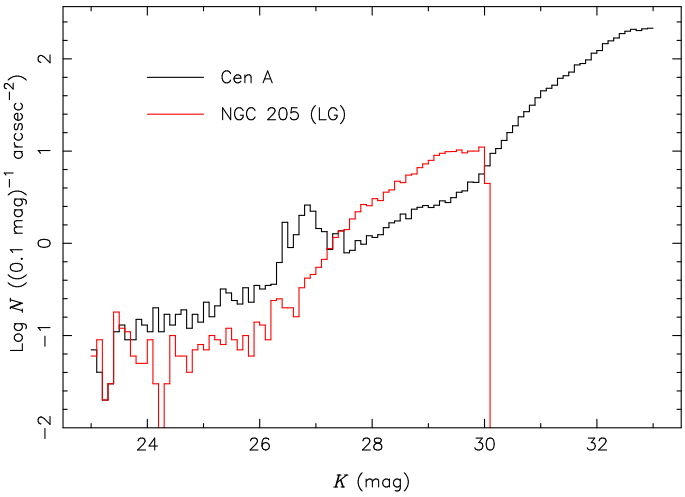

Figure 11.9: Comparison of the number density of stars as a function of K-band magnitude for Cen A (black) and NGC 205 (red) at the same V-band surface brightness of 24 mag arcsec^{-2}. Although NGC 205 is much closer than Cen A and therefore requires a much smaller *total* number of stars to produce a given surface brightness, its *differential* number counts may in fact exceed those of Cen A for a limited magnitude range.

Fig. 11.7 shows our results for the case of Cen A for which we have simulated observations at surface brightness values corresponding to the range 0.3–10 R_e. In this case the crowding effects only start dominating the magnitude errors from $\sim$2 R_e, and the 'contrast' between the magnitude limits in the galaxy's outskirts and its inner regions is not as high as in the case of M87. Both of these effects are of course due to the much lower effective surface brightness of Cen A compared to M87 (cf. Table 11.1). We find that the HB can be probed with 0.05 mag accuracy to well below 0.5 R_e, at least in the I, J and H bands. The best performance is again achieved in the J-band where the HB can be probed to well below 0.3 R_e. However, I and H again appear to be the best combination of bands.

Finally, Fig. 11.8 shows our results for the LG galaxy NGC 205. In this case we have simulated observations at surface brightness values corresponding to the very centre of the galaxy (and indeed at somewhat brighter values), and out to 5 h. For this galaxy the magnitude errors are affected by crowding only in the central region at radii smaller than 1–2 h, depending on the band. However, the limiting magnitudes then rise quite steeply towards the centre (for a reason discussed below), such that only in the J-band (which again delivers the best performance) are we able to probe the MSTO over the entire extent of the galaxy at the required accuracy of 0.05 mag. In the I-band this is possible down to $\sim$0.2 h while in the H-band the limit is $\sim$0.7 h.

We have already noted above that the shape of the limiting magnitude curve as a function of galactocentric radius obviously depends on the surface brightness profile of the galaxy. However, even at a fixed V-band surface brightness the magnitude errors are different for all three cases considered above. Going from M87 to Cen A the limiting magnitude improves at all surface brightness levels (except at the faintest, uncrowded levels where they stay the same, as they should). This is easily explained by the fact that Cen A is much closer to us than M87: the smaller the distance, the greater is the apparent brightness of the stars, and hence the lower is the density of stars that is needed to achieve a given surface brightness value. For a star at a given magnitude this reduces the Poisson fluctuation in the number of fainter (unresolved) stars 'beneath' it, thereby reducing the uncertainty of its photometry. However, when going from Cen A to NGC 205, the limiting magnitude improves for some surface brightness values as above, but it deteriorates for others (see e.g. the J-band limiting magnitudes at μ_V = 21 and 20 mag arcsec^{-2}, respectively). How is it possible that the magnitude error at a fixed magnitude and surface brightness *increases* as we *decrease* the distance? The reason for this somewhat counter-intuitive result is the steep faint-end slope of the stellar luminosity function. As demonstrated in Fig. 11.9, in a certain magnitude range the number density of stars in NGC 205 is actually *higher* than that in Cen A at the same surface brightness, due to the steep slope of the luminosity function. Hence there will be some magnitude range

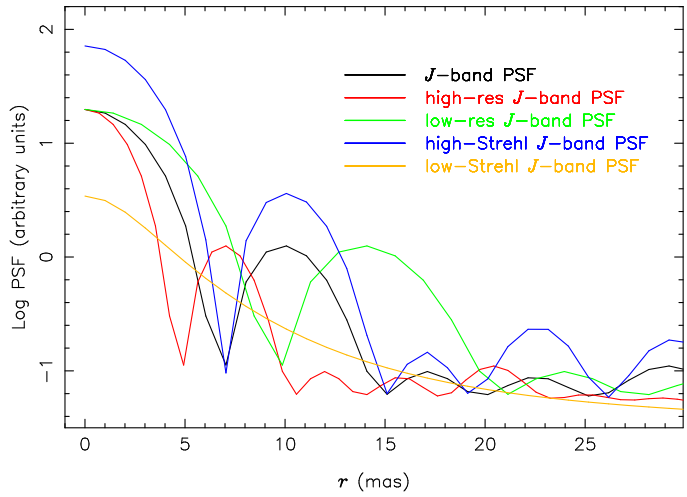

Figure 11.10: The black line shows the inner part of the (logarithmic) profile of the J-band PSF. The red and green lines show the J-band PSF again, but now scaled to the resolution that would be provided by a 60 and 30 m telescope, respectively. The blue line shows the K-band PSF profile scaled to the pixel scale of the J-band PSF, thereby mimicking a high-Strehl J-band PSF. The orange line shows the I-band PSF profile scaled to the pixel scale of the J-band PSF, thereby mimicking a low-Strehl J-band PSF.

in which the photometry of stars in NGC 205 is more badly affected by crowding than in Cen A in the same range, increasing the magnitude errors in this range. The steepness of the luminosity function also explains the steep rise of the limiting magnitude curves in Fig. 11.8 mentioned above.

Another key feature of Figs. 11.6–11.8 is the wavelength dependence of the photometric accuracy. In all cases studied here, the J-band always delivers the best results in the crowded regime, in the sense that in the J-band we can always probe the furthest down the stellar luminosity function, or, equivalently, probe to the highest stellar densities when compared to the other bands (for a fixed magnitude error). The performance gets progressively worse both when going to longer wavelengths (H and K bands) as well as when going to shorter wavelengths (I and I_C bands). This result is independent of the distance of the galaxy, the severity of the crowding, etc., but always holds as long as the photometric errors are dominated by crowding. Clearly, the explanation of this result must lie in the PSFs used for the simulations.

So how does the PSF change as a function of wavelength? Going from longer to shorter wavelengths increases the resolution (because the size of the diffraction limited core scales as λ) while the effectiveness of the AO correction (i.e. the Strehl ratio) decreases (because the inter-actuator spacing of the AO system remains fixed and any uncorrected absolute wavefront error becomes relatively more important as λ decreases). A priori one might expect both higher resolution and a higher Strehl ratio to improve the photometric performance in a crowded stellar field. Hence, when going from longer to shorter wavelengths one must evaluate whether trading AO performance for resolution is favourable or not. Evidently, this trade-off is indeed favourable when going from K to H to J but no longer when going further down to I, such that the J-band offers the best combination of resolution and Strehl.

In the following we will demonstrate the above explicitly. In order to isolate the effect of the PSF on the simulation results – in particular the effects of resolution and Strehl – we have repeated the J-band simulations for M87, but now using four different, arbitrarily constructed PSFs. First, in order to mimic the effect of varying the resolution while keeping the Strehl ratio constant, we have produced high and low-resolution versions of the J-band PSF by simply scaling the pixel size of the original J-band PSF by factors of 42/60 and 42/30, respectively, thereby obtaining PSFs that would be appropriate for a 60 and 30 m telescope, respectively. Secondly, in order to mimic the effect of varying the Strehl ratio while keeping the resolution constant, we have produced high and low-Strehl versions of the J-band PSF by simply rescaling the K and I-band PSFs to the pixel scale of the original J-band PSF. The radial profiles of all four of these newly constructed PSFs are shown in Fig. 11.10 along with the original J-band PSF.

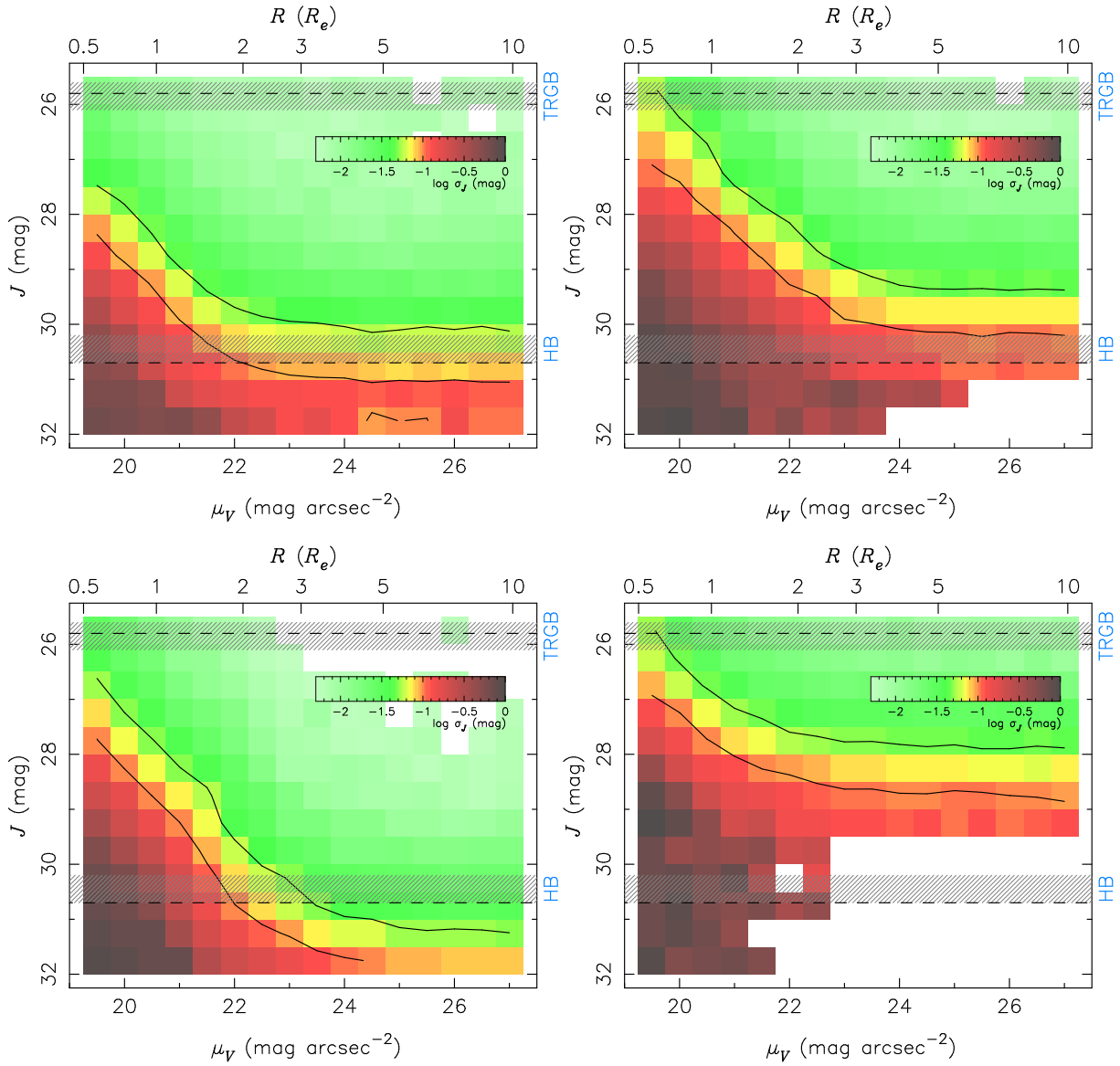

Figure 11.11: As Fig. 11.6 (J-band only) but now using different PSFs. The top panels illustrate the effect of varying the resolution, while the bottom panels show the effect of varying the quality of the AO correction. For the top panels we have used the J-band PSF scaled to the resolution and pixel scale that would be delivered by a 60 m (left) and 30 m (right) telescope (red and green lines in Fig. 11.10), simulating the effect of varying the resolution at constant AO performance. For the bottom panels we have used the K (left) and I-band (right) PSFs scaled to the pixel size of the J-band PSF (blue and orange lines in Fig. 11.10), simulating the effect of increasing and decreasing the effectiveness of the AO correction at constant resolution, respectively.

The results of the simulations using the above PSFs are shown in Fig. 11.11. By comparing the top panels of this figure with the J-band panel of Fig. 11.6 we can see that varying the resolution has a very significant effect on the limiting magnitude in the crowded regime. (There is also some effect on the limiting magnitude in the uncrowded regime but this is of no concern in the context of the present discussion.) The crowding-dominated limiting magnitude changes by almost ± 1 mag when changing the resolution by a factor of $1.4^{\pm 1}$. In contrast, from the lower left panel of Fig. 11.11 we can see that increasing the Strehl ratio has essentially no effect on the crowding-dominated limiting magnitude. However, decreasing the Strehl ratio to the level of the I-band PSF *does* have a significant effect (lower right panel): for $R < 1\,R_e$ the photometric performance using the low-Strehl PSF is very much the same as for the low-resolution

PSF with 'normal' Strehl (upper right panel). The reason for this is apparent from Fig. 11.10: the low-Strehl PSF does not have a well-developed diffraction limited core, and thus its effective resolution is somewhat worse than the resolution of a truly diffraction limited PSF.

The conclusion from the above experiment is therefore that the limiting factor that determines the photometric performance in a crowded stellar field is resolution. The quality of the AO correction only plays a role in the regime where it affects the resolution, i.e. in the regime where the AO correction is good enough to provide an improvement over seeing-limited observations, but not good enough to provide a well-developed diffraction limited PSF core (cf. Fig. 11.4 and its discussion). Once the AO performance is good enough to provide a well-developed diffraction limited core, the resolution is essentially independent of the Strehl ratio, and improving the Strehl further has no significant effect on the photometric performance in the crowding-limited regime.

The above explains why the J-band photometry is superior to that in other bands in crowded stellar fields: the J-band is the shortest wavelength at which the AO performance is good enough to provide a well-developed diffraction limited PSF core. In other words, it is in the J-band where the E-ELT will achieve its highest resolution.

Finally, we consider the effect of exposure time on our results. The exposure time affects the limiting magnitude curves in Figs. 11.6–11.8 by setting the values towards which the curves level off in the uncrowded outskirts of the galaxies.

In general, the total photometric error is always determined by the combination of the photon noise from the background (which depends on the exposure time) and the error due to crowding (which is independent of the exposure time). Therefore, in order to achieve the maximum possible depth at a given surface brightness one needs to choose an exposure time that is long enough to ensure that the background contribution to the total error budget is sub-dominant. On the other hand, integrating any longer than this would be pointless because it would not improve the photometry any further. When surveying a galaxy over a large fraction of its extent one would hence ideally always set the exposure time to its optimal value depending on the surface brightness of the field being observed. However, this will not be possible everywhere because the optimal exposure time increases rapidly as one moves from the central to the outer regions of a galaxy. The question is, where along the profile of the galaxy does it exceed a reasonable value? The answer will depend obviously on the chosen value as well as on the distance and surface brightness profile of the galaxy in question. For the case of M87 and $t_{\rm exp}$ = 100 h we saw above that the transition between background-limited and crowding-limited photometry occurs in the range 3–4 R_e. However, it is not too difficult to imagine what Fig. 11.6 would look like when using a shorter exposure time, say 10 h: the limiting magnitude curves would simply flatten off earlier at the background-dominated (i.e. uncrowded) limiting magnitude corresponding to an exposure time of 10 h.

This is exactly what is seen in Fig. 11.12, where we show the results of repeating our simulations for the M87 case, but now using a more reasonable value of $t_{\rm exp}$ = 10 h instead of the 100 h assumed in Fig. 11.6. The transition between background-limited and crowding-limited photometry now occurs in the range 2–3 R_e. Hence, for reasonable exposure times it will not be possible to probe down to the theoretically possible limit set by the crowding much beyond $\sim$3 R_e, but instead the observations in this regime will be purely limited by the available exposure time. On the other hand, in the central parts of M87 the E-ELT will be able to work at the crowding limit set by the density of stars and the telescope's resolution. In a sense, it is only in this regime where the full power of the E-ELT will be brought to bear.

In contrast, for Cen A and NGC 205 the crowding effects are less severe than in the case of M87, and hence the crowding-dominated limiting magnitudes are considerably fainter, making it harder to reach them. In fact, assuming reasonable exposure times it is likely that the photometry of the stellar populations of both Cen A and NGC 205 will never be completely dominated by the crowding. In this sense, the E-ELT will actually over-resolve these galaxies.

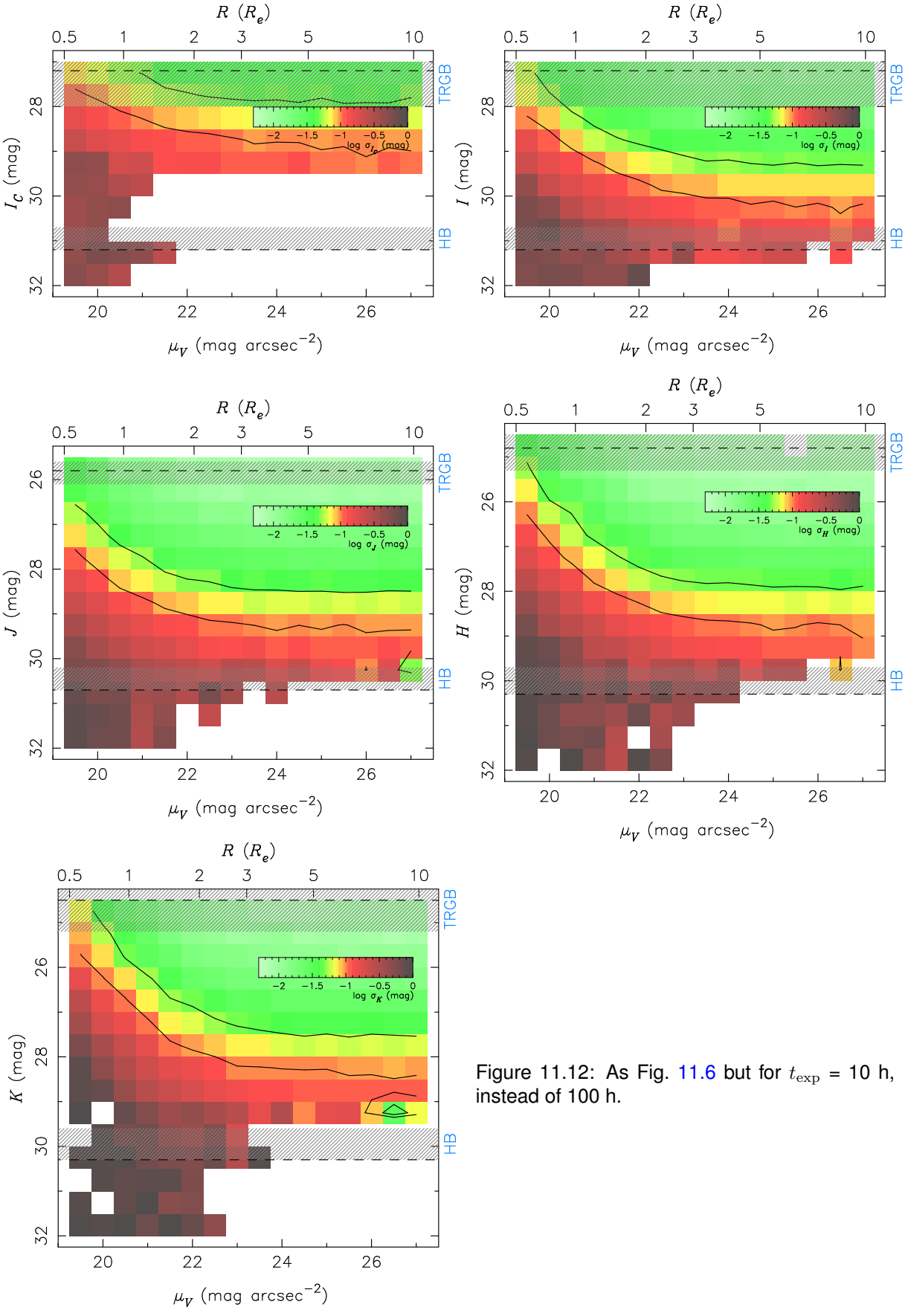

Figure 11.12: As Fig. 11.6 but for $t_{\rm exp}$ = 10 h, instead of 100 h.

11.5.3 Compliance with figures of merit

Fig. 11.12 demonstrates that in the I, J, H and K bands we can probe the stellar population of M87 in the Virgo cluster at 1 R_e with a photometric accuracy of 0.05 mag down to 1–1.5 mag below the TRGB in $\sim$10 h of exposure time. Hence we consider the main goal of this science case met.

Note that the statement above only concerns the quality of the photometry that we can achieve. Translating this into an accuracy of the physical parameters of the stellar population (such as age or metallicity) depends on the number of stars that we can populate the CMD with. The question then arises of how much area of a galaxy must be imaged (at a given surface brightness) in order to obtain a sufficiently large number of stars. The answer can essentially be read off from Fig. 11.9, where we showed the K-band number counts of the stellar population used in the simulations for μ_V = 24 mag arcsec^{-2} and for the distance moduli of Cen A and NGC 205. The number counts can be scaled to any value of μ_V by simply shifting them up (or down) by an amount 0.4 $\Delta\mu_V$, where $\Delta\mu_V$ is the difference between the desired surface brightness and 24 mag arcsec^{-2}. Similarly, the counts can be scaled to any DM by shifting them right (or left) by an amount ΔDM, as well as up (or down) by an amount 0.4 ΔDM. For example, for M87 at μ_V = 20.5 mag arcsec^{-2} ($\approx$ 1 R_e) and K = 26 mag ($\approx$ the limiting magnitude at 1 R_e) we find a density of $\sim$20 stars per arcsec^2 and per 0.1 mag bin. Similarly, at μ_V = 25 mag arcsec^{-2} ($\approx$ 5 R_e) and K = 27.5 mag ($\approx$ the 10 h limiting magnitude at 5 R_e) we find a density of $\sim$1 star per arcsec^2 and per 0.1 mag bin. Thus we conclude that in the case of M87 even a modest FoV of, say, 10 $\times$ 10 arcsec^2 would be sufficient to supply an ample number of stars with a single pointing anywhere along the galaxy's profile out to $\sim$5 R_e.

11.5.4 Sensitivity to input parameters

In Section 11.5.2.3 we have already seen how the limiting magnitude at which we can achieve an accuracy of 0.05 mag depends on magnitude, V-band surface brightness, galaxy distance and surface brightness profile, wavelength, resolution, AO performance, and exposure time. In this section we investigate the dependence of our results on a number of other scientific and technical input parameters.

Stellar population The simulations considered so far all assumed a stellar population of intermediate metallicity ([Fe/H] = -1) and with a constant SFR from 14 to 12 Gyr ago. How do the results of these simulations change when varying the metallicity, the SFH or other stellar population parameters, such as the slope of the IMF?

The crowding-induced error on the photometry depends on the stellar population because it depends on the number counts, i.e. on the surface density of stars as a function of magnitude. In particular, the error at a given magnitude depends most strongly on the number counts immediately faintwards of the magnitude in question. Varying the stellar population changes the number counts differently at different magnitudes and in different bands. The effect of varying the stellar population on the photometric accuracy will thus depend on the band and on which part of the luminosity function is being considered (TRGB, HB or MSTO). Hence, a full investigation of these effects would require us to repeat the full set of simulations for M87, Cen A and NGC 205 for a number of different stellar populations.

Here we will concentrate on our most important case, M87. In this case we are only concerned with the brightest part of the luminosity function (i.e. the TRGB). The main point to appreciate here is that when working at a fixed surface brightness the number counts at the bright end are quite insensitive to any changes in the stellar population. The reason is simply that the brightest stars dominate the surface brightness and therefore fixing the surface brightness is almost equivalent to fixing the bright number counts. Some variation must still be expected because the surface brightness is fixed in the V-band while the relevant number counts are those in the I_C to K bands.

We have checked that our results for M87 are relatively robust against changes in the stellar population by repeating the H-band simulations for M87 with a metal-poor and a metal-rich stellar population. We chose to vary the metallicity rather than the age of the stellar population because from Fig. 11.5 it is clear that the former parameter should have a larger effect on the bright number counts than the latter. Fig. 11.13 shows the results of these simulations. In the metal-poor case the bright number counts are reduced compared

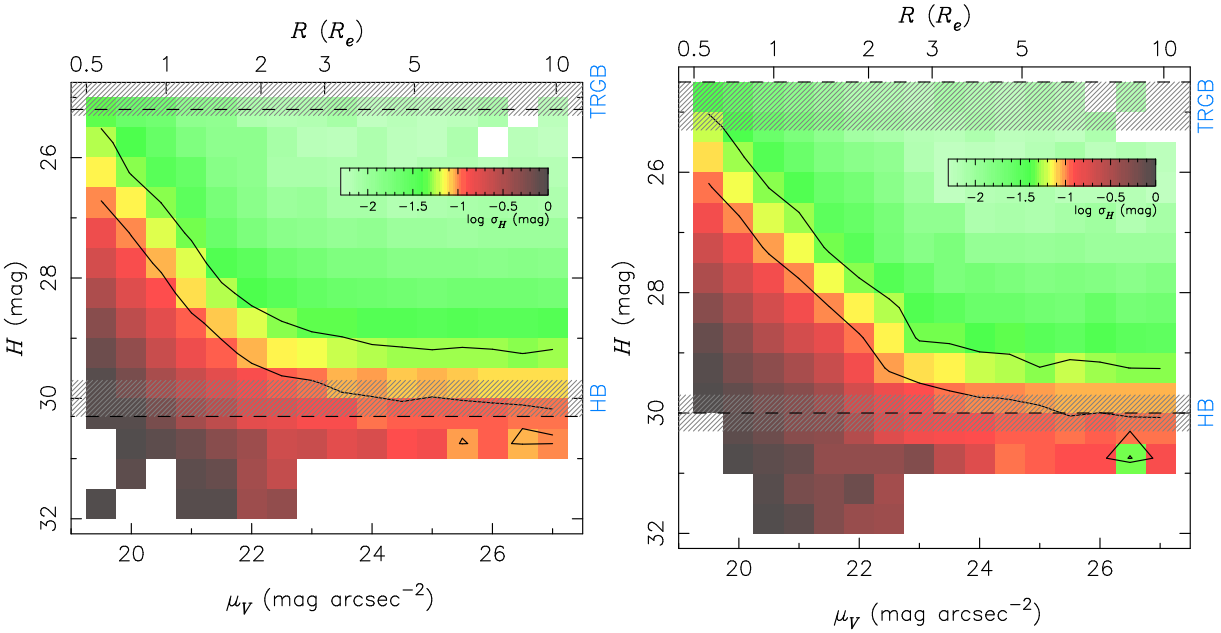

Figure 11.13: As Fig. 11.6 (H-band only) but now assuming a metal-poor ([Fe/H] = -1.8, left) and metal-rich ([Fe/H] = -0.6, right) stellar population instead of an intermediate metallicity.

to the intermediate metallicity case, and indeed, by comparing the left panel of Fig. 11.13 with the H-band panel of Fig. 11.6 we can see that the limiting magnitudes in the central parts of the galaxy have improved by 0.2–0.3 mag. However, the TRGB has also shifted towards fainter magnitudes, resulting in a net loss of depth with respect to the TRGB of $\sim$0.3 mag. In the metal-rich case the situation is essentially reversed, with both the limiting magnitudes and the TRGB shifting towards brighter magnitudes with no net loss or gain in depth. Thus we conclude that our results for M87 are quite robust with respect to any changes in the stellar population.

Telescope diameter The telescope's diameter obviously determines both its resolution (assuming the existence of an effective AO correction) as well as its photon collecting power. The top panels of Fig. 11.11 already illustrated the effect of changing the telescope's resolution on the limiting magnitude curve for the case of M87. Note that the crowding-dominated part of this curve shifted more than the background-limited part when varying the resolution. However, these simulations did not include any changes in the photon collecting power. Adding this will only affect the background-limited part of the curve such that the net result of varying the telescope diameter by a factor of 1.4 is an approximately constant shift of the entire limiting magnitude curve by $\sim$0.8 mag. This result is independent of any other parameter and only assumes that the effectiveness of the AO correction is independent of the telescope's diameter.

Coating Fig. 11.14 shows the result of repeating the simulations for M87 using $t_{\rm exp}$ = 10 h, but now assuming the bare aluminium coating instead of the protected silver/aluminium coating. The aluminium coating has a pronounced dip in its transmission function near the wavelength of the I_C-band (see Fig. 6.1 of [RD1]), and this band will be affected the most by switching the transmission function to the bare aluminium case. In addition, the emissivity of the aluminium coating is higher by a factor of 2.9 compared to the protected silver/aluminium coating, resulting in an increase of the K-band background by 0.4 mag arcsec^{-2}. These issues will affect the I_C and K-band limiting magnitudes in the uncrowded regime. Any magnitude limits that are fully dominated by crowding should not be affected.

Comparing Fig. 11.14 to the corresponding panels of Fig. 11.12 we can indeed see these effects. The limiting magnitude in the uncrowded regime in the I_C-band has become brighter by $\sim$0.4 mag, pushing the whole curve up such that only the very tip of the RGB is accessible with an accuracy of 0.05 mag, and only for $R \gtrsim 2\,R_e$. In the K-band the limiting magnitude in the uncrowded regime has become brighter

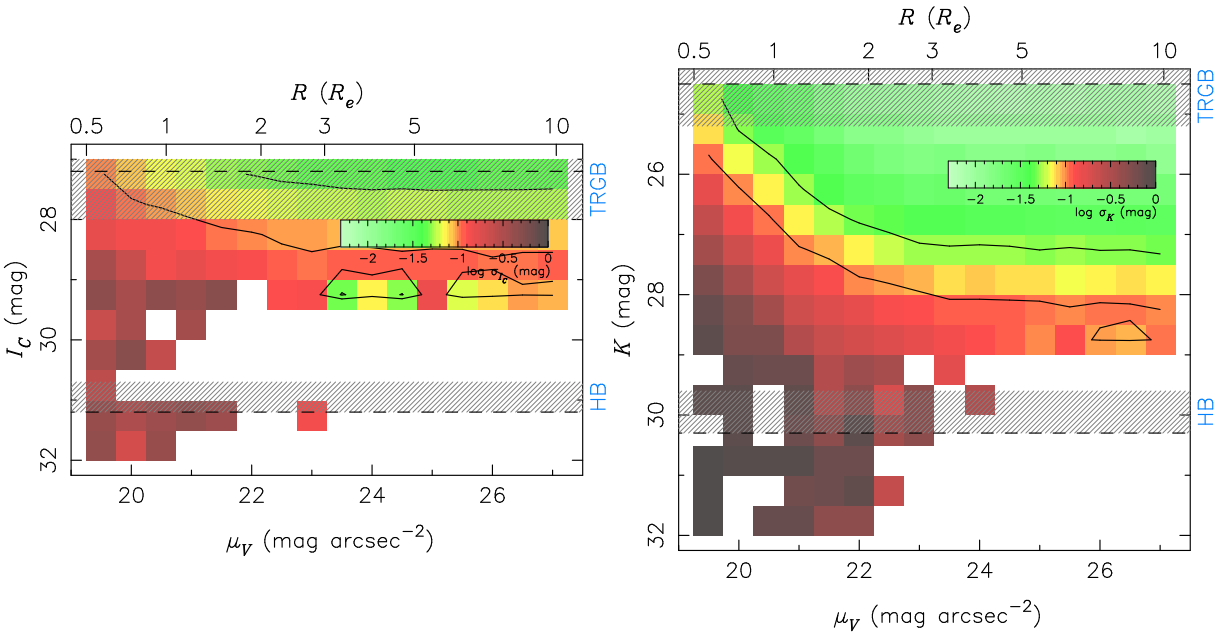

Figure 11.14: As Fig. 11.12 (I_C and K bands only) but now assuming the bare aluminium coating instead of the protected silver/aluminium coating.

by ∼0.25 mag but there are no changes in the inner part of M87 where the magnitude limit is completely dominated by crowding.

As discussed in Section 11.5.2.3, for reasonable exposure times the photometry of the stellar populations of Cen A and NGC 205 will probably never be completely dominated by crowding. Hence, any parameters that affect the background-limited performance (such as the coating) are more important for these cases than for M87.

Site Fig. 11.15 shows the result of repeating the simulations for M87 using $t_{\rm exp}$ = 10 h, but now assuming the High & Dry site instead of the Paranal-like site. The former is significantly cooler than the latter (see Table 4.1 of [RD1]) so that the thermal emission from the atmosphere and the telescope is somewhat lower, decreasing the K-band background by 0.4 mag arcsec^{-2}. This will affect the K-band limiting magnitude in the uncrowded regime. Any magnitude limits that are fully dominated by crowding should not be affected.

Indeed, comparing Fig. 11.15 to the K-band panel of Fig. 11.12 we can see that the limiting magnitude in the uncrowded regime has become fainter by ∼0.25 mag. However, the limiting magnitude curve in the inner part of the galaxy remains unchanged as expected.

As in the case of the coating above we again point out that any changes to the background-limited performance are more important to observations of Cen A and NGC 205 than for M87.

11.5.5 Calibration requirements

The most important calibration aspect for the present science case is the question of how to determine the PSF. In the simulations discussed above we have ignored this issue and simply gave StarFinder essentially the same PSF for the image analysis as was used for the image generation, implicitly assuming a perfect PSF calibration procedure. What happens if we drop this assumption? What is the effect of any differences between the true PSF and the PSF determined by the (unspecified) calibration procedure and subsequently used for the image analysis? Furthermore, for a FoV larger than a few arcsec the PSF is expected to change quite drastically across the FoV, as demonstrated in the left panel of Fig. 11.16. Similar changes are expected as a function of airmass (see Fig. 8.7 of [RD1]). To what accuracy must these changes be tracked in order not to significantly affect our results?

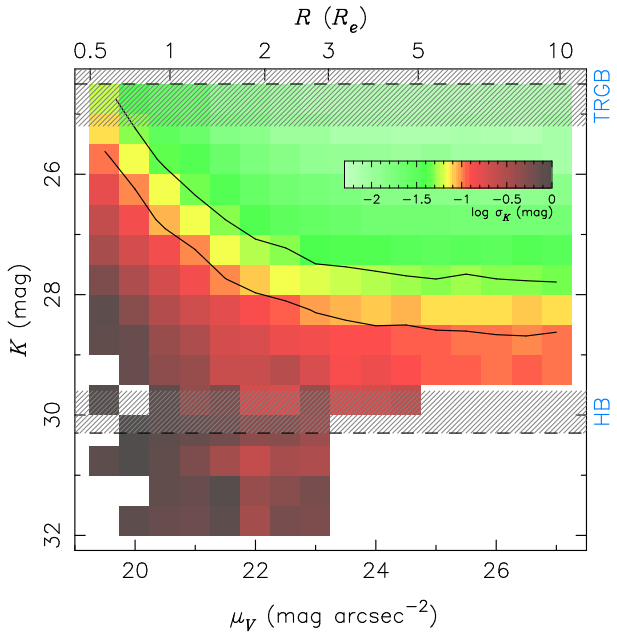

Figure 11.15: As Fig. 11.12 (K-band only) but now assuming the High & Dry instead of the Paranal-like site.

We first consider the simple case where the PSF used for the image analysis is slightly different from the true PSF, and where this difference is the same for all stars. We now gauge the effect of a constant PSF difference by re-analysing the H-band simulations for M87 using two 'perturbed' H-band PSFs. These were obtained from the original H-band PSF (which was used for the generation of the simulated images) by randomly changing the parameters of the Airy component of the analytical PSF fit by $\pm$5 % and $\pm$10 %, respectively. These 'perturbed' PSFs are shown in the right panel of Fig. 11.16 along with the original H-band PSF for comparison.

The results of the re-analyses are shown in Fig. 11.17. By comparing each of the panels of this figure to the H-band panel of Fig. 11.6 we can see that the 5 % PSF perturbation has almost no effect at all on the limiting magnitude curve, while the 10 % PSF perturbation causes a minor increase of the crowding-dominated magnitude limits by $\sim$0.2 mag. Hence a PSF calibration error that is the same for all stars has only very little effect on the *random* photometric errors.

Of course, the difference between the true PSF and the one used for the image analysis also causes a *systematic* offset in the photometry. For the PSF perturbed by 5 % the recovered magnitudes are too faint by $\sim$0.015 mag, while for the PSF perturbed by 10 % the offset is a very large $\sim$0.22 mag. However, such an offset would be unproblematic as long as it were really the same for all stars. In this case the offset would be taken care of by the calibration of the absolute photometry which will always have to be performed anyway if the CMD is to be compared with stellar population models.

The real problem arises if the difference between the true PSF and the one used for the photometry changes, e.g., as a function of position in the FoV, which is likely to be the case in reality. In this case the different offsets in the different parts of the FoV would cause a relative shift in the photometry which could be mis-interpreted as a variation of the stellar population. Although an offset of order 0.01 mag caused by a PSF variation of a few per cent is well below the random photometric error of 0.05 mag targeted in this study, it would nevertheless be easily detectable by averaging over many stars.

The issue of PSF calibration clearly requires further investigation but for now we conclude that any PSF changes (as a function of position within the FoV, airmass, star colour, etc.) will need to be tracked with an accuracy of at least a few per cent in order to be able to take full advantage of the excellent random photometric errors the E-ELT will be able to provide.

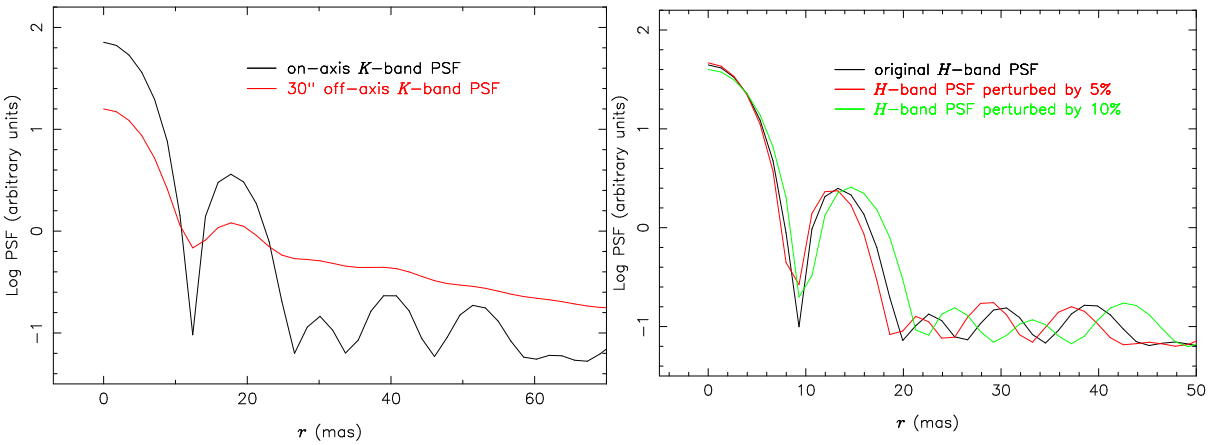

Figure 11.16: Left: Comparison of the on-axis K-band PSF (black) with a 30″ off-axis K-band PSF (red), illustrating the severity of anisoplanatism expected for LTAO. Right: The red and green lines show 'perturbed' versions of the H-band PSF obtained by randomly changing the parameters of the Airy component of the original PSF (shown in black) by $\pm$5 % and $\pm$10 %.

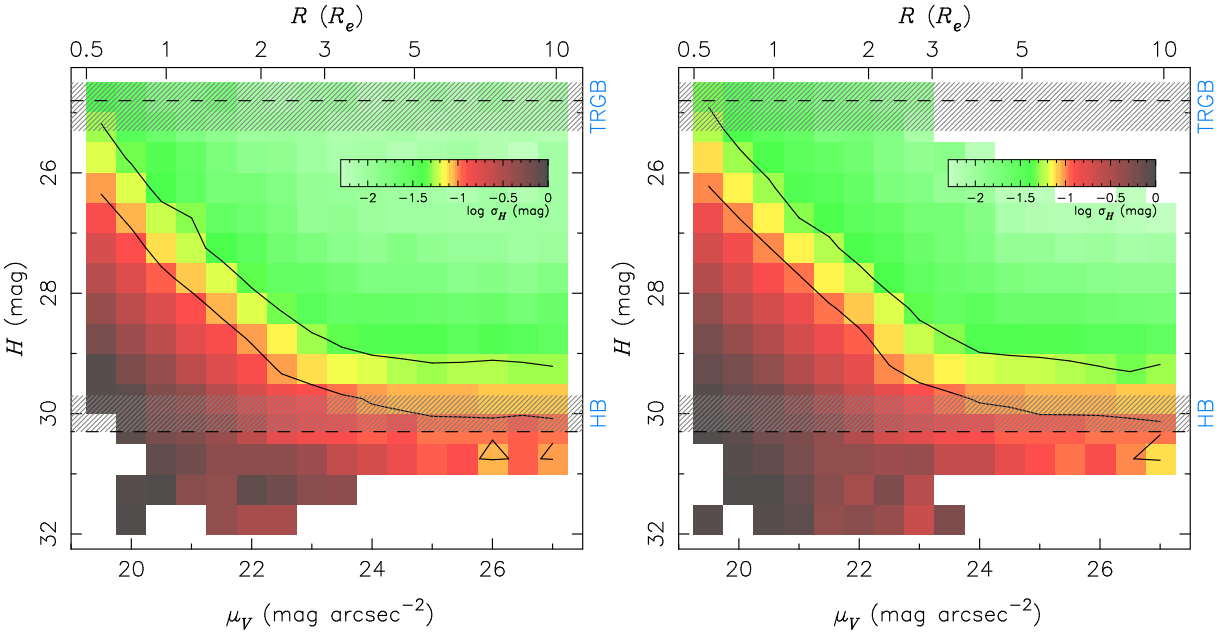

Figure 11.17: As Fig. 11.6 (H-band only) but now using the two 'perturbed' versions of the H-band PSF shown in the right panel of Fig. 11.16 for the image analysis instead of the original H-band PSF that was used for the generation of the images. The left and right panels show the results for the PSFs perturbed by 5 % and 10 %, respectively.

11.5.6 Limitations

There are several 'real-life' issues that we have ignored in the present study. First of all, we have assumed a perfect data reduction process. More precisely, we have assumed that any photometric errors induced by imperfect data reduction are small compared to those induced by the background photon noise and stellar crowding. This is not an unreasonable assumption. Note that accurate background subtraction is probably not an issue for this science case. Images analysed by StarFinder do not need to be background subtracted because StarFinder fits a local background when fitting the stars with the PSF. Evidently, this works very well.

Secondly, we have not considered the possibility of detecting and analysing stars simultaneously in multiple bands (or using the detections from a higher resolution band to analyse the image of a lower resolution band) essentially because StarFinder does not allow for this. It is clear, however, that this technique has the potential to significantly improve the limiting magnitudes derived here.

Finally, we have ignored the correlation of photometric errors across bands. If the estimated magnitude of a star in a given band is in error because of an undetected Poisson fluctuation in the number of fainter stars 'beneath' it, then the same fluctuation will cause a similar error in the estimated magnitude in any other band. This correlation significantly increases the errors on colours.

11.6 Concluding remarks

In this study we have appraised the prospects of using the E-ELT to study the resolved stellar populations of galaxies out to the distance of the Virgo cluster by constructing accurate CMDs. The key issue for the simulations was to quantify how the photometry is affected by stellar crowding as one moves to the dense inner parts of a galaxy. We summarise our results as follows:

1. For NGC 205 in the LG it will be possible to probe the MSTO over almost the entire extent of the galaxy with a photometric accuracy of no less than 0.05 mag.
2. For Cen A it will be possible to probe the HB with an accuracy of 0.05 mag all the way in to the central parts of the galaxy, down to radii of at least 0.3 R_e.
3. For M87 in the Virgo cluster the E-ELT will be able to probe the TRGB with 0.05 mag accuracy all the way in to the very dense central parts of the galaxy, down to radii of $\sim$0.5 R_e. This result is almost independent of the assumed stellar population.
4. Achieving the above results will require exposure times of a few to several tens of hours per observation.
5. The best combination of bands to use for the construction of CMDs is (Johnson) I and H, straddling the wavelength where the E-ELT will achieve its highest resolution (i.e. in the J-band).
6. The accuracy of the photometry in crowded stellar fields is entirely driven by resolution. It is independent of the quality of the AO correction as long as the correction is good enough to provide a reasonably well-developed diffraction limited core in the PSF. As long as this condition is met, the value of the Strehl ratio is immaterial.
7. Given current predictions for the AO performance, the above point will restrict studies of the resolved stellar populations of galaxies in the Virgo cluster with the E-ELT to wavelengths $\gtrsim$ 0.9 μm, below which the diffraction-limited performance is completely lost. Lowering this limit, or at least improving the AO performance in the transition region (i.e. in the I-band), is essentially the only option we have (short of increasing the telescope's diameter) if we want to improve on the results obtained in this study.
8. Accurate PSF calibration will be required. Any changes of the PSF, e.g. as a function of position within the FoV, will need to be tracked with an accuracy of at least a few per cent in order to be able to take full advantage of the excellent random photometric errors the E-ELT will be able to provide.

References

Cordier D., Pietrinferni A., Cassisi S., Salaris M., 2007, AJ, 133, 468
Diolaiti E., Bendinelli O., Bonaccini D., Close L., Currie D., Parmeggiani G., 2000, A&AS, 147, 335
McConnachie A.W., Irwin M.J., Ferguson A.M.N., Ibata R.A., et al., 2005, MNRAS, 356, 979
Mateo M. L., 1998, ARA&A, 36, 435
Olsen K.A.G., Blum R.D., Rigaut F., 2003, AJ, 126, 452
Pietrinferni A., Cassisi S., Salaris M., Castelli F., 2004, ApJ, 612, 168
Rejkuba M., 2004, A&A, 413, 903
van den Bergh S., 1976, ApJ, 208, 673

12 G4-2: The chemo-dynamical structure of galaxies

Authors: G. Battaglia, E. Tolstoy

12.1 The science case

The detailed spectroscopic study of individual stars in nearby galaxies has allowed us to determine accurate metallicity distributions and kinematic properties for stars of a range of age in a range of different nearby galaxies. One common method which has been well studied and calibrated uses intermediate resolution spectroscopy of the NIR CaT metallicity indicator (CaT). It provides accurate line-of-sight velocities ($\pm$1–2 km s^{-1}) and metallicities ([Fe/H], $\pm$0.1–0.2 dex) from the equivalent widths (EW) of CaT lines (e.g. Rutledge et al. 1997; Battaglia et al. 2008a; Starkenburg et al. 2010). This method is applied to red giant branch (RGB) stars, which have ages from > 1 Gyr to the age of the Universe, and are thus tracers of all but the most recent star formation processes in a galaxy. At the present time to even get beyond the outer halo of the Milky Way and make these measurements has required long exposures with sensitive multi-object spectrographs on the largest telescopes (e.g., FORS/VLT, FLAMES/VLT, Deimos/Keck, MIKE/Magellan). Hence the studies to date have been restricted exclusively to the nearest *dwarf* galaxies. Large numbers of careful measurements have led to a deeper understanding of the chemical evolution and also the dynamical state of these systems. The results extend from the discovery of their being dark matter dominated systems (Aaronson 1983; Lin & Faber 1983), to a detailed determination of their complex kinematics (e.g. Kleyna et al. 2003; Battaglia et al. 2006; Walker et al. 2006), and most recently using metallicities to separate out different kinematic components (Tolstoy et al. 2004; Battaglia et al. 2006, 2008b). Some of these careful studies have led to tentative detections of rotation in several nearby dwarf galaxies (Lewis et al. 2007; Battaglia et al. 2008b; Fraternali et al. 2009). We have also learnt a lot about the chemical evolution processes that dominate the evolution of these small systems (e.g. Lanfranchi & Matteucci 2007; Marcolini et al. 2008; Revaz et al. 2009; Tolstoy, Hill & Tosi 2009).

Naturally the next step in this field is to extend these studies, using the same well established and calibrated methods, to a broader range of galaxy types. This requires us to look beyond the Milky Way halo and even the Local Group. With a large telescope we can look in more detail at M 31 and its satellites, which affords a different environment within the Local Group. We can also hope to look beyond the Local Group, to the Sculptor group, for example, which contains a range of different galaxy types (from large spiral galaxies to small gas rich and gas poor dwarf galaxies) and allows us to study a different group environment in detail. Finally the ultimate goal is to be able to make a study of the resolved stellar populations in giant elliptical galaxies, of which there is no example in the Local Group, and we have to look at Cen A (at 4 Mpc) to find the closest example, of a peculiar elliptical. The best place to look at the properties of a range of elliptical galaxy types is the Virgo cluster which contains thousands of large galaxies and tens of giant ellipticals of a range of size and position in the cluster. Here we explore the possibility to carry out these studies with the E-ELT.

We explore the feasibility of multi-object spectroscopic surveys of the resolved stellar populations in a range of nearby galaxies out to the distance of Virgo, sampling entire galaxies, from inner to outer regions. This will allow us to probe the total mass of these galaxies and also to accurately determine the chemical evolution history of all components (i.e., disk, bulge and halo), including a kinematic deconvolution of these components and a comparison of their detailed properties over a range of galaxy type and environment. We will also obtain a more accurate picture of the dark matter properties of a range of different galaxy types, the effects of tidal perturbation and the ubiquity of metallicity distribution functions.

There are other metallicity indicators (e.g., Mgb lines; spectral synthesis as in Kirby et al. 2008) but we choose to concentrate on the CaT because it is red enough to be within the current specifications of the operation of the E-ELT in diffraction limited mode, and it is also at the peak of the flux distribution of an RGB star. It is also a sufficiently well used and calibrated indicator.

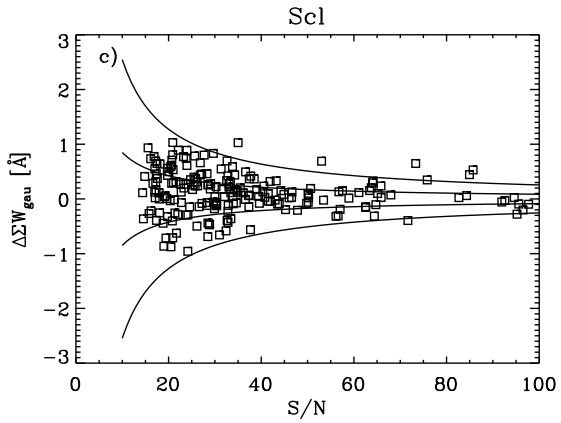

Figure 12.1: Comparison between summed EW (EW_2+EW_3) measurements of CaT lines for stars with double measurements as a function of S/N in the Sculptor dSph (figure from Battaglia et al. 2008a). The solid lines indicate the 1 and 3 σ region for an error in summed EW given by $\sigma_{\Sigma W} = 6/(S/N)$.

12.2 Goals of the DRM simulations

The primary goal of these simulations is to explore whether we can perform accurate intermediate resolution spectroscopy (R$\sim$6000) of large numbers of individual RGB stars in the region of the NIR CaT ($\sim$8600 Å) lines for galaxies out to the distance of the Virgo cluster using an E-ELT with a 42 m diameter.

The basic requirement we have defined for our goal to be achieved is to obtain accurate line-of-sight velocities (±5 km s^{-1} for dwarf galaxy types and ±20 km s^{-1} for large galaxies) and [Fe/H] ($\pm$0.3 dex) measurements from CaT lines for approximately 1000 RGB stars in each galaxy, spanning a large area and for a reasonable total exposure time. The definition of "reasonable" exposure time on source (i.e., including multiple pointings) is $<$100 hrs, which is somewhat subjective, and is defined to be the maximum time usually allocated for a normal program at the VLT.

The feasibility of achieving our goals depends mainly on the signal-to-noise (S/N) we can achieve in the final spectra which drives the errors in the determination of line-of-sight velocities and EWs of the target RGB stars (for an example from VLT/FLAMES observations of individual stars see Fig. 12.1). However it also depends on the stellar background from other resolved or unresolved stars in the galaxy. This will vary between different regions in different types of galaxies at different distances and it can be extremely high in crowded regions, especially in the central regions of dense stellar systems.

The generation of spectrographs under study for the E-ELT will almost certainly consist solely of (mono or multi) integral-field-units (IFUs). According to the physical size encompassed by the spatial resolution element (spaxel), in some objects, and at certain projected radii within the object, only the target star or a few more will fall in the spaxel, while in others there will be a considerable number of background stars which will "pollute" the spectrum of the target RGB star.

Some of the E-ELT instruments under study may offer high and low-spatial resolution modes, which respectively have a smaller field-of-view and spaxel size, or a larger field-of-view and spaxel size. It could be argued that the contribution of the stellar background could be reduced by decreasing the spaxel size and therefore preferring the high resolution spatial modes. However, other aspects also need to be considered, such as the smaller the spaxel size the larger will be the loss of flux due to point-spread-function (PSF) effects; the smaller the spaxel size the smaller will be the field-of-view of the instrument which could make the observations of large number of stars more time consuming. In these simulations we therefore address these trade-offs.

The questions that we wish to answer can be summarised as follows:

- Up to which distance are the science goals of this proposal achievable?
- At fixed distances, how does the galaxy surface brightness influence the results?

- Will the choice of site have a significant influence?
- Will the choice of the mirror coatings have a significant influence?

12.3 Metrics / figures of merit

For establishing the success of the simulations we use the following figures of merit: a) the accuracies in the line-of-sight (l.o.s.) velocity and [Fe/H] measurements from CaT lines which we can reach for a given exposure time per star; b) the total observing time on source needed to acquire these observations for a sufficient number of stars per galaxy (i.e., how many separate pointings are needed).

The required accuracies depend on the specific science goal (e.g. mass determination, general metallicity properties, discovery/analysis of kinematic and metallicity substructures, etc.). Here we will assume that the accuracies to be achieved should be similar to what is currently obtained in studies of the Milky Way and its satellites using similar techniques with current technology. We do not attempt a scientific analysis of our simulations but simply determine the required accuracies from our current experience.

For CaT metallicities we assume that the measurements should be accurate to $\pm$0.3 dex. Here we use the relation [Fe/H]$= -2.81 + 0.44[\Sigma W + 0.64(V - V_{\rm HB})]$ (Battaglia et al. 2008a), where V is the magnitude of the target star, $V_{\rm HB}$ is the V-band magnitude of the horizontal branch of the galaxy and ΣW is the sum of the equivalent width for the 2 strongest CaT lines (at 8542 and 8662 Å). Making the (reasonable) assumption that the main source of error is the determination of the CaT EW, the uncertainty in [Fe/H] will be 0.44 $\times$ the uncertainty in ΣW, therefore an error of 0.3 dex will translate roughly into an error of 0.7 Å in ΣW, which is the figure of merit we will use for the estimation of an accurate measurement of the CaT EW.

For the l.o.s. velocity we require an accuracy of 5 km s^{-1} for dwarf galaxies and 20 km s^{-1} for larger galaxies, such as spirals and ellipticals. This kind of accuracy is similar to what is achieved in current studies of the internal kinematic properties of galaxies in the Local Group from their resolved stellar populations. The limit in the accuracy required for dwarf types is more stringent than for larger galaxies, because the former are small systems, with internal velocity dispersions of a few dozen km s^{-1} and therefore the individual velocities need to be known within a few km s^{-1} so that the derivation of the internal kinematics is not not dominated by measurement errors.

In practice, the figure of merit for each experiment is given by the spread in the distribution of deviations from the input values. Thus, if $v_{\rm in}$ is the "true" (input) l.o.s. velocity of the target RGB star and $v_{\rm sim}$ is the l.o.s. velocity recovered from the simulated spectrum (at the desired resolution of R=6000, including the noise, atmosphere and the effect of the background stars in the spaxel), then we want to study the distribution of $\delta v_i = v_{\rm i,sim} - v_{\rm in}$, where i accounts for each random realisation of the spectrum in the considered spaxel performed for varying stellar background and noise.

The systematics due to the fact that we are observing an RGB star on top of a stellar background are quantified by the average value of δv_i, $\langle \delta v_i \rangle$ (this will tell us how different the simulated values are from the real one on average). The error will be given by the scatter in the distribution of δv_i, which we calculated as the median absolute deviation (m.a.d.). When discussing the achieved accuracy we take into account both systematic and non-systematic effects.

A similar approach is used when estimating the accuracy in [Fe/H].

12.4 DRM simulations

12.4.1 Methodology

The goal of the simulations is to understand how well we can recover the velocity and [Fe/H] of a target star when taking into account all the various effects involved in the simulations, as described in Section 12.2. For this we need to first simulate the corresponding spectrum and then extract from it the velocity and

[Fe/H] and compare these to the input values. Since we are interested only in the CaT lines, we consider only the wavelength range between 8000 and 9000 Å.

The methodology of the simulations can be outlined as follows (for the details of each step see Section 12.4.2 and for the scientific and technical data used see Sections 12.4.3.1 and 12.4.3.2, respectively):

1. Assuming as instrument an IFU, we simulate the spectrum between 8000 and 9000 Å in the central spaxel of the IFU (hereafter, integrated spectrum). In this spectrum we factor in the contributions from: the target RGB star (which we assume is located at the centre of the spaxel); other stars which possibly fall in the same spaxel; stars outside of the spaxel whose light is spread inside of the considered spaxel because of the wings of the PSF. The stars other than the target which contribute to the integrated spectrum will be referred to as "stellar background". The simulated spectrum includes the effects of the atmosphere, telescope, instrument, etc. For more details see Sections 12.4.2 and 12.4.3.2.

2. We derive the l.o.s. velocity, $v_{\rm i,sim}$, and the summed CaT EW, $\Sigma W_{\rm i,sim}$ (as defined in Section 12.3), from the simulated integrated spectrum. For comparison, we also derive the l.o.s. velocity and ΣW of the target RGB star (i.e. without stellar background), after including all the technical effects. This gives an idea of how the results are influenced by the stellar background.

3. We repeat points 1. and 2. 300 times, allowing for different random realisations of the stellar background and the noise.

4. Finally we study the distribution of $\delta v_i = v_{\rm i,sim} - v_{\rm in}$ and $\delta\Sigma W_i = \Sigma W_{\rm i,sim} - \Sigma W_{\rm in}$, and see how the scatter in the distributions compares to the accuracies we require.

Step 1 is divided into the following sub-steps:

a. We fix the parameters of the simulated object: galaxy type (dwarf or giant), distance modulus, surface brightness profile, internal kinematics.

b. We then associate with this object a stellar population, by choosing a star formation history, a metallicity, an α-abundance, and an initial mass function.

c. Having fixed the above parameters, a stellar catalogue is created, which contains the stars of appropriate characteristics as to produce the required stellar mix. This catalogue is created for an area much larger than the spaxel size.

d. Random coordinates, drawn from a uniform spatial distribution, are assigned to each star, while the target RGB star is placed at the centre.

e. Given the assumed type of adaptive optics, airmass, seeing and therefore corresponding PSF, we calculate which stars contribute to the light in the central spaxel and with which percentage of their original flux, on the basis of their distance from the central spaxel.

f. A spectrum, taken from a synthetic spectral library, is associated with each of the stars contributing to the light in the spaxel. These spectra are then summed up together, the technical effects included and this finally creates the integrated spectrum (for more details see Sections 12.4.2 and 12.4.3.2).

These simulations are performed for elliptical galaxies, both dwarfs and giants, and we assume that they are exclusively pressure supported, i.e. with no net rotation. The data are simulated for galaxies covering a range of distance, from the Local Group to the Virgo cluster; also, several projected radii, and therefore surface brightnesses, are explored within each system, in order to assess the effect of crowding (i.e. the different amount of stellar background) on the recovered properties of the target RGB stars.

The decision to superimpose a target RGB star on an underlying stellar population as done in step 1., rather than randomly selecting stars from a stellar population, is done in order to better understand the effect of the underlying, varying, stellar background on a specific, unchanging, target. This is equivalent to assuming that the location and magnitude of the target RGB stars are known from previous E-ELT imaging, as it is done in all slit and fibre spectroscopy today, where the spectroscopic targets are extracted from photometry in order to optimise the number of targets to be observed. In the specific case

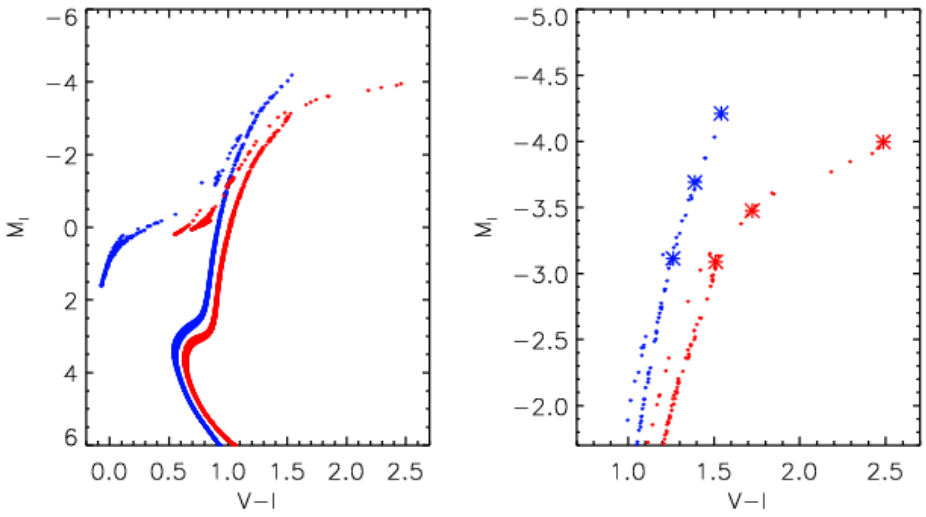

Figure 12.2: Example of colour-magnitude diagram for an ancient stellar population of constant star formation rate from 14 to 12 Gyr ago (blue points: MP population; red points: MR population). The right panel zooms in on the region of the RGB. The asterisks in the right panel show the location on the CMD of the target RGB stars for the two populations. They are located at the tip of the RGB, 0.5 mag below the tip and 1 mag below (the corresponding parameters are listed in Table 12.2).

of E-ELT observations with an IFU, a priori knowledge of the characteristic of the stellar population in various regions of the target galaxy would allow the IFU to be positioned in such a way as to minimise the contamination from the stellar background and to maximise the number of individual spectra of RGB stars. In Section 12.5.3 we calculate how many target RGB stars would be likely available in the simulated objects, given their surface brightness, distance and stellar population, and we discuss how many pointings and how much total exposure time are required to meet the science goals.

12.4.2 Pipeline

The pipeline used to carry out these simulations is divided into 3 main steps: 1. creation of a stellar population; 2. creation of a catalogue of stars for the stellar population; 3. creation of the integrated spectrum corresponding to a bright RGB in addition to the mix of (generally fainter) stars in the catalogue. For steps 1. and 2. we use the same codes used to simulate images of resolved galaxies, developed by J. Liske and E. Tolstoy. For the details of these codes see Section 11. Here we summarise the salient features of the codes used for steps 1. and 2. and the extra steps involved to adapt them to the aim of this work, and we describe the procedures we developed to create the integrated spectrum.

12.4.2.1 The stellar population

The stellar population is created using the BaSTI stellar evolution tracks[1] (e.g. Pietrinferni et al. 2004; Cordier et al. 2007). The input parameters are: the [α/Fe] abundance; the chemical composition, mainly identified with the iron abundance, [Fe/H]; the star formation history; the IMF. The values used for the simulations are summarised in Table 12.1. For simplicity we assume single metallicity and α-abundance. Note that we explore a "metal poor" (MP) case, [Fe/H]$= -1.8$ and a "metal rich" (MR) case, [Fe/H]$= -1.0$. These values for the [Fe/H] abundance were chosen so as to be well within the range of metallicities where the empirical calibration between CaT EW and [Fe/H] abundance is known to be valid.

A weight is assigned to each of the points in the stellar population model, according to the IMF and star formation history, and this gives the probability of finding a star in a certain evolutionary stage.

[1] http://193.204.1.62/

Table 12.1: Parameters of the stellar population used for the simulations.

	Stellar population parameters
SFH	constant
SFH period	14–12 Gyr ago
$[\alpha/\mathrm{Fe}]$	solar
$[\mathrm{Fe/H}]$	-1.0 (MR), -1.8 (MP)
IMF	Salpeter

Table 12.2: Characteristics of the target RGB stars used for the simulations. The last column lists the probability to find a star in the magnitude range (in the I-band) from the tip of the RGB down to 0.1 mag below the magnitude of the target.

Case	$M/M_\odot$	$\log(L/L_\odot)$	$\log T_e$	$\log g$	M_V	$(V-I)$	M_I	Prob
MR-tip	0.88	3.413	3.546	-0.018	-1.508	2.487	-3.995	1.27e$-$05
MP-tip	0.81	3.353	3.607	0.270	-2.670	1.542	-4.212	1.49e$-$05
MR-0.5 mag below	0.86	3.109	3.574	0.458	-1.752	1.723	-3.475	8.49e$-$05
MP-0.5 mag below	0.82	3.128	3.623	0.611	-2.302	1.389	-3.691	8.46e$-$05
MR-1 mag below	0.88	2.925	3.594	0.750	-1.579	1.510	-3.089	0.000149
MP-1 mag below	0.79	2.883	3.639	0.922	-1.852	1.262	-3.114	0.000203

Table 12.3: Physical parameters of the simulated galaxies. When the light profile is de Vaucouleurs, then R_e is the effective radius and Σ_e is the surface brightness at the effective radius, while when the profile is exponential, R_e stands for scale radius and Σ_e for the central surface brightness. All the values referring to photometric properties are from V-band observations. The corresponding references are listed in Sections 12.5.2.1, 12.5.2.2 and 12.5.2.3, except for the ellipticity values which were taken from the GOLDMine database, NED and Choi, Guhathakurta & Johnston (2002) for M87, Cen A and NGC 205, respectively.

	Virgo (M87)	Cen A (NGC 5128)	LG (NGC 205)
DM	31.2	27.92	24.58
kpc/arcsec	0.084	0.0186	0.00396
kpc/arcmin	5.055	1.116	0.238
Profile	de Vaucouleurs	de Vaucouleurs	Exponential
R_e	104.53″	330″	102″
Σ_e [mag arcsec^{-2}]	20.58	22.15	20.4
ellipticity	0	0.23	0.43
dispersion	100 km s^{-1}	100 km s^{-1}	35 km s^{-1}

This stellar population is then used to create a catalogue of stars. An example for the resulting colour-magnitude diagram of a simulated stellar population for an ancient elliptical galaxy can be seen in Fig. 12.2.

12.4.2.2 The stellar catalogue

We want to create a catalogue which includes a bright RGB star (our target) plus the underlying stellar background. The characteristics of the target RGB stars are summarised in Table 12.2 (see also right panel in Fig. 12.2). The background stellar mix is produced for a galaxy at a given distance modulus and at several projected distances from the centre of the galaxy (corresponding to different surface brightnesses) using the stellar population file described above. The number of stars in the catalogue is normalised using the distance modulus, surface brightness in V-band and area considered.

A catalogue contains a list of stars with available information on their mass, luminosity, effective temperature, gravity, magnitudes, colours and l.o.s. velocity. Random coordinates extracted from a uniform distribution are assigned to each star; here we make the assumption that the surface brightness of the galaxy can be considered constant over the simulated area. The target star is placed at the centre.

We explore galaxies located at the outskirts of the Local Group (NGC 205), at the distance of the Centaurus group (NGC 5128, hereafter also Cen A) and in the Virgo galaxy cluster (M87) assuming they all have the same underlying stellar population (in terms of age and chemical composition). The physical characteristics of the galaxies we simulate are summarised in Table 12.3.

For the creation of the background stellar mix and to take into account the possible contamination from stars outside the spaxel because of PSF effects, we create a catalogue containing the stars which, from an area $5'' \times 5''$, contribute to the central 50 mas $\times$ 50 mas spaxel with more than 5×10^{-3} of their original flux and adjust their magnitude according to the flux effectively contributing to the final spectrum. Three random realisations of the stellar background are simulated for each case. Here we are using the LTAO PSF.

To include the broadening due to the intrinsic velocity dispersion of the galaxy under consideration, a l.o.s. velocity is randomly assigned to each of the stars in the catalogue, extracted from a Gaussian distribution centred at zero velocity and with dispersion of 100 km s^{-1} for M87 and Cen A and 35 km s^{-1} for NGC 205.

These simulations do not include the effect of extinction on the magnitude of the targets. This is because the galaxies under consideration are general examples to study the effects of distance and surface brightness and not with the intent of making accurate predictions for these galaxies in particular. But it will be an issue which will have to be considered carefully in a real galaxy.

12.4.2.3 The integrated spectrum

We now need to create the integrated spectrum of the background stellar population in a spaxel, and of the target star, and include the "observational" effects.

This is done in the following steps:

a. We fix the set of technical parameters (telescope diameter, obstruction, spaxel size, telescope and instrument transmission, mirror coating, site, airmass, resolution) and the exposure time.

b. We associate with each star in the stellar catalogue a synthetic spectrum with the closest matching values of gravity, effective temperature, [M/H] and [α/Fe]. The synthetic spectra are from the Munari et al. (2005) synthetic spectral library at R=20 000 (for more details see Section 12.4.3.1).

c. Each of these spectra is red/blue-shifted according to the l.o.s. velocity of the star with which it is associated and the flux is corrected to reflect the apparent magnitude of the star.

d. The spectra are summed on a wavelength grid which covers the region from 8000–9000 Å, the broader wavelength range which includes the NIR CaT lines.

e. The effects of the atmospheric, telescope and instrument transmission are factored in and the background emission is added (see Section 12.4.3.2 for details).

f. The spectra are convolved with a Gaussian to the desired resolution, sampled with 2 pixels per spectral resolution element, and photon and read-out noise are added (see Section 12.4.3.2 for details). To assess the effect of noise, we produce 100 Montecarlo realisations of the noise for each of the stellar catalogues. The sky spectrum is also separately produced, to be then used for the sky-subtraction.

g. The spectra are sky-subtracted. Here we assume a perfect data reduction in the sense that the spectra only contain random noise but no systematics (e.g. from the flat-fielding, sky subtraction, wavelength calibration, etc.). The procedure to calculate the stellar continuum in the spectrum and extract velocities and equivalent widths is described in Battaglia et al. (2008a). Thus a continuum, S/N, l.o.s. velocity v_i and CaT summed EW ΣW_i, are derived for each of the simulated spectra. The EWs of the CaT lines are found by integrating the flux of the CaT lines over 15 Å around the wavelength corresponding to the l.o.s. velocity of the spectrum, which is previously determined. This integration range has been tested in resolved stellar population studies in Milky Way satellites. However, in these studies, the EWs used for the analysis were from Gaussian fits to the line profiles, corrected for a factor which accounted for the non-Gaussian wings of the lines, especially at high metallicities. Here, for simplicity, we adopt the EW from the flux integration. This approach is appropriate for most of the analysis, but causes some underestimation of the EW in the case of NGC 205 (for discussion on this effect see Section 12.5.2.1).

h. Each measured v_i and ΣW_i are then compared to the input values for the target RGB star. The distribution of $v_i - v_{\rm in}$ and $\Sigma W_i - \Sigma W_{\rm in}$ are used to quantify the performance of the simulations for the different galaxies and runs. More specifically the average value of the distribution quantifies the systematics, while the scatter (e.g. the median absolute deviation) quantifies the error.

The approach described above is slightly modified for the case of M87 and for the central regions of Cen A (at R = 0.5 and 1 R_e) in order to reduce the computing time, given the much larger number of stars involved in the simulations. Instead of searching the closest matching spectrum for each of the stars, we consider all the faint main sequence stars with $M_V \geq +4.5$ to be represented by the same template spectrum with $T_{\rm eff} = 5000$ K and $\log g = 4.5$; we calculate how many main sequence stars are present in the simulated area in magnitude bins of 0.2 mag down to $M_V = +8$ mag and normalise the template spectrum to account for the magnitude of the main sequence stars present; the template is also broadened according to the input velocity dispersion of the galaxy.

12.4.3 Inputs

12.4.3.1 Scientific data

Objects: See Table 12.3 and Section 12.4.2.

Synthetic spectra: The synthetic spectra used in this work are from the Munari et al. (2005) spectral library. We used their fluxed spectra at the highest available resolution, i.e. R=20 000. For the grid of parameters available in the library see Table 2 in Munari et al. (2005). Here we note that several rotational velocities and micro-turbulent velocities are available in the models. However for these simulations, we use the models with micro-turbulent velocities = 2 km s^{-1} and rotational velocities = 2 km s^{-1} for stars with $T_{\rm eff} \leq 6000$ K, and 10 km s^{-1} for $T_{\rm eff} > 6000$ K. The latter choices are based on Table B2 from Gray (2005).

12.4.3.2 Technical data

The technical data used for the simulations were all taken from [RD1].

Site and background: We follow [RD1]. We use the background emission and atmospheric transmission models for both the Paranal-like and High & Dry sites. However, in the CaT region (8400–8700 Å) the site is likely to have a negligible effect given that for the Paranal-like site the transmission is 0.98 and for the High & Dry site this is almost 1. We consider an airmass of 1.0 for all the simulations.

Table 12.4: Runs performed for the NGC 205 case with the corresponding parameters listed. Columns: (1) run; (2) object; (3) metallicity assumed for the stellar population (see Table 12.1); (4) site; (5) exposure time on source; (6) mirror coating; (7) explored projected distance from the centre in units of R_e; (8) magnitude of the target RGB star.

Run	Galaxy	Metallicity	Site	Exp. time	Coating	R/R_e	Target
NGC205_1	NGC205	MR	Paranal	2 h	Bare	0,1,2,3,4,5	tip
NGC205_2	NGC205	MP	Paranal	2 h	Bare	0,1,2,3,4,5	tip
NGC205_3	NGC205	MR	Paranal	1 h	Bare	0,1,2,3,4,5	tip
NGC205_4	NGC205	MP	Paranal	1 h	Bare	0,1,2,3,4,5	tip
NGC205_5	NGC205	MR	Paranal	30 m	Bare	0,1,2,3,4,5	tip
NGC205_6	NGC205	MP	Paranal	30 m	Bare	0,1,2,3,4,5	tip
NGC205_7	NGC205	MR	Paranal	20 m	Bare	0,1,2,3,4,5	tip
NGC205_8	NGC205	MP	Paranal	20 m	Bare	0,1,2,3,4,5	tip
NGC205_9	NGC205	MR	Paranal	2 h	Bare	0,1,2,3,4,5	0.5 below tip
NGC205_10	NGC205	MP	Paranal	2 h	Bare	0,1,2,3,4,5	0.5 below tip
NGC205_11	NGC205	MR	Paranal	1 h	Bare	0,1,2,3,4,5	0.5 below tip
NGC205_12	NGC205	MP	Paranal	1 h	Bare	0,1,2,3,4,5	0.5 below tip
NGC205_13	NGC205	MR	Paranal	30 m	Bare	0,1,2,3,4,5	0.5 below tip
NGC205_14	NGC205	MP	Paranal	30 m	Bare	0,1,2,3,4,5	0.5 below tip
NGC205_15	NGC205	MR	Paranal	20 m	Bare	0,1,2,3,4,5	0.5 below tip
NGC205_16	NGC205	MP	Paranal	20 m	Bare	0,1,2,3,4,5	0.5 below tip
NGC205_17	NGC205	MR	Paranal	2 h	Bare	0,1,2,3,4,5	1.0 below tip
NGC205_18	NGC205	MP	Paranal	2 h	Bare	0,1,2,3,4,5	1.0 below tip
NGC205_19	NGC205	MR	Paranal	1 h	Bare	0,1,2,3,4,5	1.0 below tip
NGC205_20	NGC205	MP	Paranal	1 h	Bare	0,1,2,3,4,5	1.0 below tip
NGC205_21	NGC205	MR	Paranal	30 m	Bare	0,1,2,3,4,5	1.0 below tip
NGC205_22	NGC205	MP	Paranal	30 m	Bare	0,1,2,3,4,5	1.0 below tip
NGC205_23	NGC205	MR	Paranal	20 m	Bare	0,1,2,3,4,5	1.0 below tip
NGC205_24	NGC205	MP	Paranal	20 m	Bare	0,1,2,3,4,5	1.0 below tip

Telescope: As described in [RD1]. We take into consideration both of the mirror coatings offered by [RD1] to explore the effect of the coating on the final result. The 'standard' bare aluminium (hereafter, bare Al) coating has a considerable dip between 7000 and 9000 Å, which covers the region of the NIR CaT lines, whereas the coating with the protected combination of silver and aluminium (hereafter, Ag/Al) has a more regular and higher reflectivity over the wavelength range of interest.

Instrument: The present science case requires a large number of resolved stars to be observed and so we choose as baseline an instrument having a relatively large field of view. In [RD1] two options were available: a single-field IFU, with a $10'' \times 10''$ field-of-view and 50 mas/spaxel scale; and a multi-IFU design, with a circular patrol field of $5'$ diameter and individual IFUs with a $1.3'' \times 1.3''$ f.o.v. and 50–75 mas/pix scale. Hence, we fixed the spaxel dimension in our simulations to 50 mas. The optimum field-of-view will be considered in Section 12.5.3 where we discuss the number of available targets in each galaxy per f.o.v.

We adopted a spectral resolution of R=6000, which is the typical resolution of CaT studies in the recent literature. Since we are only interested in the behaviour of the NIR CaT lines, we restrict the wavelength range to 8000–9000 Å. Each resolution element consists of 3 pixels.

An instrumental efficiency was not available at the time this work was carried out. We choose to use a constant value of 0.278 (including the detector QE) so that the total telescope and instrument efficiency is $\sim$0.25 (assuming the Ag/Al coating).

Adaptive Optics and PSFs: For AO mode we assume LTAO. This choice is driven by the need to include as much flux as possible in the chosen spaxel area. In the I-band the PSF effects are particularly strong, with a loss of about 94% of the flux for an area corresponding to 50 mas$\times$50 mas (for a detailed description see [RD1]).

Table 12.5: Runs performed for the NGC 5128 case with the corresponding parameters listed. Columns: (1) run; (2) object; (3) metallicity assumed for the stellar population (see Table 12.1); (4) site; (5) exposure time on source; (6) mirror coating; (7) explored projected distance from the centre in units of R_e; (8) location on the RGB of the target star.

Run	Galaxy	Metallicity	Site	Exp. time	Coating	R/R_e	Target
cenA_old_1	cenA_old	MR	Paranal	5 h	Bare	0.5,1,2,3.5,5,7,10	tip
cenA_old_2	cenA_old	MP	Paranal	5 h	Bare	0.5,1,2,3.5,5,7,10	tip
cenA_old_3	cenA_old	MR	Paranal	2 h	Bare	0.5,1,2,3.5,5,7,10	tip
cenA_old_4	cenA_old	MP	Paranal	2 h	Bare	0.5,1,2,3.5,5,7,10	tip
cenA_old_5	cenA_old	MR	Paranal	1 h	Bare	0.5,1,2,3.5,5,7,10	tip
cenA_old_6	cenA_old	MP	Paranal	1 h	Bare	0.5,1,2,3.5,5,7,10	tip
cenA_old_7	cenA_old	MR	Paranal	5 h	Ag/Al	0.5,1,2,3.5,5,7,10	tip
cenA_old_8	cenA_old	MP	Paranal	5 h	Ag/Al	0.5,1,2,3.5,5,7,10	tip
cenA_old_9	cenA_old	MR	Paranal	2 h	Ag/Al	0.5,1,2,3.5,5,7,10	tip
cenA_old_10	cenA_old	MP	Paranal	2 h	Ag/Al	0.5,1,2,3.5,5,7,10	tip
cenA_old_11	cenA_old	MR	Paranal	1 h	Ag/Al	0.5,1,2,3.5,5,7,10	tip
cenA_old_12	cenA_old	MP	Paranal	1 h	Ag/Al	0.5,1,2,3.5,5,7,10	tip
cenA_old_13	cenA_old	MR	High & Dry	5 h	Ag/Al	0.5,1,2,3.5,5,7,10	tip
cenA_old_14	cenA_old	MP	High & Dry	5 h	Ag/Al	0.5,1,2,3.5,5,7,10	tip
cenA_old_15	cenA_old	MR	High & Dry	2 h	Ag/Al	0.5,1,2,3.5,5,7,10	tip
cenA_old_16	cenA_old	MP	High & Dry	2 h	Ag/Al	0.5,1,2,3.5,5,7,10	tip
cenA_old_17	cenA_old	MR	High & Dry	1 h	Ag/Al	0.5,1,2,3.5,5,7,10	tip
cenA_old_18	cenA_old	MP	High & Dry	1 h	Ag/Al	0.5,1,2,3.5,5,7,10	tip
cenA_old_19	cenA_old	MR	Paranal	5 h	Bare	0.5,1,2,3.5,5,7,10	0.5 below tip
cenA_old_20	cenA_old	MP	Paranal	5 h	Bare	0.5,1,2,3.5,5,7,10	0.5 below tip
cenA_old_21	cenA_old	MR	Paranal	2 h	Bare	0.5,1,2,3.5,5,7,10	0.5 below tip
cenA_old_22	cenA_old	MP	Paranal	2 h	Bare	0.5,1,2,3.5,5,7,10	0.5 below tip
cenA_old_23	cenA_old	MR	Paranal	1 h	Bare	0.5,1,2,3.5,5,7,10	0.5 below tip
cenA_old_24	cenA_old	MP	Paranal	1 h	Bare	0.5,1,2,3.5,5,7,10	0.5 below tip
cenA_old_25	cenA_old	MR	Paranal	5 h	Ag/Al	0.5,1,2,3.5,5,7,10	0.5 below tip
cenA_old_26	cenA_old	MP	Paranal	5 h	Ag/Al	0.5,1,2,3.5,5,7,10	0.5 below tip
cenA_old_27	cenA_old	MR	Paranal	2 h	Ag/Al	0.5,1,2,3.5,5,7,10	0.5 below tip
cenA_old_28	cenA_old	MP	Paranal	2 h	Ag/Al	0.5,1,2,3.5,5,7,10	0.5 below tip
cenA_old_29	cenA_old	MR	Paranal	1 h	Ag/Al	0.5,1,2,3.5,5,7,10	0.5 below tip
cenA_old_30	cenA_old	MP	Paranal	1 h	Ag/Al	0.5,1,2,3.5,5,7,10	0.5 below tip
cenA_old_31	cenA_old	MR	High & Dry	5 h	Ag/Al	0.5,1,2,3.5,5,7,10	0.5 below tip
cenA_old_32	cenA_old	MP	High & Dry	5 h	Ag/Al	0.5,1,2,3.5,5,7,10	0.5 below tip
cenA_old_33	cenA_old	MR	High & Dry	2 h	Ag/Al	0.5,1,2,3.5,5,7,10	0.5 below tip
cenA_old_34	cenA_old	MP	High & Dry	2 h	Ag/Al	0.5,1,2,3.5,5,7,10	0.5 below tip
cenA_old_35	cenA_old	MR	High & Dry	1 h	Ag/Al	0.5,1,2,3.5,5,7,10	0.5 below tip
cenA_old_36	cenA_old	MP	High & Dry	1 h	Ag/Al	0.5,1,2,3.5,5,7,10	0.5 below tip

Table 12.6: Runs performed for the M87 case with the corresponding parameters listed. Columns: (1) run; (2) object; (3) metallicity assumed for the stellar population (see Table 12.1); (4) site; (5) exposure time on source; (6) mirror coating; (7) explored projected distance from the centre in units of R_e; (8) location on the RGB of the target star.

Run	Galaxy	Metallicity	Site	Exp. time	Coating	R/R_e	Target
M87_1	M87	MR	Paranal	50 h	Bare	2,5	tip
M87_2	M87	MP	Paranal	50 h	Bare	2,5	tip
M87_3	M87	MR	Paranal	20 h	Bare	2,5	tip
M87_4	M87	MP	Paranal	20 h	Bare	2,5	tip
M87_5	M87	MR	Paranal	50 h	Ag/Al	2,5	tip
M87_6	M87	MP	Paranal	50 h	Ag/Al	2,5	tip
M87_7	M87	MR	Paranal	20 h	Ag/Al	2,5	tip
M87_8	M87	MP	Paranal	20 h	Ag/Al	2,5	tip
M87_9	M87	MR	High & Dry	50 h	Ag/Al	2,5	tip
M87_10	M87	MP	High & Dry	50 h	Ag/Al	2,5	tip
M87_11	M87	MR	High & Dry	20 h	Ag/Al	2,5	tip
M87_12	M87	MP	High & Dry	20 h	Ag/Al	2,5	tip

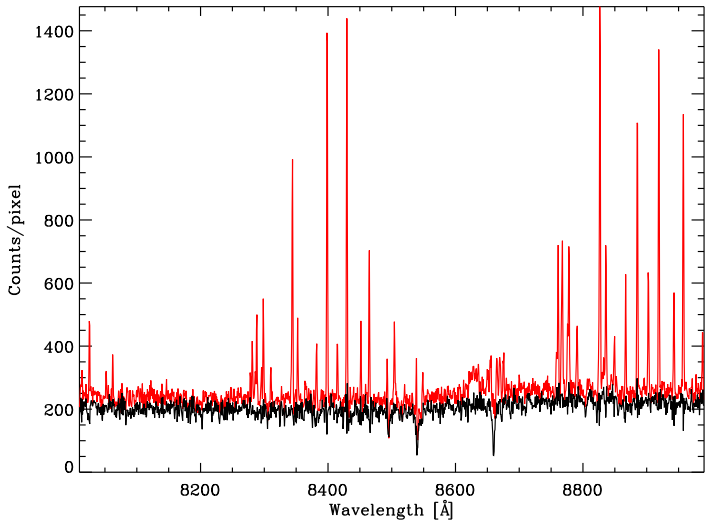

Figure 12.3: Spectrum of a MR ([Fe/H]= -1.0) star at the tip of the RGB, at the distance of NGC 205 (824 kpc), as "observed" with an E-ELT located at a Paranal-like site, with a bare Al coating, in a 20 m exposure. The red includes the sky contribution, while the black shows the sky-subtracted spectrum. The CaT lines can be clearly seen.

12.4.4 Outputs

The output of the simulations simply consists of a list of l.o.s. velocities and ΣW for the simulated integrated spectra and the target RGB star.

12.5 Results of simulations

12.5.1 Simulation runs

The various simulation runs and their parameters are specified in Tables 12.4, 12.5 and 12.6, and are described in the following sections.

12.5.2 Analysis

12.5.2.1 NGC 205

NGC 205 is a nucleated dSph belonging to the M31 system and is located at a distance of (824$\pm$27) kpc from the Sun (McConnachie et al. 2005). The internal velocity dispersion of the system varies along its major and minor axis, but it has an average value of 35 km s^{-1} (Geha et al. 2006). The surface brightness profile in the V-band is well fitted by an exponential law with an effective radius R_e =1.7′ and central surface brightness Σ_e = 20.4 mag arcsec^{-2} (see Mateo 1998, and references therein).

The parameters explored in the simulations for this galaxy are summarised in Table 12.4. The explored exposure times are 20 m, 30 m, 1 h and 2 h per pointing; we consider RGB targets at the tip, 0.5 mag below and 1.0 mag below the tip of the RGB and we explore pointings located at projected radii equal to 0 (centre), 1, 2, 3, 4, 5 R_e (corresponding to a range in surface brightness in the V-band from 20.4 to 25.8 mag arcsec^{-2}). Examples of the spectra obtained for an exposure time of 20 m are shown in Figs. 12.3 and 12.4.

In Fig. 12.5 we show the results of the simulations regarding the accuracy in the recovered l.o.s. velocity and CaT ΣW. We plot the average and scatter in the distribution of velocity and CaT ΣW differences. In this figure we show both the results from the integrated spectrum (target + stellar background) and the

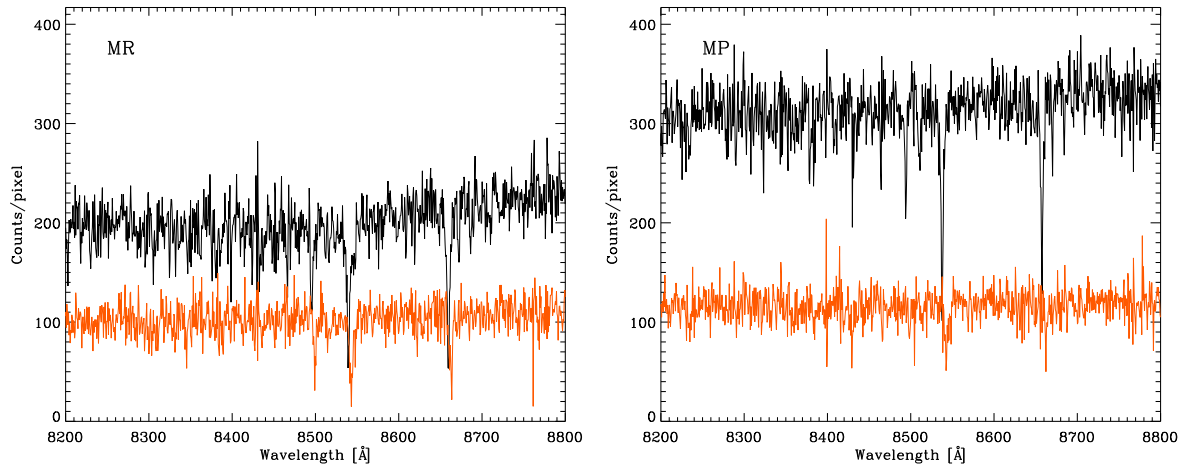

Figure 12.4: Sky-subtracted spectrum of a star at the tip of the RGB (black) and about one magnitude below the tip (orange), at the distance of NGC 205, as "observed" with an E-ELT at R∼6000 for a 20 m exposure time on a Paranal-like site using a bare Al mirror coating (left: MR case; right: MP case). The S/N/pixel of these spectra are 12.8, 17.2 (MR and MP at the tip of the RGB), 9 and 9 (MR and MP 1 mag below the tip of the RGB).

spectrum of the target RGB star alone. Comparison between the two, in terms of systematic and scatter in the recovered values, show that, given the low surface brightness of the object and the small physical area covered by a spaxel of 50 mas per side, crowding and therefore the effect of stellar background is not an issue for this galaxy, not even in the centre. This means that with the instrumental configuration used, stars are resolved in NGC 205. This statement can be extended to the other dSphs and dE in the outskirts of the Local Group, given that they have comparable distances from the Sun and comparable (even lower) surface brightnesses to NGC 205.

Since the stellar background has no effect on the values of the recovered velocity and [Fe/H] (e.g. does not cause appreciable changes in the shape and location of the CaT lines), it is not needed to explore different values of the internal velocity dispersion. In the following we only show the results for the central regions of the galaxy ($R = 1\ R_e$).

In Fig. 12.6 we show how the accuracy in the derived l.o.s. velocity and summed CaT EW varies as a function of the explored exposure times and magnitude of the target RGB star, for both the MR and the MP case. The l.o.s. velocities can be derived very accurately at the distance of NGC 205, both for the MR and the MP case and for all the RGB magnitudes explored. No systematics are present in the determination. For the shortest explored exposure time (20 m) the accuracy varies from about 2 to 3 km s^{-1} from the tip of the RGB to one magnitude below. This is well within the required figure of merit. For the longest exposure time considered (2 h), the accuracy improves to about 1 km s^{-1}. Given that already the less demanding site/coating combination (i.e. Paranal-like and bare) gives very good results, we do not consider the results for the other combinations.

We remind the reader that 94% of the flux of the target star in the I-band falls outside of the central spaxel and is lost because of the PSF shape. Given that the stellar background is not an issue for this galaxy, we note that by co-adding signal from surrounding spaxels it would be possible to reach similar accuracies even with shorter exposure times than have been explored here or it would be possible to target fainter stars with the same exposure times.

The panels on the right-hand side in Fig. 12.6 show the results for the distribution of differences in [Fe/H] values. In general, it appears that the summed CaT EW derived from the simulated spectra is underestimated of 0.1–0.2 dex in the simulated spectrum, i.e. the simulated spectrum appears to be more metal poor than the correct value. For a fixed magnitude, the deviation from the real value becomes smaller for longer exposure times, which suggests that this effect is related to the S/N of the spectrum. We checked that this effect is related to the way the EW of each CaT line is estimated, i.e. by summing the flux over

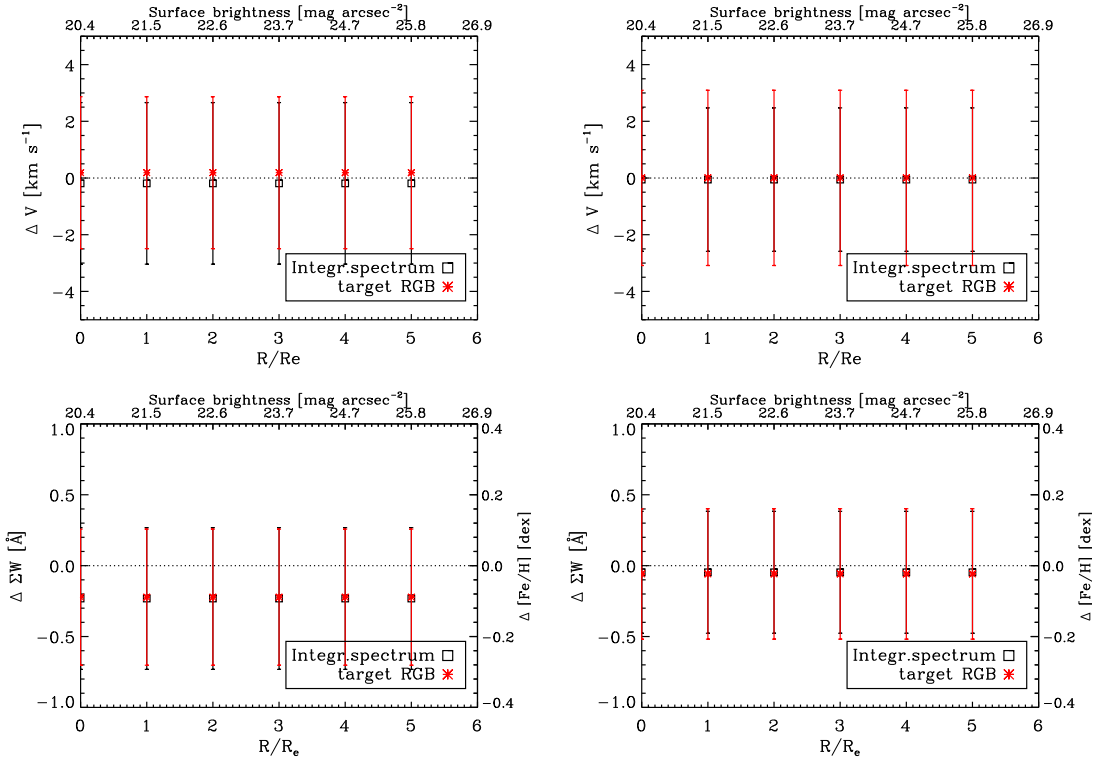

Figure 12.5: Difference in l.o.s. velocities (top) and in summed equivalent widths (bottom) for the MR (left) and MP (right) cases between the values obtained from the simulated spectra and the input values for a target RGB star with 1 mag below the tip of the RGB at the distance of NGC 205, as a function of projected distance from the centre expressed in R/R_e. The exposure time is 20 m, for a Paranal-like site and bare Al mirror coating. The squares show the results for the integrated spectrum in the spaxel and the asterisks for the target RGB star. Comparison between the results given by the squares and the asterisks tell us that the stellar background is unimportant for this case. The symbols and the error-bars show the average difference and its m.a.d. from 300 random realisations (3 for different stellar background mix and 100 each for the noise).

a 15 Å range around each of the CaT lines. Two issues appear to be present: a) a small part of the flux is lost for MR targets because the CaT lines for the MR targets are broader than 15 Å; b) the continuum level appears to be slightly underestimated in correspondence of the CaT lines, and this effect is larger the stronger and/or deeper the line (and therefore for more MR targets and for brighter targets); this provokes an underestimation of the EW. An improvement to the estimate of the EW would be, for example, to perform a Gaussian fit to the individual CaT lines and correct their EW by a factor taking into account the non-Gaussian wings of the lines, as it was done in the VLT/FLAMES observations for which the original version of the pipeline was developed and used. However, since in this case the systematic in the EW determination is due only to the way the EW is determined (and not to the presence of stellar background), we can consider the average value as the zero-point and simply consider that the accuracy in the measurement is just given by the scatter around the average value. Fig. 12.6 shows that also the determination of the summed CaT EW, and therefore [Fe/H], is within the required figure of merit even for the shortest explored exposure time (20 m), the largest errors being 0.25 dex.

We conclude that, as expected, the NGC 205 case does not represent a problem for the E-ELT in terms of achieving the required accuracy. In Section 12.5.3 we discuss the availability of targets per pointing and the total exposure time required.

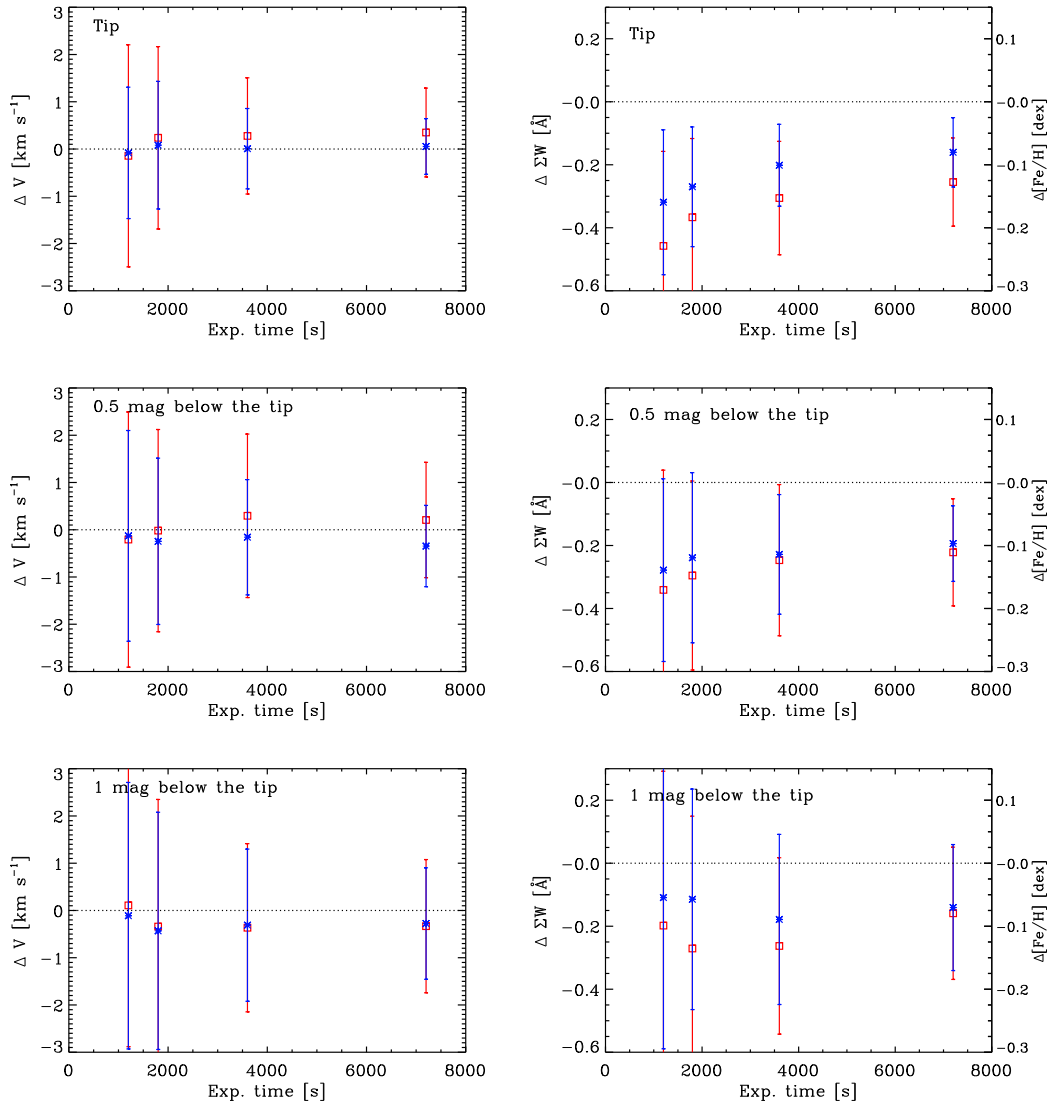

Figure 12.6: Difference in l.o.s. velocity (left) and summed CaT EW (right) between the values obtained from the simulated spectra and the input values for a target RGB star at the distance of NGC 205, as a function of the exposure time on source. For these simulations we have adopted a resolving power $R \sim 6000$, a Paranal-like site and a bare Al mirror coating. The squares show the results for the metal rich case and the asterisks for the metal poor case. From top to bottom, the results are for a target star at the tip of the RGB, about half a magnitude and a magnitude below the tip. The symbols and the error-bars show the average difference and its m.a.d. from random realisations of the noise.

12.5.2.2 Centaurus A

Located at a distance of (3.84±0.35) Mpc from the Sun (Rejkuba 2004), Centaurus A (Cen A), or NGC 5128, is the nearest giant elliptical galaxy and it is the central galaxy of the Centaurus A group. For the internal velocity dispersion of the system we assume 100 km s^{-1}, close to the value suggested by the kinematics of its globular cluster system out to a projected distance of 15 kpc from its centre (Woodley et al. 2007). The surface brightness profile in V is well fitted by a de Vaucouleurs law with an effective radius $R_e = 5.5'$ and effective surface brightness $\Sigma_e = 22.15$ mag arcsec^{-2} (van den Bergh 1976).

In order to explore the effect of crowding we consider projected distances from the centre of 0.5, 1, 2, 3.5, 5, 7, 10 R_e (corresponding to a range in the V-band surface brightness from 20.8 to 28.6 mag arcsec^{-2}).

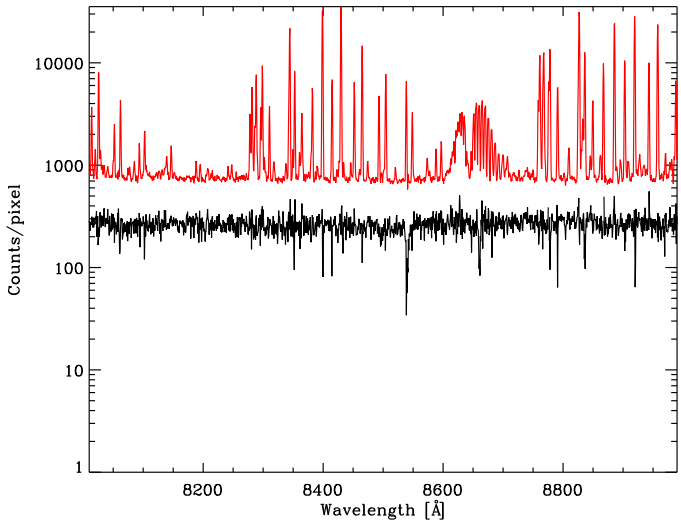

Figure 12.7: Spectrum of a MR ([Fe/H]= -1.0) star at the tip of the RGB, at the distance of Cen A, as "observed" with an E-ELT at R$\sim$6000 for a 5 h exposure on a Paranal-like site using an Ag/Al mirror coating. The red line includes the sky contribution, while the black line shows the sky-subtracted spectrum.

We do not consider larger distances from the centre because from our simulated stellar catalogue we find that at 10 R_e only 2 or 3 stars contribute to the light in the spaxel containing the target star, which means that the stellar background is surely not an issue at larger projected radii. The target RGB stars have a magnitude equal to the tip of the RGB or 0.5 mag below the RGB tip. The parameters explored for the different runs are listed in Table 12.5.

An example of a spectrum obtained in an exposure time of 5 h is shown in Fig. 12.7.

Fig. 12.8 shows an example of a f.o.v. of $5'' \times 5''$ at $R = 2$ and $5\ R_e$, with the target star at the centre of the field. As expected there is more crowding at $R = 2\ R_e$ than at $R = 5\ R_e$. However, when looking at the spectrum obtained for the central spaxel (see Fig. 12.9) almost no difference is visible in the continuum level at the two distances from the centre, suggesting that the target star contributes most of the flux to the spectrum in the considered spaxel, while the stellar background has a negligible effect at these distances.

In Appendix A we show the results on the average and scatter (m.a.d.) for the distribution of l.o.s. velocity differences for the MP and MR case for a target star 0.5 mag below the tip of the RGB (Figs. 12.14 and 12.15, respectively) and for the difference in ΣW (Figs. 12.16 and 12.17). These figures show that, for the explored instrumental configuration, the stellar background has no influence on the measurements for NGC 5128 at all the considered projected radii $\geq 2\ R_e$. At $R = 0.5$ and $1\ R_e$ the influence of the stellar background can be noticed as a large systematic difference in the determination of [Fe/H] from the integrated spectrum with respect to the determination for the individual target star: at $R = 0.5\ R_e$ the difference ranges from about -0.4 dex for the MP case to and -0.8 dex for the MR case, respectively, while at $R = 1\ R_e$ the differences range from about -0.2 dex to -0.4 dex (MP and MR, respectively). The systematic in the velocity determination instead appears negligible at $R = 1\ R_e$, while for the innermost explored radius it is less than 10 km s^{-1}, and therefore – including the error-bars – still within the figure of merit. At $R = 0.5\ R_e$ the velocity difference (recovered velocity $-$ input) for the MP case is about -5 km s^{-1} (here the target has a positive velocity with respect to the systemic velocity of the galaxy), while for the MR target is about 10 km s^{-1} (here the target has a negative velocity with respect to the systemic velocity of the galaxy): this trend indicates that the presence of the stellar background is pushing the recovered velocity towards the systemic velocity of the system, as it would be expected. Most likely the average values obtained will have some dependence on the actual value of the velocity of the target star, but especially this should depend on the fraction of light that the target star contributes to the integrated spectrum with respect to the stellar background, i.e. for fixed stellar backgrounds the velocity differences will be more significant for fainter targets.

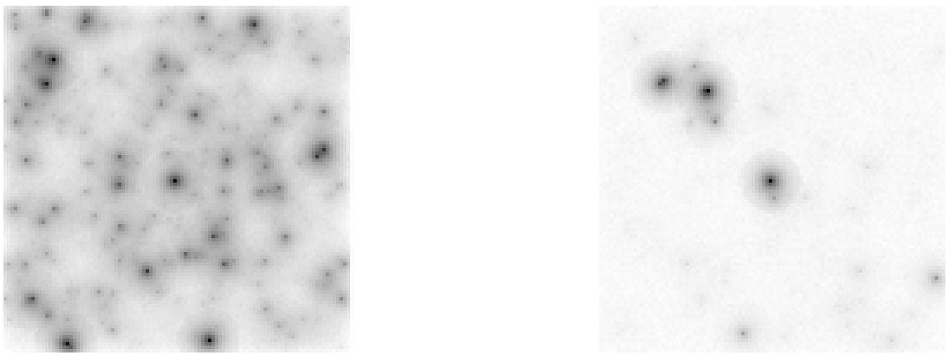

Figure 12.8: Example of a $5'' \times 5''$ stellar field at the distance of the galaxy Cen A, for two different projected distances from the centre of the galaxy (left: 2 R_e; right: 5 R_e). The star at the centre of the field-of-view is the target RGB star, with a magnitude 0.5 mag below the tip of the RGB. The pixel size corresponds to 50 mas, i.e. the spaxel size used throughout this work.

The figures in Appendix A also show that, for a fixed exposure time, the site does not play a significant role in improving the accuracy in the measurements, while the mirror coating improves considerably the achieved accuracies.

A summary of the results for all the different runs is given in Fig. 12.10 which shows how the S/N/pixel and errors in the velocity and [Fe/H] measurements from the integrated spectrum vary as a function of the exposure time per pointing, and for all the different combinations of site and coating. These plots are for $R = 2\ R_e$, where crowding does not have an effect, and the results do not differ much for the larger projected radii explored, as shown in Appendix A. The errors in velocity and [Fe/H] measurements are given as the root-mean-square of the distribution of velocity and [Fe/H] differences, and in this way they include both systematic and random effects.

For the MP target star at the tip of the RGB, the l.o.s. velocities can be derived with accuracies within the figures of merit for all explored parameters, except for the Paranal-like case with bare Al mirror coating and 1h exposure. However, for the bright MR target, a 2 h exposure is needed and with an Ag/Al coating. A 2 h exposure with a bare Al mirror coating does not result in accurate enough velocities. For a target star 0.5 mag below the tip of the RGB, at least a 2 h exposure using an Ag/Al mirror coating is needed to recover velocities with accuracy close to the required one for both the MR and MP case.

The derivation of accurate [Fe/H] measurements instead necessitates of longer exposures. Only the simulations for a MP target at the tip of the RGB with a 5 h exposure and Ag/Al mirror coating do result in recovered [Fe/H] values within the required accuracy. For the same parameters but considering the target at 0.5 mag below the tip of the RGB, the recovered [Fe/H] is about 0.4–0.5 dex from the input value.

As in the case of NGC 205, as most of the target flux falls outside of the central spaxel, the determination of the velocities and metallicities could benefit from summing up the flux from the target in surrounding spaxels, over regions small enough that the crowding can still be neglected. As an example, we consider the MR case at 2 R_e for a target 0.5 mag fainter than the tip of the RGB with a 5 h and 2 h exposure for a Paranal-like site with Ag/Al coating and we derive the integrated spectrum over an area of 250 mas $\times$ 250 mas, i.e. 5 spaxels $\times$ 5 spaxels including the central one, and an area of 90 mas $\times$ 90 mas, i.e. 3 spaxels $\times$ 3 spaxels including the central one. For these simulations we consider again 100 random simulations for the noise in each of the spaxels and consider just 1 random realisation for the stellar background.

We first derive the integrated spectrum for each of the 25 spaxels as described in Section 12.4.2, and then we sum their flux up. Even though a region of 250 mas $\times$ 250 mas around the central spaxel should contain approximately 34.5% of the flux with respect to the 6.5% of when considering the central spaxel alone, the gain in S/N/pixel is not very large: for a 5 h exposure the S/N/pixel is 7.9 over 250 mas $\times$ 250 mas with respect to 6.0 over 50 mas $\times$ 50 mas, and for a 2 h exposure the S/N/pixel is 5.2 over

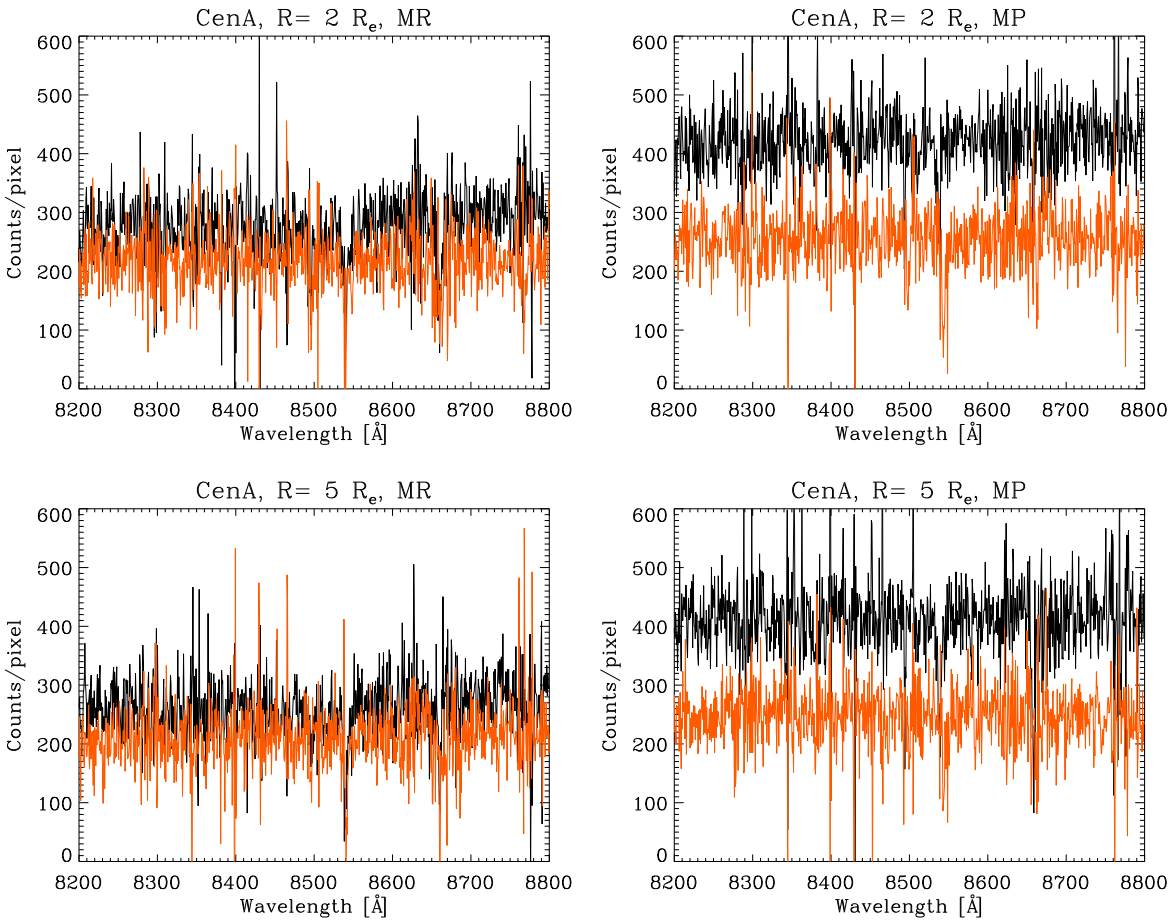

Figure 12.9: Sky-subtracted spectrum of a star at the tip of the RGB (black) and about 0.5 mag below the tip (orange), at the distance of Cen A, as "observed" with an E-ELT at R∼6000 for a 5 h exposure on a Paranal-like site using an Ag/Al mirror coating (left: MR case; right: MP case). The S/N/pixel of these spectra are: 7.2, 10.1 (MR and MP at the tip of the RGB), 5.7, 6.4 (MR and MP 0.5 magnitude below the tip of the RGB). The top panels show the spectra at a projected distance from the centre of 2 R_e and the bottom panels of 5 R_e. The spectra in the inner and outer parts of the galaxy are very similar, hinting that the contribution of the stellar background is small already at 2 R_e (see also Figs. 12.14, 12.15 for the l.o.s. velocity and 12.16, 12.17 for the determination of EW).

250 mas $\times$ 250 mas with respect to 3.9 over 50 mas $\times$ 50 mas. This is likely due to the large noise introduced by the sky-subtraction in each spaxel.

The velocity determination improves slightly when considering the integrated spectrum from a larger area, be it 250 mas $\times$ 250 mas or 90 mas $\times$ 90 mas, making it possible to achieve the required accuracy. However, for the metallicity determination the improvement is quite negligible and the recovered metallicity values are still not within the figure of merit. It appears that with the considered instrumental parameters exposure times longer than 5 h per pointing may be necessary to improve the metallicity determinations.

The fraction of encircled energy within one spaxel appears to be a very important parameter in determining the accuracy in the measurements. As an example, we repeat the simulations for the MR case with a target at 0.5 mag below the tip of the RGB (at an High & Dry site with Ag/Al mirror coating) by increasing the encircled energy within the central spaxel to 18%: already with a 2 h exposure the accuracy in velocity improves to 4 km s^{-1} and the accuracy in [Fe/H] to about 0.25 dex, which is within the figure of merit.

We conclude that the Cen A case does not represent a problem for the E-ELT in terms of achievable accuracies for velocity determinations down to 0.5 mag below the RGB. Accurate [Fe/H] measurements

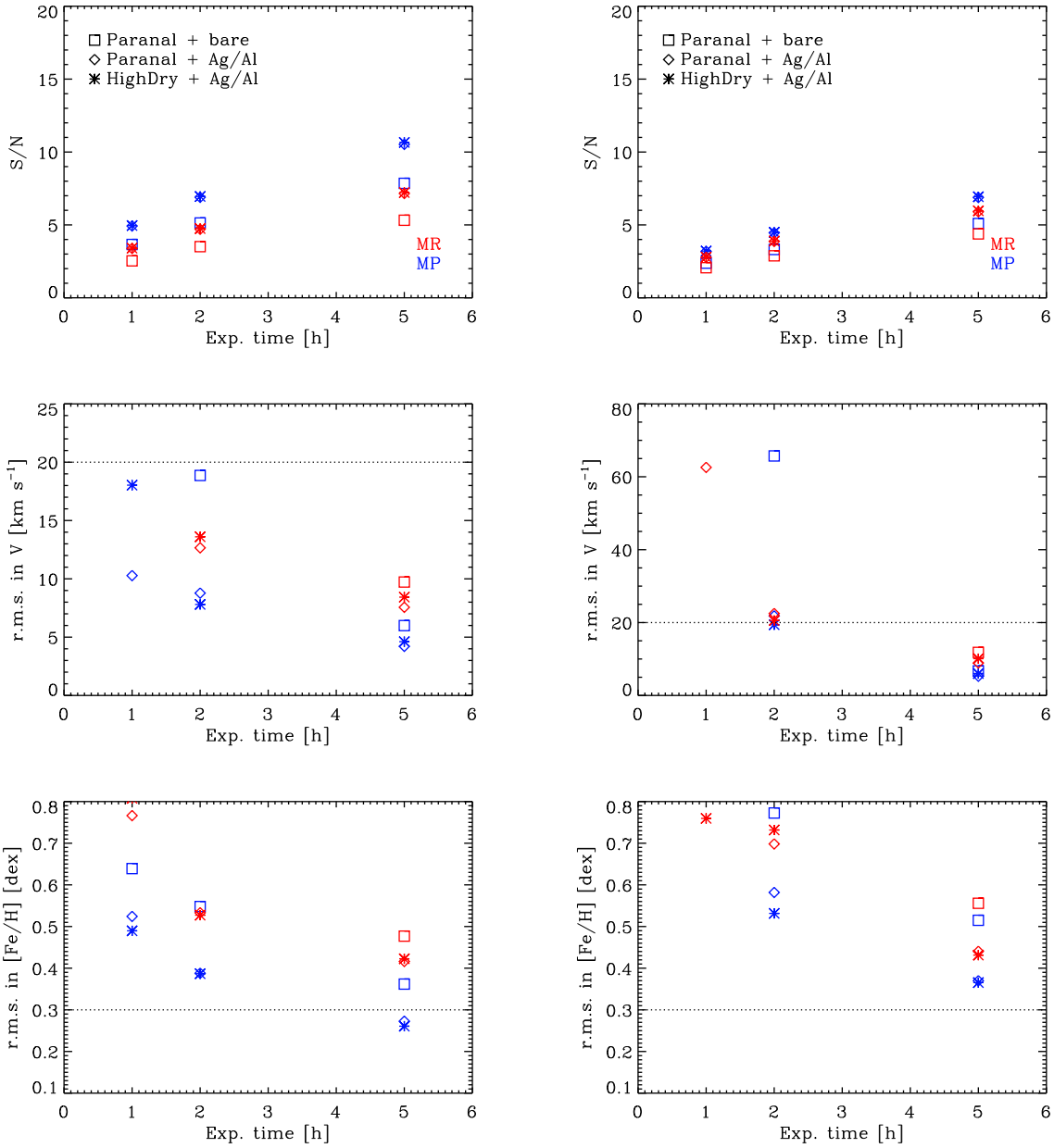

Figure 12.10: Results from the simulations on Cen A as a function of exposure time. Left: for a target star at the tip of the RGB. Right: for a target star 0.5 mag fainter than the tip of the RGB. From top to bottom: S/N per pixel, root-mean-square (r.m.s.) of the distribution of velocity differences, root-mean-square of the distribution of [Fe/H] differences. The red and blue symbols show the results for a MR and MP target, respectively. The squares, diamonds and asterisks show the results for different combinations of site and mirror coating: Paranal-like - bare Al, Paranal-like - Ag/Al, High & Dry - Ag/Al, respectively. The horizontal lines show the cut-off for the acceptable measurement errors.

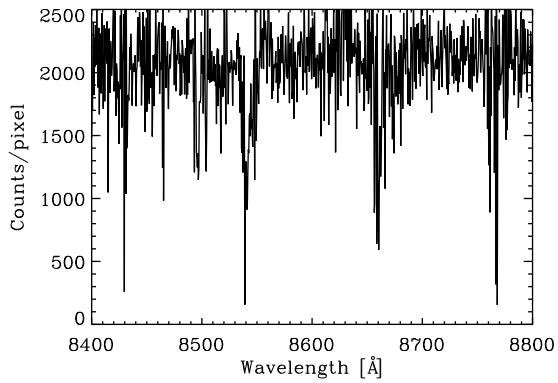

Figure 12.11: Co-added (sky-subtracted) spectrum for 25 spaxels for a target star 0.5 mag fainter than the tip of the RGB at the distance of Cen A, as "observed" with an E-ELT at R∼6000 for a 5 h exposure on a Paranal-like site using an Ag/Al mirror coating. These simulations are at projected distance $R = 2\ R_e$.

are possible for targets at the tip of the RGB, although with exposure times per pointing rather long within our assumptions for the various instrumental parameters; this accuracy could be improved if a larger fraction of encircled energy in the I-band retained in the central spaxel could be achievable. In Section 12.5.3 we discuss the availability of targets and the required total exposure time to survey the galaxy.

12.5.2.3 M87

With distance modulus $(m - M)$∼31, the galaxies in the Virgo cluster represent the biggest challenge for this science case. To explore the feasibility of CaT studies at this distance we perform simulations taking as an example M87, the central Virgo elliptical galaxy. For the the V-band surface brightness of this galaxy we use a de Vaucouleurs profile with an effective radius R_e = 1.742′ (= 104.53″) and an effective surface brightness $\Sigma_e = 20.58$ mag arcsec^{-2} (values are taken from http://goldmine.mib.infn.it/db/).

In this galaxy the effects of crowding will vary greatly going from the central to the outer regions, varying the contribution of stars in neighbouring spaxels to the light in our target spaxel. To address these effects, we explore 2 projected radii in M87, 2 and 5 R_e (corresponding to surface brightness of 22.2 and 24.8 mag arcsec^{-2} in the V-band, respectively). As target stars, we use only those at the tip of the RGB.

In the following, we first use our standard approach of deriving the l.o.s. velocity and CaT [Fe/H] from the spectrum of the central spaxel, as done for NGC 205 and Cen A; then, we try to improve the S/N of the final spectrum by co-adding the flux on a larger area than the single central spaxel.

Standard approach: one spaxel We simulate spectra for exposure times of 20 h and 50 h (the parameters are listed in Table 12.6). Given the much larger number of stars involved at these distances, in order to reduce the computing time, we simulate 3 random realisations of the stellar background and for each 1 random simulation of the noise per spaxel. Therefore, hereafter, instead of providing the scatter in the distribution of recovered velocities and [Fe/H], we consider that the recovered velocity and [Fe/H] are in agreement with the input values when they are within 2σ from the input values (where the σ is 20 km s^{-1} in velocity and 0.3 dex in metallicity as in the previous sections).

At $R = 2\ R_e$, the S/N/pixel of the integrated spectrum changes from 4.4 to 2.2 for the MP case and from 4.3 to 2.15 for the MR case, going from an exposure of 50 h at an High & Dry site with Ag/Al coating, to an exposure of 20 h at a Paranal-like site with bare coating. For comparison, the S/N/pixel of the target RGB star ranges from 1.6 to 0.8 and from 1.0 to 0.5 for the MP case and MR case, respectively: the larger S/N of the integrated spectrum with respect to the spectrum of the target star is due to the contribution of the stellar background, which is considerable at this projected radius even in one spaxel.

The S/N between the integrated spectrum in the MP and the MR case are very similar to each other. This is because the main contributor to the flux of the spectrum is not the target but the stellar background, which is set in the simulations mainly by the value of the surface brightness.

From the results of the simulations no clear trends emerge between the explored parameters and the success rate in the recovery of the velocity and [Fe/H] values.

For the MP case, the velocity is recovered 3 times out of 3 for a 50 h exposure at an High & Dry site with Ag/Al mirror coating and for a 20 h exposure at a Paranal-like site with Ag/Al coating, 1 time for a 20 h exposure at an High & Dry site with Ag/Al coating and none of the times for a 20 h exposure at a Paranal-like site with bare Al coating. The [Fe/H] values are instead recovered 1 out of 3 times in each case.

For the MR target, the velocity is recovered 2 out of 3 times for a 50 h exposure at an High & Dry site with Ag/Al mirror coating, 1 time for a 20 h exposure at an High & Dry site with Ag/Al coating and none of the times in the other cases. The [Fe/H] values are recovered only 2 times in total.

We remind the reader that these simulations assume a single exposure per pointing, while for real observations the total exposure time will have to be split in much shorter observing blocks and this will further decrease the accuracy in the final measured values.

At $R = 5\ R_e$ the S/N/pixel of the integrated spectrum is slightly larger than the S/N of the target RGB star alone, indicating that there is still some contribution of the stellar background, but that it is small. The values for the integrated spectrum range from 1.8 to 0.9 and from 1.3 to 0.7 going from an exposure of 50 h at an High & Dry site with Ag/Al coating, to an exposure of 20 h at a Paranal-like site with bare coating, for the MP case and MR case respectively. The S/N values for the target star alone are as at $R = 2\ R_e$ given that the target stars are the same ones.

At this distance from the centre, there are no systematics in the recovered value of [Fe/H] and the measurements disagree by more than 2σ with the input values in the great majority of cases. The same behaviour is present for the velocities.

We conclude that at both distances from the centre, within the adopted strategy of data reduction, such observations are very challenging.

Maximising the S/N: co-addition of multiple spaxels Given the large loss of flux in the I-band due to the small spaxel size, we analyse the result of co-adding the flux from spaxels surrounding the central one, which also contain flux from the target RGB star, to try and improve the S/N of the final spectrum. However, the more we increase the spatial area, the more we also increase the contribution of neighbouring stars to the final spectrum. We followed two cases: co-adding the flux over 9 spaxels and over 25 spaxels, and we performed this operation at both 2 R_e and at 5 R_e.

2 R_e: Fig. 12.12 shows the comparison between the spectrum for the central spaxel and the co-added spectrum of 25 spaxels for a MR and MP target, for a 50 h exposure time, for a High & Dry site with Ag/Al coating.

Clearly the S/N of the co-added spectrum improves considerably with respect to the spectrum of the central spaxel alone. With an exposure time of 50 h, the S/N/pixel is around 11.5 when summing up the flux over 25 spaxels and around 7.5 when summing up over 9 spaxels. For an exposure time of 20 h (MP case), we have a S/N/pixel in the range 7.5–10 when summing up 25 spaxels, and in the range 5–6.5 when summing up 9 spaxels. In both cases the lowest S/N is for the Paranal-like site with bare Al coating and the largest is for the High & Dry site with Ag/Al coating, where the mirror coating is the parameter that plays the largest role in affecting the S/N, while the site has barely any effect.

Although the quality of the spectrum clearly improves by summing up multiple spaxels, Fig. 12.12 shows that it is the stellar background that accounts for most of its flux. The CaT lines of the spectrum from multiple spaxels are considerably displaced in wavelength and have a very different shape with respect to the target RGB star. In the majority of the simulation – independent of the metallicity, exposure time, site, coating, number of spaxels summed – the velocity found for the co-added spectrum is shifted towards the

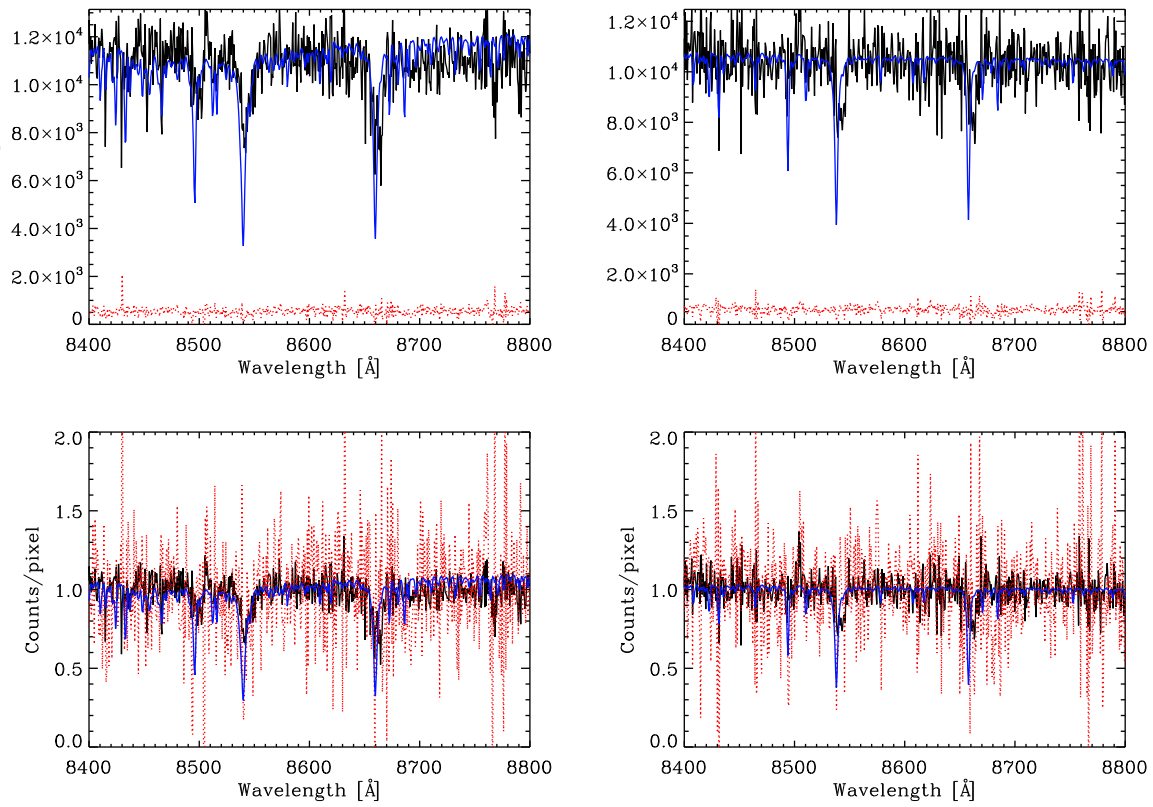

Figure 12.12: Top: Co-added (sky-subtracted) spectrum for 25 spaxels (black), and the central spaxel alone (red), for a target star at the tip of the RGB at the distance of M87, as "observed" with an E-ELT at R∼6000 for a 50 h exposure on a High & Dry site using an Ag/Al mirror coating (left: MR case; right: MP case). The blue line shows the input spectrum for the target star at the tip of the RGB, convolved at R∼6000 and scaled to match the continuum of the co-added spectrum. The flux in the central spaxel alone is about one order of magnitude lower than when co-adding the signal over 25 surrounding spaxels. Bottom: as top panels, but for normalised spectra. These simulations are at projected distance $R = 2\ R_e$.

systemic velocity of the whole galaxy, giving rise to very large systematic differences with respect to the velocities of the target star.

The differences between the ΣW of the summed spectra and the input values are also very large and range mostly between -1 dex and -2 dex.

It is thus likely that neither reducing or increasing the number of spaxels used to form a spectrum would bring a considerable improvement in the results: in the first case the S/N of the total spectrum would drop, producing similar results to the integration of one single spaxel (point a); while in the second case the contribution of the stellar background would increase, washing out even more the signal from the target RGB star.

We conclude that at projected distances $\leq$ 2 R_e from the centre of M87, the techniques used for resolved stellar populations cannot be applied, and we are effectively in the situation where we need to use techniques from integrated stellar population studies. These results can be extended to galaxies at similar distances at positions where the surface brightness is similar to that we considered here, i.e. 22.2 mag arcsec^{-2}.

5 R_e: At this distance from the centre of the galaxy the influence of the stellar background represents less of an issue. However the S/N of the co-added spectrum is not always sufficient to obtain reliable measurements and to establish trends with exposure time, coating, etc.

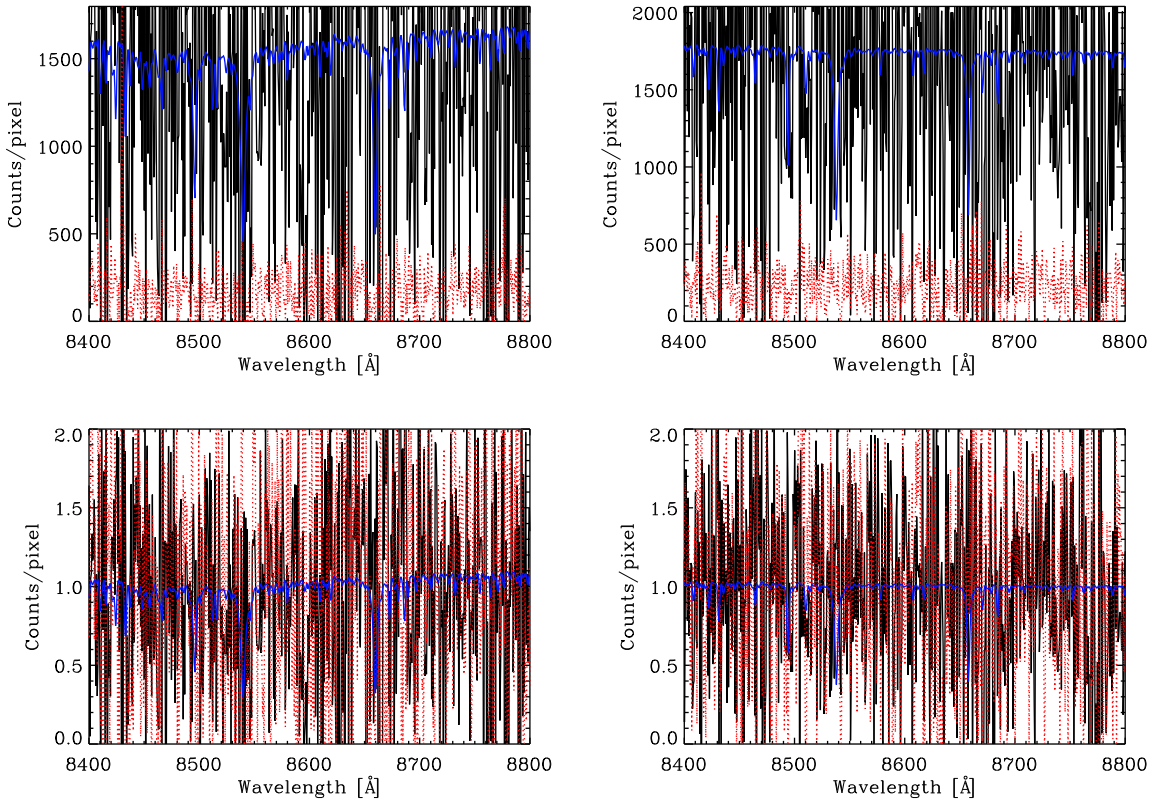

Figure 12.13: As in Fig. 12.12 but for a projected distance $R = 5\ R_e$.

From our simulations for the parameters listed in Table 12.6, we find that the ΣW of the two strongest CaT lines cannot be measured within the desired accuracy in the majority of the cases. The typical difference from the input value is about ± 2 Å, resulting in [Fe/H] differences of about ± 1 dex, but for some simulations the [Fe/H] differences can be as large as several dex .

In the great majority of the cases the input l.o.s. velocity cannot be recovered within 2σ, where $\sigma = 20$ km s^{-1} is the maximum allowed error in velocity, both when co-adding 9 or 25 spaxels. The few times the velocity is recovered no trend with coating, site and exposure time is apparent indicating that statistical effects, due to low S/N and therefore large errors in the measurements, are playing a role.

From these simulations it appears that, at this projected distance from the centre of M87, the determination of l.o.s. velocities is challenging even with exposure times of 20 h and 50 h for very bright RGB stars (corresponding to the tip of the RGB for a MP old population). It is possible that co-adding even more spaxels, or adopting a more sophisticated strategy for the data reduction, may improve the determinations.

It is likely that more random simulations of the stellar background would be needed to assess under which conditions CaT observations of the outer regions of galaxies at the distance of Virgo can be carried out. However, something else should be taken into account for these challenging observations. The simulations carried out here are somewhat idealised as, for example, the instrumental performance is assumed to be similar to the existing one for FORS and the exposure time is considered as one single exposure of 20–50 h. When splitting up the exposure time in smaller observing blocks, most likely one hour long, more noise will be added because of the read-out, which will lower the S/N of the final spectrum.

Instrument-simulators, which provide a closer match to what may be the final performance of the instrument, should be used to assess the feasibility of CaT studies in the outer regions of galaxies at the distance of Virgo. This case is clearly at the edge of what we currently think is possible.

Table 12.7: Number of possible target stars in an area of $10'' \times 10''$ for NGC205, at various projected distances from the centre of this galaxy and in different magnitude ranges (see also Table 12.2 for the values of probability to find stars in these magnitude ranges).

NGC205		MR			MP		
R/R_e	$\Sigma(R/R_e)$ [mag arcsec^{-2}]	tip	0.5 below tip	1 below tip	tip	0.5 below tip	1 below tip
0	20.40	5.93± 0.01	39.67± 0.06	69.47± 0.10	5.17± 0.01	29.41± 0.05	70.66± 0.12
1	21.49	2.17± 0.01	14.54± 0.04	25.46± 0.06	1.90± 0.01	10.78± 0.03	25.89± 0.07
2	22.57	0.80± 0.00	5.38± 0.02	9.41± 0.04	0.70± 0.00	3.98± 0.02	9.58± 0.04
3	23.66	0.29± 0.00	1.97± 0.01	3.45± 0.02	0.26± 0.00	1.46± 0.01	3.51± 0.03
4	24.74	0.11± 0.00	0.73± 0.01	1.28± 0.01	0.10± 0.00	0.54± 0.01	1.30± 0.02
5	25.83	0.04± 0.00	0.27± 0.00	0.47± 0.01	0.03± 0.00	0.20± 0.00	0.48± 0.01

Table 12.8: As Table 12.7 but for the case of Cen A. Here we exclude the region at $R < 1\ R_e$ since the stellar background starts to contribute significantly to the light of the integrated spectrum.

Cen A		MR		MP	
R/R_e	$\Sigma(R/R_e)$ [mag arcsec^{-2}]	tip	0.5 below tip	tip	0.5 below tip
1	22.15	25.66± 0.02	171.59± 0.12	22.38± 0.02	127.18± 0.10
2	23.74	5.93± 0.01	39.67± 0.06	5.17± 0.01	29.41± 0.05
3.5	25.25	1.48± 0.00	9.87± 0.03	1.29± 0.00	7.32± 0.02
5	26.32	0.55± 0.00	3.69± 0.02	0.48± 0.00	2.73± 0.02
7	27.43	0.20± 0.00	1.33± 0.01	0.17± 0.00	0.98± 0.01
10	28.71	0.06± 0.00	0.41± 0.01	0.05± 0.00	0.30± 0.01

Table 12.9: As Table 12.7 but for the case of M87.

M87		MR	MP
R/R_e	$\Sigma(R/R_e)$ [mag arcsec^{-2}]	tip	tip
4	24.10	87.34± 0.03	76.17± 0.03
5	24.80	45.84± 0.02	39.98± 0.02

12.5.3 Compliance with figures of merit

In Section 12.5.2 we estimated the accuracies in the determination of l.o.s. velocities and ΣW from CaT lines for galaxies in the Local Group, Cen A and Virgo. We established the exposure time needed per pointing so that the accuracies of the measurements would comply with the accuracies needed to meet the science goals.

In order to evaluate the compliance of the required sample size to the figure of merit, we now need to estimate if the total exposure time needed to observe a sample of about 1000 stars per galaxy, spread over a large fraction of the galaxy, is within the required limit of 100 h.

For this we need to calculate the number of available targets per galaxy and the number of pointings required for a reasonable field-of-view.

We consider first a single IFU with field-of-view of $10'' \times 10''$ and calculate the average number of RGB stars per pointing available as targets in the desired magnitude range[2].

The probabilities to find a star in a certain magnitude range are listed in Table 12.2. In order to find the number of possible target stars, we first simulate a catalogue for the given stellar population and surface brightness for an area corresponding to the field-of-view of the instrument; we then multiply the total number of stars found in the field-of-view, $N_{\rm fov}$ by the probability of finding a star in the required magnitude range, p: this gives us an estimate of the number of targets, $N_{\rm tar} = p \times N_{\rm fov}$. We assume that a Poissonian error applies for $N_{\rm fov}$, i.e. $\sigma_{\rm fov} = \sqrt{N_{\rm fov}}$; using error propagation, we estimate that the error in $N_{\rm tar}$ is $\sigma_{\rm tar} = p \times \sigma_{\rm fov}$.

[2]As magnitude range we consider stars with I magnitudes between the RGB tip and 0.1 mag below the magnitude of the target RGB star.

The number of targets per field-of-view in the specific magnitude ranges for NGC 205, Cen A and M87 are listed in Tables 12.7, 12.8 and 12.9, respectively, for several projected radii.

As shown in the previous section, crowding is a problem at distances $\leq$2 R_e in M87 and most likely, even though we did not perform the simulations, still is around 3 R_e; therefore we will exclude those radii from the calculation of the number of pointings needed.

NGC 205: It was possible to reach 1 mag below the tip of the RGB with an exposure time of 20 m per field. At 0, 1, 2 R_e, with the adopted surface brightness profile and stellar population, approximately 70, 25, 10 targets are available per field. This means that to collect about 1000 targets in total, about 4, 12, 30 pointings at around 0, 1, 2 R_e are sufficient with an instrument of $10'' \times 10''$ field-of-view. This would amount to about 15 h exposure time on source.

At the distance of NGC 205, a projected radius of 2 R_e corresponds to approximately 0.8 kpc from the centre of the galaxy. In order to explore the outer regions of such an object, which obviously will contain a smaller number of targets per pointing, either an instrument with a larger field-of-view would be desirable, or a longer exposure time so as to reach fainter magnitudes and increase the number of possible targets per pointing.

The physical size encompassed by the instrument field-of-view at the distance of NGC 205 is small, about 40 pc. It is known that the kinematic and metallicity properties of even small systems such as dwarf galaxies vary considerably throughout these objects; therefore, in order to quantify spatial variations in those properties, one would have to cover a larger fraction of the galaxy. An instrument with a larger field-of-view would allow to cover a larger fraction of the galaxy and keep the total exposure time short.

Clearly, the case for NGC 205 complies with the figure of merit both in terms of the required accuracy in the measurements and the ability to obtain a large sample of measurements in an acceptable total exposure time.

Cen A: We estimated that an exposure time of 5 h per field would yield the required accuracies in l.o.s. velocity for stars as faint as 0.5 mag below the tip of the RGB and for [Fe/H] measurements for MP targets at the tip of the RGB. Instead, an exposure time of 2 h per field would yield the required accuracies in l.o.s. velocity for stars as faint as 0.5 mag below the tip of the RGB but not accurate enough [Fe/H] measurements.

At 1, 2, 3.5 R_e, approximately 150, 35, 8 targets are available per field. This means that to collect about 1000 targets in total, about 2, 8 and 35 pointings at 1, 2, 3.5 R_e are sufficient. Considering a 5 h exposure per field, this would amount to about 225 h exposure time on source. This observing time is larger than what we had set as reasonable.

However this could be improved, either by increasing the number of targets that could be collected per pointing – for example using an instrument with a larger field-of-view or multi-IFUs – and/or by recovering part of the flux lost because of PSF effects. Regarding the latter point, our attempt of summing up the flux from surrounding spaxels containing signal from the target star did not allow to significantly improve the S/N and the resulting accuracies in the measurements because the sky-subtraction in each of the co-added spaxels introduces a large amount of noise. Perhaps a more sophisticated analysis when co-adding the spaxels would improve the determinations. Alternatively, we showed that a larger fraction of encircled energy in the area of the spaxel would improve the results: if an encircled energy within the central spaxel as large as 18% in I-band would be achievable, this would allow l.o.s. velocity and [Fe/H] measurements within the figure of merit for the faintest explored targets at this distance in 2 h exposure per pointing. This would decrease the total exposure time to about 90 h, within what we had set as reasonable.

It should finally be noted that, in order to explore regions with surface brightness similar to the internal regions in Cen A, the spaxel size should remain small to avoid the effects of crowding. In this sense, instruments like HARMONI and OPTIMOS-EVE could be complementary for observations of a galaxy like Cen A, where HARMONI would be better suited to observe the inner regions and OPTIMOS for the outer regions.

We conclude that the case for Cen A complies with the figure of merit in terms of the required accuracy in the velocity measurements and in terms of the total exposure time needed to observe a suitably large sample of stars with accurate velocities. Gathering large samples of accurate metallicity measurements appears to require total exposure times larger than our figure of merit within our standard assumptions for the parameters that we used in this work.

Virgo/M87: The observations at the distance of M87 are challenging. At distances from the centre of the galaxy where crowding is not an issue, e.g. about 4–5 R_e, there are about 80 and 40 targets, respectively, in a $10'' \times 10''$ field-of-view at the tip of the RGB from which accurate l.o.s. velocities can be obtained. Therefore, of the order of 5 and 10 pointings at those distances should provide the desired number of targets. For an exposure time of 20 h, this would amount to a large total exposure time of 300 h, while for 50 h a total of 750 h would be needed. However, these observations appear to be difficult even with the large exposure times explored in our idealised simulations.

As in the previous cases, the total exposure time could be decreased by observing a larger number of targets at once, again for example with an instrument with a field-of-view larger than $10'' \times 10''$, and/or by increasing the fraction of encircled energy within the area of one spaxel.

12.5.4 Sensitivity to input parameters

Site: In the explored wavelength range, between 8000 and 9000 Å, the differences in the transmission curves for a Paranal-like site and an High & Dry site are negligible. Therefore, in this work, the figures of merit are not sensitive to the choice of site.

Mirror coating: In general, there is an improvement in the simulation results when using a Ag/Al coating rather than bare Al, due to the higher transmission curve of this mirror coating in the wavelength range of CaT.

In the case of Local Group galaxies, the choice of one coating over the other does not play a role in the feasibility or optimisation of the studies. For the Cen A case, both the determination of l.o.s. velocities and CaT ΣW benefit from using an Ag/Al mirror coating over a bare Al one. For example, for the fainter considered target at the distance of Cen A for a 2 h exposure per pointing, using an Ag/Al mirror coating we obtain velocities very close to the figure of merit, while when using the bare Al mirror coating the r.m.s. in the recovered velocities is about 60 km s^{-1} instead of the maximum allowed 20 km s^{-1}. For the M87 case, for which the observations are the most challenging, usually only the best combination of parameters was able to provide a useful recovery success rate.

Instrument efficiency: The instrument efficiency assumed in these simulations resembles the efficiency of current slit spectrographs such as FORS, but it may be too optimistic a value, and it is assumed constant with wavelength. This is most likely not an issue for the explored cases of galaxies in the Local Group and in Cen A, but it will certainly play an important role in determining if CaT studies of resolved RGB stars can be carried out at the distance of Virgo.

Spaxel size: The spaxel size used throughout this work is 50 mas. For smaller sizes, even more flux from the target star would fall outside of the spaxel and be lost because of the PSF shape, resulting in longer exposure times to reach the required accuracies. It is possible that for the inner regions of the M87 case, where crowding is an issue, a smaller spaxel would reduce the pollution from the stellar background. However, at the same time, a smaller fraction of the light from the target would be retained inside the spaxel. The trade-offs should explored with simulations.

Larger spaxel sizes would allow to collect more signal from the target star, assuming that the LTAO could still be used. Otherwise, one should consider the trade-off between the increased spaxel size and the AO type available for that instrumental configuration. Another aspect to consider would be the increased contribution from the stellar background.

Field-of-view: As discussed in Section 12.5.3, a single-IFU instrument with $10'' \times 10''$ field-of-view is sufficient to acquire the desired number of measurements in the explored cases.

However, the physical size encompassed by such a field-of-view is small compared to the size of the galaxies we considered, especially for galaxies at the distance of Cen A and even for dwarf galaxies in the Local Group. One of the goals of CaT studies is to determine the large-scale properties of the resolved stellar populations over all spatial scales in a galaxy. Several studies have shown that, even in objects as small as dwarf spheroidal galaxies, the kinematical and chemical properties vary spatially within these systems. This means that studies limited to a small fraction of these galaxies may provide biased views of their properties and therefore it is advisable to cover as much of the projected area of the galaxies as possible. This means that an instrument providing a larger spatial coverage than $10'' \times 10''$ would be preferable (here we assume that the spaxel size would still be 50 mas or not considerably larger, as to keep the same crowding conditions as in the simulations here examined).

PSF: As discussed in Section 12.5.2.2, a larger fractional encircled energy in the considered band within the spaxel containing the target star would be desirable for the Cen A case, as it would allow to collect large samples of accurate l.o.s. velocities and [Fe/H] for the faintest considered targets with total exposure times within our figures of merit. Even though we did not explore it with simulations, it is very likely that also the observations for targets at the distance of M87 would benefit by that.

12.5.5 Limitations and possible improvements

Among the possible limitations of the simulations, and possible or required improvements, the following should be considered:

- In these simulations we assume a perfect data reduction.
- In the regimes of very low S/N such as in the M87 case, it might become an issue to separate the total exposure in several shorter exposures, for example as short as 1h, the typical length of observation blocks at ESO telescopes.
- The effect of airmass is not considered.
- A more realistic instrument efficiency may be needed.
- Given that 94% of the flux in I-band falls outside the spaxel, a more sophisticated data analysis technique could be adopted for NGC 205 and Cen A, such as summing up the flux from the target in spaxels surrounding the considered one.
- In regions where crowding is an issue, such as the inner radius considered for the M87 case, a more sophisticated data reduction may allow to disentangle the signal of the target from the contribution of the stellar background.

12.6 Concluding remarks

We have carried out simulations to determine how well determinations of large samples of l.o.s. velocities and EWs from intermediate resolution spectroscopy (R$\sim$6000) in the region of the NIR CaT for individual RGB stars can be achieved in three different galaxies at a range of distances – from the Local Group out to the Virgo cluster – and a range of surface brightnesses (and therefore projected radii) within each object.

Different sites, mirror coatings and exposure times have been considered in the simulations.

A number of assumptions had to be made to perform the simulations. Some of them concern the physical properties of the objects under consideration (such as the properties of the underlying stellar population, in terms of chemistry and kinematics); others concern the performance of the telescope and the instrument. The latter assumptions will be subject to change as the designs of the telescope and instruments progress.

We find that the science goal of obtaining accurate l.o.s. velocities and CaT EWs for a large number of individual stars, spread over a wide area, in Local Group galaxies can be achieved using the E-ELT

well within the equivalent of the observing time allocated for a normal program at the VLT. Accurate l.o.s. velocities for large number of individual stars, spread over a wide area, in galaxies out to the distance of Cen A can also be acquired in the equivalent of the observing time allocated for a normal program at the VLT, but large samples of metallicity measurements appear to require longer observing times.

Observations of individual stars at the tip of the RGB at the distance of the Virgo cluster require much longer exposure times and are necessarily limited to the outer parts of these galaxies, under the explored parameters. Achieving the goal of collecting measurements for about 1000 individual stars in one single galaxy at that distance will likely require observing time equivalent to an ESO Large Program.

Acknowledgements

Giuseppina Battaglia gratefully acknowledges Mike Irwin for kindly making his velocity and EW extraction pipeline for FLAMES/GIRAFFE data available for adaptation to this work. GB thanks Joe Liske for providing the code for generating stellar population catalogues and helpful discussions. GB also acknowledges A. Calamida for help with IRAF routines, M. Rejkuba for useful information on the characteristics of Cen A and E. Emsellem for useful suggestions.

References

Aaronson M., 1983, ApJ, 266, L11
Battaglia G., et al., 2006, A&A, 459, 423
Battaglia G., Irwin M.J., Tolstoy E., Hill V., Helmi A., Letarte B., Jablonka P., 2008, MNRAS, 383, 183
Battaglia G., Helmi A., Tolstoy E., Irwin M.J., Hill V., Jablonka P., 2008, ApJ, 681, L13
Choi P.I., Guhathakurta P., Johnston K.V., 2002, AJ, 124, 310
Cordier D., Pietrinferni A., Cassisi S., Salaris M., 2007, AJ, 133, 468
Fraternali F., Tolstoy E., Irwin M.J., Cole A.A., 2009, A&A, 499, 121
Geha M., Guhathakurta P., Rich R.M., Cooper M.C., 2006, AJ, 131, 332
Gray D.F., *"The Observation and Analysis of Stellar Photospheres"*, 3rd edition, Cambridge, UK: Cambridge University Press, 2005
Kirby E.N., Guhathakurta P., Sneden C., 2008, ApJ, 682, 1217
Kleyna J.T., Wilkinson M.I., Gilmore G., Evans N.W., 2003, ApJ, 588, L21
Lanfranchi G.A., Matteucci F., 2007, 468, 927
Lewis G.F., et al., 2007, MNRAS, 375, 1364
Lin D.N.C., Faber S.M., 1983, ApJ, 266, L21
Marcolini A., D'Ercole A., Battaglia G., Gibson B.K., 2008, MNRAS, 386, 2173
Mateo M., 1998, ARA&A, 36, 435
McConnachie A.W., et al., 2005, MNRAS, 356, 979
Munari U., Sordo R., Castelli F., Zwitter T., 2005, A&A, 442, 1127
Pietrinferni A., Cassisi S., Salaris M., Castelli F., 2004, ApJ, 612, 168
Rejkuba M., 2004, A&A, 413, 903
Rejkuba M., et al., 2005, ApJ, 631, 262
Revaz Y., et al., 2009, A&A, 501, 189
Rutledge G.A., et al., 1997, PASP, 109, 883
Starkenburg E., et al., 2010, A&A, 513, A34
Tolstoy E., et al., 2004, ApJ, 617, L119
Tolstoy E., Hill V., Tosi M., 2009, ARA&A, 47, 371
van den Bergh S., 1976, ApJ, 208, 673
Walker M.G., et al., 2006, ApJ, 642, L41
Woodley K.A., et al., 2007, AJ, 134, 494

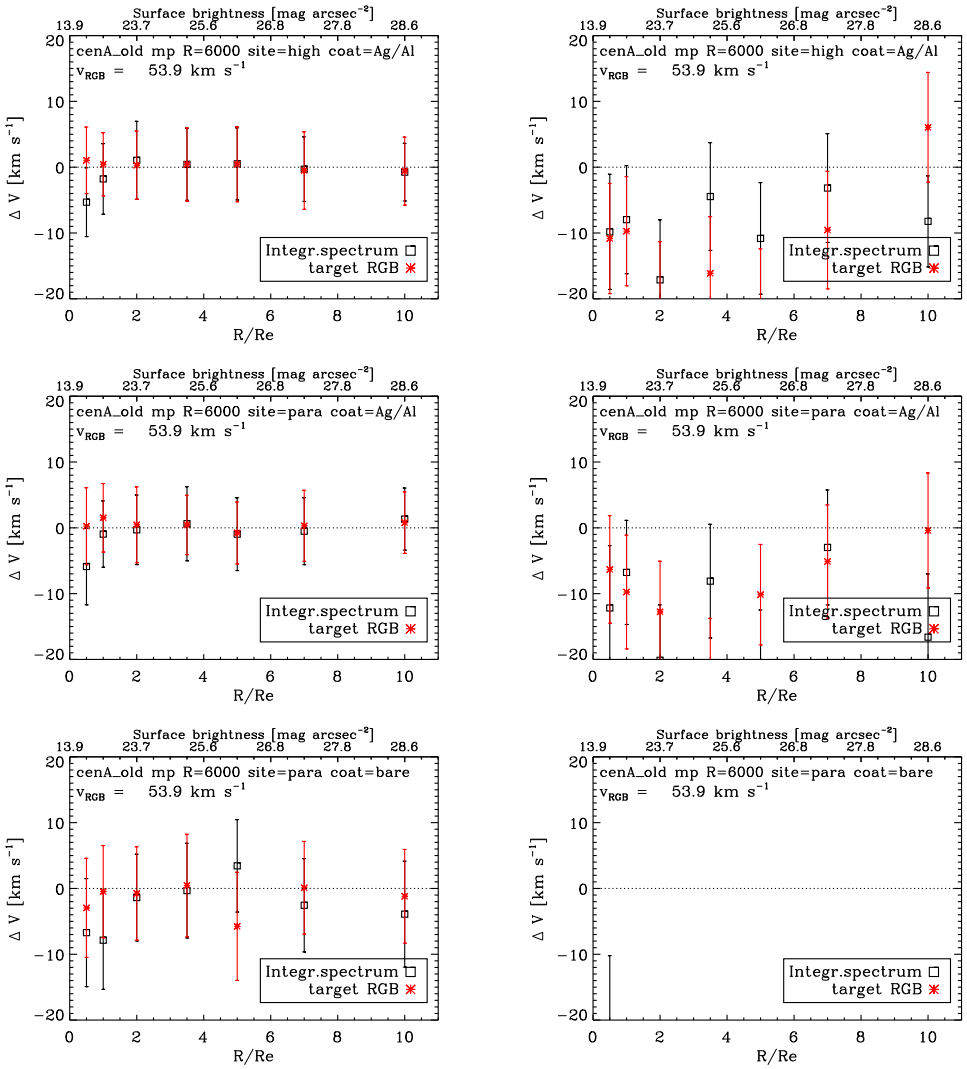

Figure 12.14: Difference in l.o.s. velocities for the Cen A MP case between the values obtained from the simulated spectra and the input values for a target RGB star 0.5 mag below the tip of the RGB at the distance of Cen A, as a function of projected distance from the centre expressed in R/R_e. From left to right, the exposure time is 5 h, 2 h. From top to bottom, the combination of site and mirror coating explored are respectively: High & Dry - Ag/Al, Paranal-like - Ag/Al, Paranal-like - bare Al. The black squares show the results for the integrated spectrum in the spaxel and the red asterisks for the target RGB star alone. Comparing the squares and the asterisks tells us about the importance of the effect of stellar background. The symbols and the error-bars show the average difference and its m.a.d. from 3 random realisations of the stellar background, and for each of these 100 random realisations of the noise.

Appendix A

In this Appendix we show the results on the average and scatter (m.a.d.) for the distribution of l.o.s. velocity differences for the MP and MR case for a target star 0.5 mag below the tip of the RGB (Figs. 12.14 and 12.15, respectively) and for the difference in ΣW (Figs. 12.16 and 12.17). In these figures we plot both the results for the integrated spectrum (target RGB star + background) and the results for the target RGB star alone. All the figures show that, for the explored instrumental configuration, the stellar background has no influence on the measurements for NGC 5128 at all the considered projected radii $\geq 2\ R_e$. At $R = 1\ R_e$ the influence of the stellar background starts being noticeable as a systematic difference in the determination of the velocity and [Fe/H] from the integrated spectrum with respect to the determination for

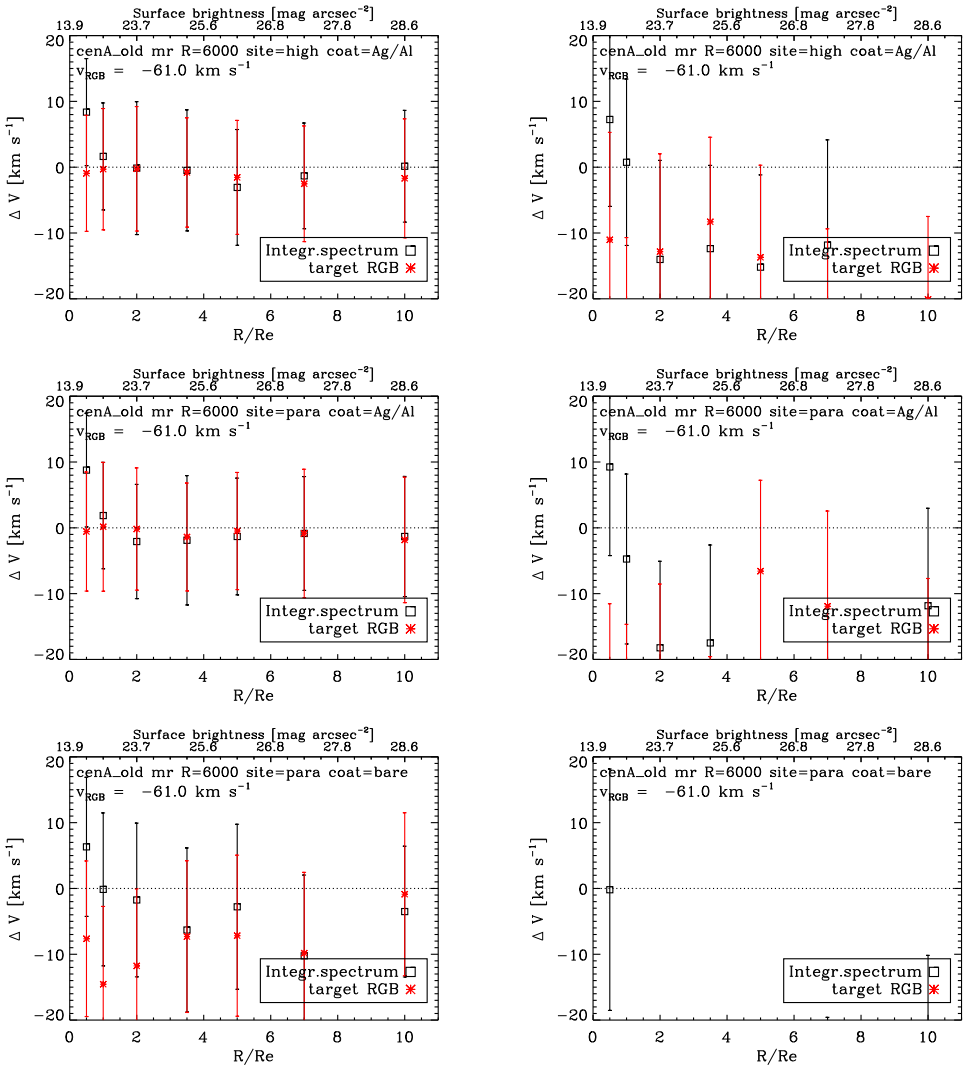

Figure 12.15: As Fig. 12.14 but for the MR case. Here the target RGB star is fainter than in the previous figure.

the individual target star. Since the contribution of stellar background is negligible for most of the cases, we do not consider the effect of variations in the internal kinematics of the object, or in other words of allowing for a different internal velocity dispersion.

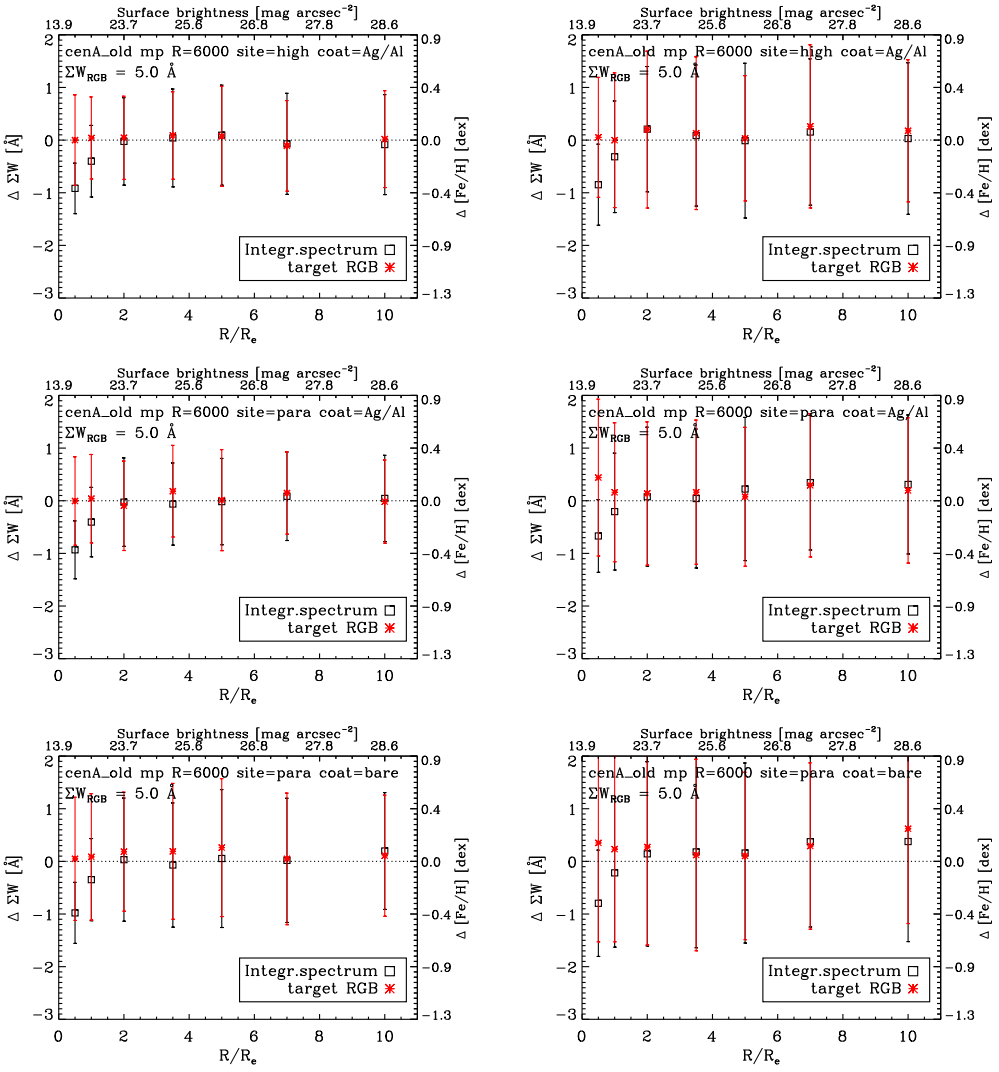

Figure 12.16: As Fig. 12.14 but for the difference in ΣW for a MP target star.

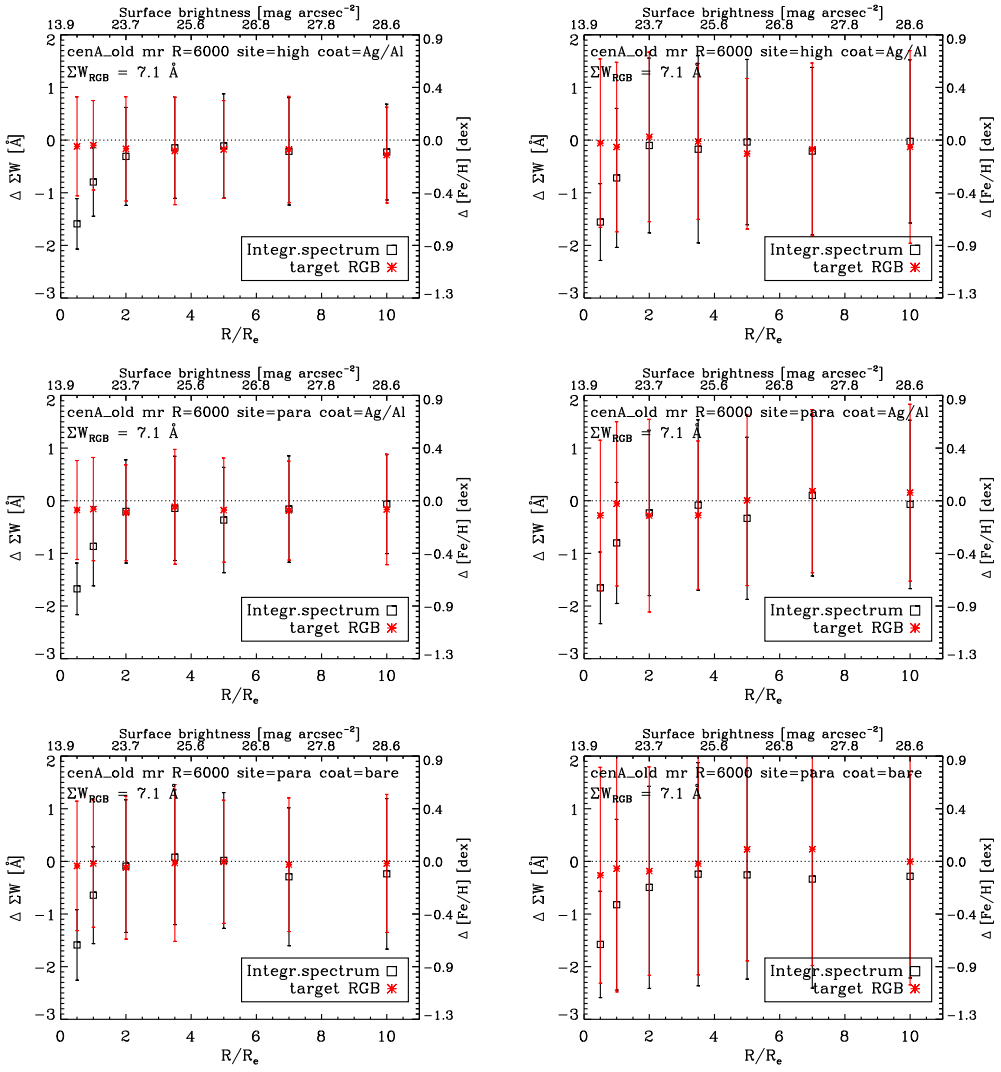

Figure 12.17: As Fig. 12.16 but for the MR case.

13 G4-3: First stars relics in the Milky Way and satellites

Authors: G. James, S. Feltzing, P. François

13.1 The science case

There are two basic ways to probe the very first stellar formation events: look for population III stars at high redshift, or observe, locally, the nucleosynthetic imprints of these first stars. These stars still display in their atmospheres the imprints of metal-enrichment by population III stars, and allow us to gain insight on the nature and nucleosynthesis of the earliest chemical enrichment processes, provided that the detailed abundance patterns of these stars can be derived (Cayrel et al. 2004; Cayrel 2006). Here we focus on this latter approach.

Searches for the most metal-poor stars in the galactic halo are advancing fast (e.g., Christlieb et al. 2008; Schörck et al. 2009, and references therein) and in the coming decade we expect to have large samples of extremely metal-poor (EMP) stars identified throughout the Milky Way halo (see Fig. 13.1). Additionally, large surveys (e.g., SDSS) have revealed a plethora of ultra-faint dwarf spheroidal galaxies (see, e.g., Belokurov et al. 2007). These are currently being explored in more detail and we are starting to see a fair number of quite metal-poor (giant) stars as well as unusual abundance patterns (interesting examples include the studies by Koch et al. 2008; Feltzing et al. 2009; Simon et al. 2010; Frebel et al. 2010; Norris et al. 2010, which have targeted such galaxies as Hercules, Boötes I, Ursa Major II, etc.). Also in the classical dwarf spheroidal galaxies we are starting to find both very metal-poor stars (Aoki et al. 2009; Cohen & Huang 2009) as well as odd abundance ratios (Fulbright, Rich, & Castro 2004).

However, neither in the halo stars nor in the stars in the nearby dwarf galaxies are we able to probe those elements that are of prime importance for discussing the most important aspects of early nucleosynthesis. Due to the faintness of the stars ($V > 17$–18 mag), the spectra we can obtain are both low-S/N and low-resolution. By the time ELTs will come into operation, future EMP surveys should have been extended to even larger volumes and depth, reaching magnitudes of B = 20–21 mag, which means there will be significant samples outside the Milky Way (see Table 13.1), making this search even more challenging. In this DRM case we will show that only with a 42 m telescope equipped with a high-resolution spectrograph that is sensitive both in the optical and in the blue are we able to make progress on the study of the nature and nucleosynthesis of the earliest chemical enrichment processes.

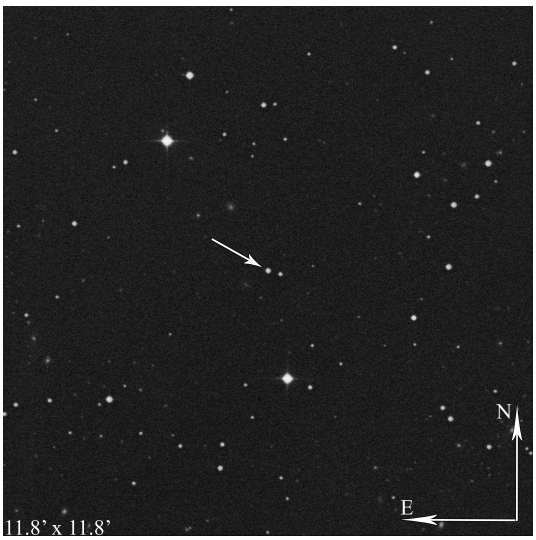

Figure 13.1: $11.8' \times 11.8'$ field of view around HE 0107-5240, one of the two most Fe-deficient stars found up to now. EMP stars are usually isolated objects in the Galactic halo (and in the few nearby galaxies where they have been seen up to now) and make thus quite easy targets once they have been identified.

Table 13.1: Ongoing and future EMP surveys for future high spectral resolution follow-up, and expected number of stars with $[\mathrm{Fe/H}] < -3$ and -5, the so-called hyper metal-poor stars (extrapolated from the HES survey, courtesy N. Christlieb).

Survey	Hemisphere	Start	Eff. sky coverage (deg^2)	Eff. mag. limit	$N < -3$ (EMP)	$N < -5$ (HMP)
HES	South	1989	6400	$B < 17.0$	400	4
SEGUE	North	2005	1000	$B < 19.0$	1000	10
LAMOST	North	2007	10 000	$B < 19.0$	10 000	100
SSS	South	2007	20 000	$B < 18.0$	5000	50

The major scientific questions that could be answered by an in-depth, spectroscopic abundance study of the most metal-poor stars in the Milky Way and its satellites include:

- What is the true lowest-Z boundary for low-mass stars?
  - How does star-formation proceed close to zero metallicity?
  - What is the role of extremely massive population III stars in the build-up of the metals?
  - Is the outer Halo preferentially populated by EMP stars (as suggested by the finding of [Fe/H] = -5 stars[3] in the HES sample, a category of stars absent from more nearby surveys)?
  - Is the lowest metallicity reached in a particular galaxy correlated with the baryonic or dark matter mass of the galaxy (Helmi et al. 2006)?
- Can the Lithium abundances (both ^{6}Li and ^{7}Li) observed in extremely metal-poor dwarf stars be reconciled with Big Bang nucleosynthesis under the concordance cosmological model?
- What is the role of cosmic ray spallation in the early Milky Way production of Be and ^{6}Li?
- What is the age of the most metal-poor stars (radioactive age-dating using Th and U, e.g., Hill et al. 2002; Frebel et al. 2007)?

In order to answer these and related questions we need to consider which elements give us information about what. Here we briefly discuss which elements probe what.

Oxygen is the most abundant element after He. Because the production of oxygen depends on the mass of the star, oxygen abundances enable us to put constraints on the nature of populations III stars.

We still do not know where **r-process elements** are formed. Because of the rare nature of the r-process, it is a very good probe of mixing of SNe products in the early galactic gas. Additionally, some heavy elements produced through this process, e.g., U and Th, are radioactive and decay over time and thus provide a clock which can be used to date very metal-poor stars.

Beryllium probes cosmic-ray spallation in the earliest phases of Galactic evolution. Reaching down to $B \sim$ 15–15.5 mag will turn the current handful of stars ([Fe/H] < -2.5) where Be is measurable into a whole population (some 100 turn-off stars are currently known in the galactic halo with [Fe/H] < -3 below B of about 15 mag).

Nitrogen is one of the current major puzzles for early nucleosynthesis. It is a secondary element that should therefore be under-abundant in metal-poor stars but instead shows a primary behaviour and is even enhanced in the record-holders -5 metallicity stars.

The lithium abundance in old stars provides constraints on Big Bang nucleosynthesis and Ω_{baryon}. So far, its abundance has only been measured in the Milky Way. The E-ELT will for the first time give us the opportunity to measure it in another galaxy (other DM halo).

Iron abundances do not only yield the metallicity of stars but also provide an independent mean to constrain the stellar parameters such as the stellar surface gravity – or luminosity – and hence distance. It is important to obtain data for both neutral and singly ionised iron to achieve this (ionisation equilibrium).

[3]Here we use the classical notation where $[\mathrm{A/B}] = \log(N_\mathrm{A}/N_\mathrm{B}) - \log(N_\mathrm{A}/N_\mathrm{B})_\odot$, and $\log(\mathrm{A/H}) \equiv \log \epsilon_\mathrm{A} = \log(N_\mathrm{A}/N_\mathrm{H}) + 12$.

13.2 Goals of the DRM simulations

Can we reach the optimal combination of S/N and spectral resolution needed to explore the early nucleosynthesis and build-up of metals? If yes, is this feasible in a reasonable amount of observing time for a facility that will certainly be highly demanded like the E-ELT?

Due to the faintness of the targets, achieving a sufficient S/N at high resolution is not possible with today's equipment. For example, the recent studies of the ultra-faint dwarf spheroidal galaxies (Koch et al. 2008; Feltzing et al. 2009; Simon et al. 2010; Frebel et al. 2010; Norris et al. 2010) never reach S/N per pixel higher than 10–40 at 450 nm and 60–90 around 650 nm (in the very best cases). Typical resolutions used in these studies vary from 20 000 to 40 000. Such combinations (together with the uncertainties in the atmospheric parameters) result in typical errors of 0.1–0.2 dex (sometimes more) in the abundance determinations (see also Aoki et al. 2009), which is far from what is needed if we want to achieve, e.g., precise age dating using U and Th lines. One of the various scientific goals of the next generation of telescopes will be to achieve a direct measurement of the age of Galactic halo stars with an accuracy of no more than 0.2 Gyr compared to $\geq$ 1 Gyr using globular clusters nowadays (see, e.g., Gratton et al. 2003, and subsequent works). This is roughly equivalent to uncertainties of less than 0.05 dex on the U and Th abundances. Today, with modern spectrographs and 8–10 m class telescope, r-process enriched EMP stars can be observed in the Galactic halo with resolutions close to R = 80 000 and very high S/N ($\geq$ 200), but this is still not sufficient to get the uncertainties down to better than 0.1 dex for U or Th abundances, yielding an uncertainty on the age not better than a few Gyr (Hill et al. 2002; Cayrel et al. 2004; Frebel et al. 2007). Moreover these "extreme" observations can hardly be done within a reasonable observing time already in the Galactic halo: with a magnitude of V = 11.7 mag, the Uranium-star CS 31082–001 needed close to 10 hours of observing time with VLT/UVES (at $R \sim$ 80 000) to reach the required S/N and cover all necessary UV and optical wavelength domains for a thorough abundance analysis. This alone lets us imagine the prohibitive exposure times that current telescopes would need to reproduce such measurements outside of the Milky Way!

Before answering the question regarding how high S/N we need to achieve to make headway, we need to consider which elements and which spectral lines are of interest to us. Given the predicted as well as observed strengths of these lines in extremely metal-poor dwarf and giant stars (see references above and also Fig. 13.2), we will then be able to set the figures of merits that we need to meet in order to accomplish our goals.

In the table below we give a brief summary of the wavelength ranges we will consider, and which elements are observable (for metal-poor stars) in these regions.

310 – 370 nm	Be	O and N	r-process elements
370 – 520 nm	Fe	α and iron peak elements	r-process elements
670.7 nm	Li		

In the following we briefly expand on this:

310 – 370 nm:

- Beryllium has only a couple of useful lines at 313 nm. This lines are extremely weak in metal-poor stars and will require typically a resolution of at least 50 000 and $S/N \geq 150$ for EMP dwarfs (see, e.g., Rich & Boesgaard 2009, and references therein).
- Many r-process elements can only be detected bluer than 370 nm, most notably Nb, Mo, Ru, Ag and Ir, Pt, Bi. This needs both very high resolution ($\geq$ 50 000) and very high S/N ($\geq$ 100), as can be seen in any classical abundance analysis of EMP stars (e.g., François et al. 2007, and references therein).
- Oxygen has a few transitions in the V, R and H bands, but all of these become undetectable at the lowest metallicities, leaving the OH bands in the UV, around 310 nm, the only probe of oxygen abundance in these stars (e.g., Rich & Boesgaard 2009). Since these bands are quite sensitive to variations in the O abundance, they are easily detectable and need only $S/N \sim$ 50 (but still $R \sim$

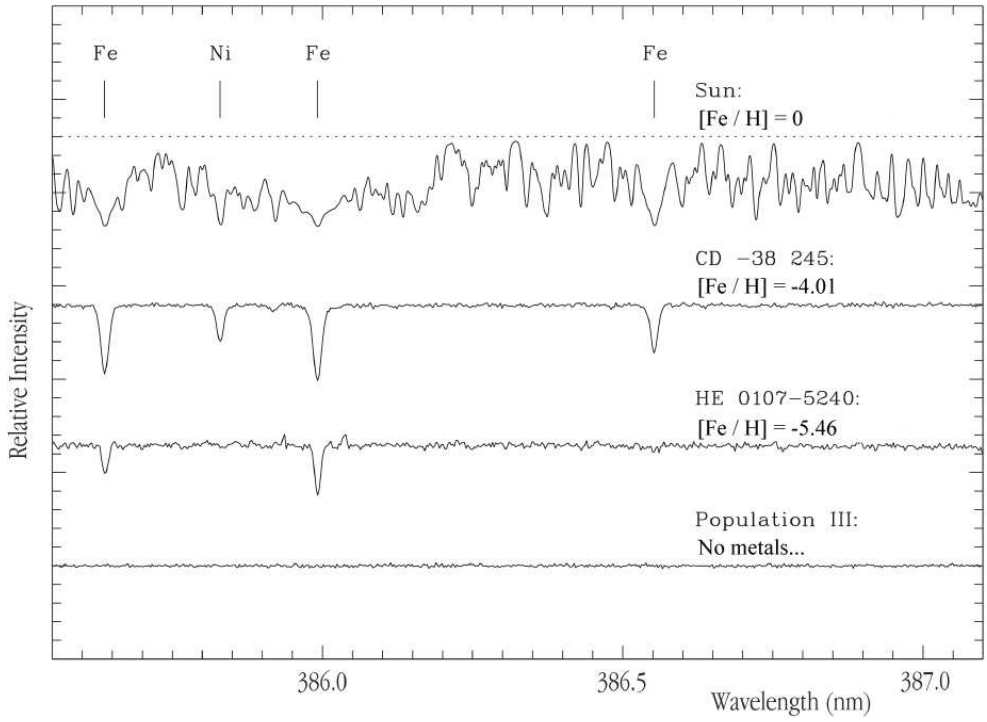

Figure 13.2: Comparison of spectral line strengths in the blue for stars of different Fe abundances. The stars targetted here are of $[\mathrm{Fe/H}] < -3$ to -4. In the red wavelength domain, only a very few and exceedingly weak lines persist at these metallicities.

50 000 to disentangle them from the various other lines/species present around this wavelength range), e.g. in EMP giants.

- Similarly, nitrogen is only detectable at the lowest metallicities from the NH bands around 336 nm. The bandhead is quite wide, and a successful detection can be achieved already with $R = 20\,000$ and $S/N \sim 70$ (or at R =50 000 with $S/N \geq 40$) in EMP stars (see, e.g., Spite et al. 2005).

370 – 520 nm: Since all lines become weaker and weaker with decreasing metallicity (see Fig. 13.2), most of the elements in extremely metal-poor stars can only be measured in the blue. These include:

- Numerous lines of Fe to constrain the metallicity and stellar parameters. As seen in classical abundances analyses (see references above), this requires already a fairly high resolution ($> 40\,000$) and decent S/N (≥ 80) in order to lower the uncertainties on the atmospheric parameters as much as possible.
- α-elements (Mg, Ca, Ti) and iron-peak elements (Mn, Co, Ni, Zn) allow us to constrain the nucleosynthetic origin of the metal-enrichment and hence the nature and metal-production of population III stars. The detection and analysis of lines of these elements needs roughly the same R and S/N as the ones given above for Fe.
- The radioactive elements Th and U (formed through the r-process) can be used for age-dating stars. Their most prominent lines are at 401.9 nm and 385.9 nm, respectively. Both are extremely weak, the Th line reaches some 30 mÅ while the U line reaches at most a few mÅ and will only be clearly visible in a few cases of r-process enriched stars if $[r/\mathrm{Fe}] \gg +1$ dex (Hill et al. 2002; Frebel et al. 2007). The minimal requirements for these lines are obviously the ones set by today's limits, which are $R > 70\,000$ and $S/N > 200$. In order to reach the optimal case where an age determination could be done with the highest precision, we would need $R \geq 100\,000$ and $S/N \gg 200$.

670.7 nm – Li: Li is detected by its 670.7 nm doublet, and will be measurable in the relatively nearby Sagittarius dwarf galaxy stars. The lines are quite weak at very low $[\mathrm{Fe/H}]$, but at the same time very isolated, so that they are easily resolved with $R \sim 50\,000$ at least in the case of dwarf (turn-off) stars, as soon as a S/N of about 60 is reached around this wavelength, yielding an accuracy of at least ± 0.1 dex.

Spectrographs with enough resolution, stability and spectral coverage going from the UV to the near-IR exist already on several medium-sized or large telescopes, but they all lack the capacity of going really deep in magnitude and thus to carry on with accurate abundance determinations in EMP stars outside the Milky Way, and even in distant parts of our own Galactic halo. In particular, sensitive cases like cosmochronometry cannot be achieved yet even in the nearest satellite galaxies. In principle, similar observations could be done with VLT/ESPRESSO in its 4-UT combined mode, provided it goes all the way to UV wavelengths, which is technically very difficult given the light path to the common focus. Even in the (unlikely) event that the combined 4-UT mode reaches the equivalent of a real 16 m telescope, the limiting magnitudes that would be reachable there would still be too shallow to reach the outer Galactic halo stars and the nearest satellite galaxies (LMC, SMC, Sagittarius, Sculptor, Draco, ...). A telescope of the size of the E-ELT equipped with the proper high-resolution spectrograph is thus particularly crucial for the success of this science because of its incredible photon collecting power.

The goal of these simulations is to demonstrate which parts of the science case will be feasible depending on input parameters such as the E-ELT diameter, coating, or observing site, and the characteristics of an associated high-resolution spectrograph (resolution, wavelength coverage, efficiency, ...). That means that our simulations should clearly indicate if and when the required combinations of resolution and S/N will be reached. The simulations should also show how deep we can go with the analysis of EMP stars in the Galactic halo and in nearby satellite galaxies depending on the same input parameters.

13.3 Metrics / figures of merit

Past and current works have already shown what kind of EMP stars we expect to find in the Galactic halo and in nearby satellite galaxies (see references above), but many of the actual targets that will be observed by the E-ELT and other future facilities still need to be identified by the ongoing EMP surveys (see Table 13.1). Since the main issue today is the photon collecting power (i.e. the size of the telescope), S/N combined with the spectral resolution R is one major figure of merit that we need to know in order to decide if we can achieve the scientific goals discussed above.

A first figure of merit could already be Fig. 13.3, where we summarise in a visual way the combination of resolution and S/N needed for a successful detection of a given chemical element and its associated transitions (OH and NH bands, Li, Fe, Fe-peak elements, α-elements, Beryllium, r-process elements, ...). The cyan part of the graph shows the most common values and sets a first lower limit for our constraints: for most cases we will need $R \geq 50\,000$ (at all wavelengths) and at least $S/N > 80$ at $\lambda > 400$ nm if we want accurate determinations of the target's atmospheric parameters (using Fe lines). This is the minimal requirement for a basic abundance analysis, but such limits also guarantee that the conditions needed for other cases are matched (at least OH, NH, Fe-peak, α-elements, and even some r-process elements). A higher constraint on the S/N comes from the Be abundance, which needs (at the same resolution) a $S/N \sim 150$ in the UV, but cosmochronometry using U and Th lines sets an even higher constraint, as it not only requires the same kind of very high S/N in the blue, but it also increases the need in spectral resolution to $R \geq 70\,000$ (shown in green in Fig. 13.3). With this figure, we can easily understand that we need a high-resolution spectrograph capable of reaching at least $R \sim 70\,000$ combined with a telescope (or rather a system Telescope+Instrument+Detector[4]) that can go to S/N values up to at least 200 around the Uranium line at 385.9 nm (in a "reasonable" exposure time, which we set to ≤ 10 hours of integration for a given target to match what is actually still acceptable with a large telescope such as the VLT) if we want to be able to successfully study all the cases described here.

The outcome of our simulations should be figures of merit indicating clearly if the required resolution and S/N, and hence precision on the abundance determination, have been reached for a successful realisation of this science case depending on the telescope diameter, coating, site, and the characteristics

[4] In the text we will use the notation TEL+INS+DET for Telescope+Instrument+Detector.

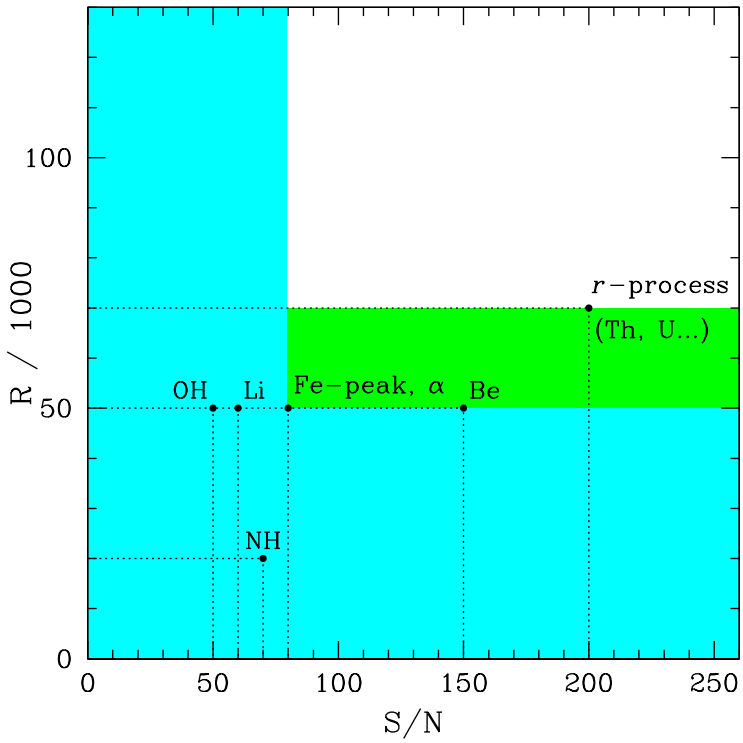

Figure 13.3: General requirements on resolution and S/N needed to obtain accurate abundance determinations in EMP stars ($\leq \pm 0.1$ dex).

of the high-resolution spectrograph. We can identify three kinds of figures that will help deciding wether the goals have been reached:

1. spectral synthesis fitting of absorption lines (e.g. U, Li, Be, OH, ...) showing the accuracy reached at least at the optimal combination of R and S/N for different kinds of EMP stars;
2. S/N as a function of the target's magnitude with identification of the optimal S/N needed for each case, depending on the input parameters for the system TEL+INS+DET;
3. a realistic estimate of the exposure time needed to reach our objectives as a function of the target's magnitude, depending also on these input parameters.

13.4 DRM simulations

13.4.1 Methodology

We will split the simulation into two main parts:

Simulation of final products and synthetic spectral line fitting: First of all, we created so-called "final products", i.e. stellar spectra resulting from the observation of different kinds of EMP stars at a given resolution and with a given S/N. This was done by spectral synthesis using existing model atmospheres and by addition of the proper noise and spectral sampling in order to match realistic observing conditions. These final spectra were fitted with theoretical synthetic spectra to check to which accuracy we will be able to determine chemical abundances in EMP stars according to the resolution, S/N and spectral type of the target.

Simulations using the E-ELT Spectroscopic ETC: We used the E-ELT Spectroscopic Exposure Time Calculator (ETC) to check the feasibility of this science case depending on various parameters, such as

the spectral type and brightness of the target (hence the limiting magnitudes), the telescope size, coating and selected site, the sky conditions (seeing, airmass), but also depending on the instrumental setup (spectral resolution and coverage, AO mode if required).

The results presented in the sections hereafter will give a good preview of several spectral regions that should be analysed in EMP stars in order to answer the questions asked in Section 13.1, and also an estimate of the observational limits and hence the feasibility of this science case.

13.4.2 Pipeline

The simulations of the final products and the synthetic spectral line fitting were carried out as follows:

1. Stellar spectra of different EMP stars were created using MARCS model atmospheres combined with the spectral synthesis code `turbospectrum` (B. Plez, priv. comm., for a description see, e.g., Hill et al. 2002, and subsequent works). We used several linelists for molecules and chemical elements to reproduce the content of the different analysed wavelength ranges, with the latest available atomic data found in the VALD database, and considered all known species present in the studied wavelength domains. The resolution (and thus also the step in λ) of these spectra was imposed by the choice of the spectrograph and was a parameter in this version of `turbospectrum`.

2. Different levels of (Poisson) noise were added to these spectra using a simple `Fortran` routine in order to reproduce the (optimal) observing conditions of these EMP stars, and to study the optimal requirements for an accurate abundance determination depending on the spectral resolution of the instrument, the wavelength domain and the predicted line strengths for all chemical elements.

3. The abundance analysis of these "observed" EMP stars was performed using classical spectral line synthesis tools as those described in, e.g., François et al. (2007), and also using the original version of `turbospectrum`. We have used the same kind of MARCS model atmospheres and spectral synthesis codes as Hill et al. (2002) used for their analysis of the first discovered Uranium-star (CS 31082–001) and all subsequent works of the "First Stars" series of papers.

The simulations using the E-ELT Spectroscopic ETC were carried out as follows:

1. The E-ELT Spectroscopic ETC [RD1] was used to determine the S/N ratio obtained in a 3600 s exposure of typical EMP stars with a given reference apparent magnitude. The main parameters are listed below:

   Target input flux distribution: MARCS stellar models and full synthetic spectra matching classical EMP stars were downloaded from the MARCS website and incorporated in an offline version of the ETC. The targets are all point sources. The ETC is run with an arbitrary apparent magnitude, which will be used as a reference point to extrapolate further results using simple relations between S/N, apparent magnitude, exposure time, telescope size, etc.

   Telescope setup: The reference site is a Paranal-like (2600 m) place, but we also ran simulations for a High & Dry site (5000 m). The telescope reference diameter is 42 m, but other sizes were also considered. The mirror coating is also a parameter here.

   Sky conditions: We performed these simulations only for a clear sky with a seeing of 0.8″ (close to the average yearly seeing of Paranal without any specific ground layer correction). Not knowing if the E-ELT will technically be able to point exactly at the zenith, we chose to run the ETC using an airmass of 1.15 instead of 1.0, which is close enough to optimal observing conditions anyway.

   Instrument setup: Since this science case does not have any requirements on the spatial resolution and therefore on AO, the simulations were done in seeing limited (FWHM=0.8″) mode. The ETC was used to estimate the S/N ratio at 4 different wavelengths (310, 380, 440 and 670 nm) corresponding to regions where the transition lines of several important chemical elements (Be, U, Fe-peak elements, Li, etc.) are seen. The radius of the S/N reference area, the number of spectra on the detector and the actual spectral resolution are parameters, and their choice will be explained in the next sections.

Exposure time: We set DIT = 3600 s to get a reference point for the S/N obtained at a given wavelength for a given EMP star (MARCS model and fixed apparent magnitude).

2. Optical detectors are clearly less efficient in the UV than in the optical or in the near-IR, but this is not taken into account in [RD1], which assumes a detector QE of 90% at all optical wavelengths. In fact the ETC works in a slightly different way and assumes directly a constant overall efficiency of the system TEL+INS+DET equal to 25% (excluding the atmospheric transmission). A first correction to the ETC results was applied so that the overall efficiency matched a more realistic case, and was also dependent on the sensitivity of the detector, and thus on the wavelength. This was done by applying a correction factor to the observed magnitude, issued from the relative variations between this constant overall efficiency and real overall efficiencies of well-known current TEL+INS+DET systems, and will be explained in the next sections.
3. Since the mirror coating is not a parameter in the ETC, which assumes only a classical aluminium coating, an additional correction was applied to its results to simulate a protected silver/aluminium coating. The method used here is similar to the one described above for the detector efficiency, meaning that we chose to apply a correction factor to the observed target's apparent magnitude. This was done simply by comparing the transmission curves for both coatings taken from [RD1], and this will also be explained in more detail in the next sections.
4. Using the (corrected) result of the ETC, we computed the variation of the S/N ratio as a function of the target's apparent magnitude for different types of EMP stars. We extended the simulations to different telescope diameters, observing sites and the two types of coating.
5. Using the previous results for a single 3600 s exposure, we computed the exposure times needed to reach the optimal S/N ratios required to answer each part of this science case (see previous sections) for different EMP stars, depending on the wavelength of the observed lines, the telescope diameter, the site, and the coating type.

13.4.3 Inputs

13.4.3.1 Scientific data

The stellar parameters that we used as input for the spectral synthesis and the ETC simulations are summarised in Table 13.2. These were chosen using typical values seen in all large EMP stars studies in the Milky Way, and match what is expected – and already found on average – in nearby dwarf spheroidals (see references in Section 13.1). We call our 3 EMP stars simply Giant I, Giant II, and Dwarf.

Since the raw MARCS stellar models produce by default only solar-like abundance patterns, several corrections had to be considered when running the spectral synthesis codes. For example, we set the typical $[\alpha/\mathrm{Fe}]$ enhancement to +0.4 dex. We also set more realistic (predicted or commonly observed) abundance ratios for several crucial elements, such as C, N, O (in the 3 EMP stars Giants I, II and Dwarf), Be, Li (in the EMP Dwarf), and added r-process enhancement when we had to simulate an r-enriched EMP star (Giant I) where we want to be able to detect elements such as U or Th.

- The reference star (CS 31082–001), was analysed by Hill et al. (2002) with the same abundance and spectral synthesis codes that we used here. Fig. 13.4 shows a comparison of the Uranium line

Table 13.2: Stellar parameters used to reproduce the spectra of typical EMP stars. The parameters of the Uranium star CS 31082-001 are also shown here for comparison.

EMP Stars	T_{eff} [K]	$\log g$	[Fe/H]	v_{t} [km s^{-1}]	[α/Fe]	[r–process/Fe]
Giant I	4550	1.0	-3.0	2.0	$+0.4$	$\gg$ +1.0
CS 31082-001	4825	1.5	-2.9	1.8	$+0.4$	+1.5–1.7
Giant II	5050	1.8	-4.0	2.0	$+0.4$	–
Dwarf	6250	4.5	-3.0	1.5	$+0.4$	–

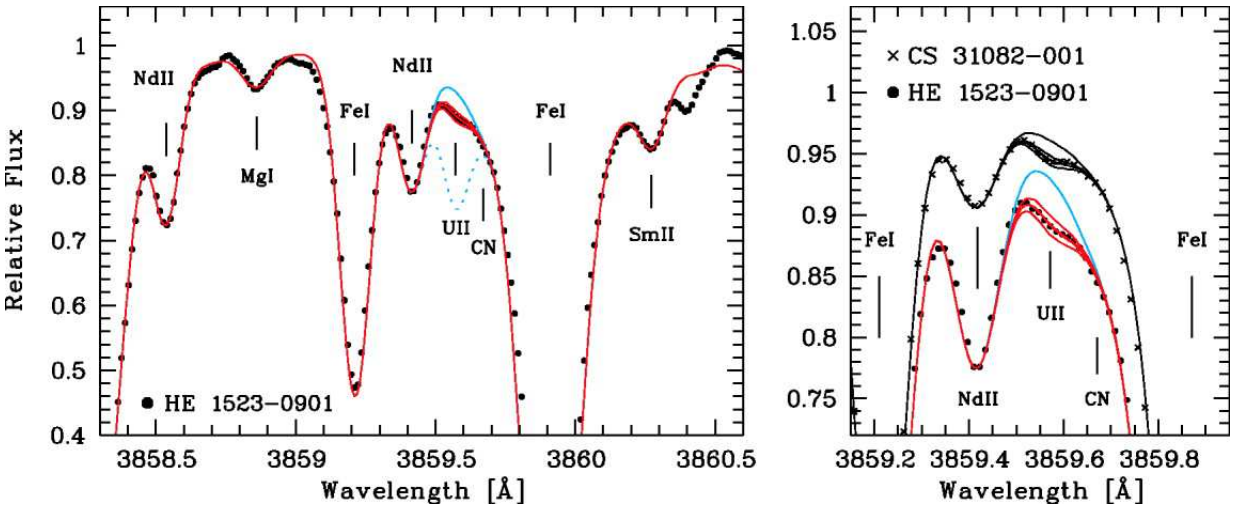

Figure 13.4: Comparing the strength of the Uranium line at 385.96 nm in the only two stars where a detection has been confirmed up to now (Frebel et al. 2007).

in the two well-known Uranium stars HE 1523–0901 and CS 31082–001. These stars were used as references to check wether different levels of r-process enhancements in our models were realistic.

- The work of Rich & Boesgaard (2009) combined with predictions from Smiljanic et al. (2009, and references therein) were used to correct the models for typical Beryllium abundances in EMP dwarfs.
- Several papers from the series "First Stars" were used to adjust the models for typical C, N, O abundances in the simulated EMP giants (Cayrel et al. 2004; Spite et al. 2005) and dwarfs (Bonifacio et al. 2009).
- The Lithium abundance was set at the value of the Lithium Plateau (Spite & Spite 1982a,b) as seen in Galactic EMP dwarfs (see, e.g., Bonifacio et al. 2007).

13.4.3.2 Technical data

The technical data used for these simulations were mainly taken from [RD1], to which we have added a few extensions.

Site and background: As described in [RD1]. We make use of the background emission and atmospheric transmission models for both the Paranal-like and High & Dry sites. We assume an airmass of 1.15 (i.e. a zenith distance of 30°).

Telescope: As described in [RD1]. We also consider primary mirror diameters of 30 and 60 m. We use both the bare aluminium and protected silver/aluminium coatings. Since the ETC uses only the classical Al coating, it was not possible by default to run simulations for the protected Ag/Al coating. We found a simple solution considering that changing the coating has an obvious direct influence on the number of collected photons: if the ETC returns a S/N ratio for a given magnitude V_0 using the classical Al coating, the same S/N would be reached for a star of apparent magnitude $V_0 + \Delta\mathrm{mag}(\tau)$ if the protected Ag/Al coating was used, where

$$\Delta\mathrm{mag}(\tau) = -2.5 \log \left(\frac{\tau_{\mathrm{Al}}}{\tau_{\mathrm{Ag/Al}}} \right). \tag{13.1}$$

τ_{Al} and $\tau_{\mathrm{Ag/Al}}$ are the throughputs of the corresponding coatings after 5 reflections. Using [RD1], we computed the corrections summarised in Table 13.3. This correction procedure is only valid as long as we are operating in the regime where the noise is dominated by target photon noise, which is, however, true for all simulations considered here.

Table 13.3: Corrections to the reference magnitude due to the effect of changing the coating from bare Al to protected Ag/Al.

λ [nm]	310	380	440	670
$\tau_{\rm Al}$	0.662	0.678	0.680	0.583
$\tau_{\rm Ag/Al}$	0.001	0.751	0.852	0.861
$\Delta{\rm mag}(\tau)$	-7.05	+0.11	+0.24	+0.42

Instrument: We need a high-resolution spectrograph (HRS) that goes if possible all the way to blue and even UV wavelengths, while performing observations of single objects with a high efficiency. The optimal resolution should be at least 70 000 in order to detect the most sensitive features described in the science case. A large wavelength coverage in a single shot would also clearly be an advantage. There is no need for special requirements on the spatial resolution (seeing limited observations with $\mathrm{FWHM} = 0.8''$ are sufficient).

Too high a spectral resolution can be an issue in terms of luminosity: the higher the resolution, the more light we might loose in the instrument (i.e. there is a risk of lowering the S/N for a given exposure time). Since this can easily be fixed by binning the final products (or in a binned read-out mode during the exposures), we can continue with a resolution as high as we want. In fact resolutions higher than 70 000 might open the door to other science cases, such as studying isotopic ratios of various chemical elements in other galaxies. In other words, the higher the resolution, the better, because we can still recover some S/N. On the contrary, too low a resolution would simply not work for this science case. Hence we decided to use the HRS with a resolution of R = 100 000 for our simulations in order to match the highest resolution available in the ETC for the sky spectra.

We will assume that our HRS is equipped to reach the UV atmospheric cut-off at around 300 nm.

Using the current design of the CODEX instrument concept study as a guideline, we use $R_{\rm ref} = 400$ mas for the radius of the circular S/N reference area in the ETC, and $N_{\rm spec} = 280$ for the number of individual spectra on the detector over which the light from the S/N reference area is spread (see also the description of the E-ELT Spectroscopic ETC in [RD1]).

The QE assumed in the ETC and in [RD1] is 90% at all wavelengths for an optical detector. Moreover the ETC uses directly a constant overall efficiency of the system TEL+INS+DET of 25% for all wavelengths (excluding the atmospheric transmission). This is clearly not realistic and will immensely underestimate exposure times in the UV and blue wavelengths. Even if combinations of an E-ELT, an efficient spectrograph and future generations of detectors could reach a maximum overall efficiency of 25% at redder wavelengths, it is very unlikely that this value can even be approached in the UV. Since reaching UV and blue wavelengths (at least the bluer part of the optical domain) is crucial for this science case, we could not continue using such simple assumptions and had to apply a correction factor depending on the analysed wavelength range.

Among the currently available high-resolution spectrographs on 8–10 m class telescopes, UVES is certainly one of the most efficient and luminous. We rescaled the UVES overall efficiency curve so that the peak of the curve reached 25%. Since a CODEX-like spectrograph will be most likely fibre-fed, even this assumption might be a bit too optimistic despite all possible improvements in optical detectors in the next 10 years. Nevertheless, we decided to choose this approach rather than the simply too optimistic results given by the original ETC in the blue and UV. The rescaled UVES efficiency curve is used in a similar way as described above to correct the ETC results. If the original ETC gives a S/N for an input magnitude V_0, the "real" magnitude that would give the same S/N assuming a more realistic overall efficiency is in fact $V_0 + \Delta{\rm mag}(\eta)$, where

$$\Delta{\rm mag}(\eta) = -2.5 \log \left(\frac{\eta_{\rm H}}{\eta_{\rm U,n}} \right). \tag{13.2}$$

$\eta_{\rm H}$ is the total efficiency of the system E-ELT+HRS+DET assumed in the ETC (= 25%) and $\eta_{\rm U,n}$ is the total efficiency of the system VLT+UVES+DET rescaled so that its peak value is 25%. Table 13.4 lists the corrections we have computed. Again, this correction procedure is only valid as long as we are operating in the regime where target photon noise dominates.

Table 13.4: Corrections to the reference magnitude due to a more "realistic" overall system efficiency.

λ [nm]	310	380	440	670
$\eta_{\rm H}$	0.250	0.250	0.250	0.250
$\eta_{\rm U}$	0.060	0.105	0.110	0.155
$\eta_{\rm U,n}$	0.091	0.159	0.167	0.235
$\Delta{\rm mag}(\eta)$	-1.10	-0.49	-0.44	-0.07

Adaptive Optics and PSFs: There are no special requirements on the spatial resolution since EMP stars are typically expected to be isolated point-like objects easily identified and observed with a single-object spectrograph. Hence we will assume that our HRS will be operated in seeing-limited conditions (recall also our assumption of a 0.8″ diameter entrance aperture above), and we use the noAO PSFs of [RD1], which were simulated for 0.8″ seeing.

Efficiency/throughput curves for the different parameters discussed above are reproduced in Figs. 13.5 and 13.6, where we show: the total atmospheric transmission (Fig. 13.5 shows the sky transmission of the Paranal-like site, and Fig. 13.6 shows the High & Dry site), the telescope throughput (both coatings), some of the instrument characteristics (flux losses at the entrance, instrument efficiency), a detector QE, and also resulting S/N spectra for the complete model. In this figures, we also indicate as grey shaded areas the three wavelength ranges considered in this science case: 310–370 nm for OH and NH bands, some r-process elements and Beryllium; 370–520 nm for Fe, Fe-peak elements, α-elements, more r-process elements (including U and Th); 670 nm for the Li doublet.

13.4.4 Outputs

The outputs of our simulations are:

- A set of simulated (reduced) spectra for different types of EMP stars (Giants I, II and Dwarf, see Table 13.2) at the resolution of the HRS ($R = 100\,000$) and with S/N ratios high enough to detect the most important features needed to check the feasibility of this science case at a given wavelength. The spectra are limited to just a few Å around these features, and they are simple ASCII tables listing pairs of wavelength and normalised relative intensity values.
- For each of our EMP stars, a prediction of the variation of the S/N as a function of the apparent magnitude for an exposure time of 3600 s using the ELT+HRS. The S/N ratios are given at 4 different wavelengths where important absorption lines for this science case are expected.
- Also for each of our EMP stars, a prediction of the exposure time needed to reach the optimal S/N ratios at the important wavelengths for this science case, as a function of the apparent magnitude.

13.5 Results of simulations

13.5.1 Simulation runs

Simulation of final products and synthetic spectral line fitting: We simulated observed spectra for our 3 EMP stars (atmospheric parameters listed in Table 13.2 of Section 13.4.3.1) following the procedure described in Section 13.4.2. The abundances of particular elements like C, N, O, Li and Be, as well as typical α enhancements for Halo stars, and r-process enrichment (whenever necessary), were adjusted to match observations/predictions in EMP stars (see also Section 13.4.3.1 for more details). The spectral regions that we analysed consisted always of a few Å around the lines listed in Table 13.5. The spectral resolution was set at $R = 100\,000$, with a resolution element spread over 4 pixels, so that we had to slightly adjust the step in λ depending on the spectral region.

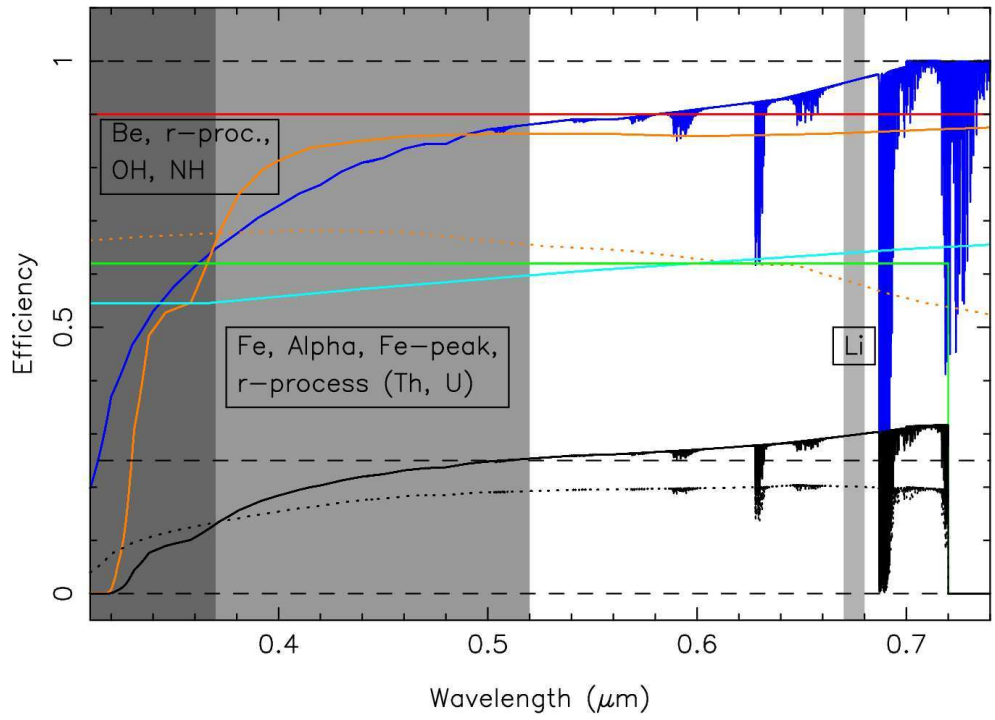

Figure 13.5: Transmission curves for the E-ELT. Blue: total atmospheric transmission in the optical for a Paranal-like site (airmass 1.15, $R = 100\,000$); orange: telescope reflectivity (solid: Ag/Al coating; dotted: Al coating); cyan: flux loss at the entrance of the instrument; green: instrument efficiency (excluding entry losses and detector QE); red: detector QE; black: resulting S/N spectra at $R = 100\,000$ (solid: Ag/Al coating; dotted: Al coating). (This figure was produced using a code written by J. Liske.)

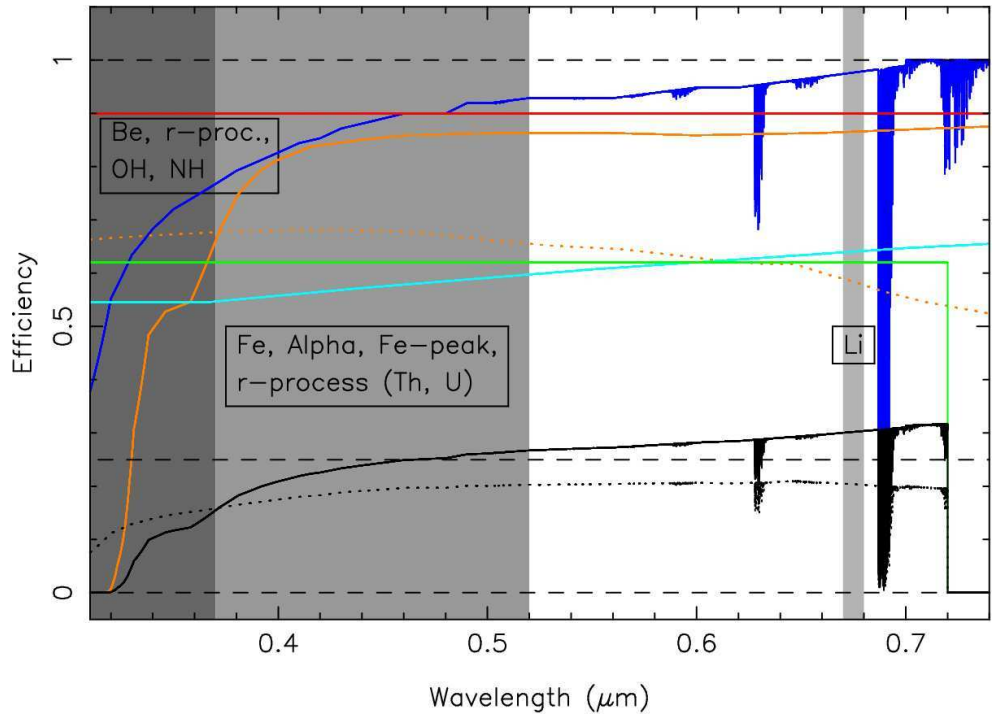

Figure 13.6: Same as Fig. 13.5, but for the High & Dry site.

Table 13.5: Absorption lines of several important elements for the abundance analysis of EMP stars, with indication of the spectral resolution and signal-to-noise ratios that were used for the simulations.

Element	λ [Å]	$R = \lambda/\Delta\lambda$	S/N (@ λ)	Type
OH bands	∼3100	100 000	50	Giant II
Be II	3130.420, 3131.065	100 000	150	Dwarf
U II	3859.571	100 000	200	Giant I
Li I	6707.761, 6707.912	100 000	60	Dwarf

Simulations using the E-ELT Spectroscopic ETC: The simulations were run for flux distributions corresponding to our 3 EMP stars following the procedure described in Section 13.4.2 (see also Sections 13.4.3.1). Each star was given an arbitrary reference apparent magnitude (e.g., $B_0 = 15$ or $V_0 = 15$) to start the ETC. This value is not particularly important, as it is simply a reference point, and this result can easily be extended to any other magnitude using well-known relations between S/N, flux (or apparent magnitude), exposure time and telescope diameter. Our targets were obviously considered as point sources.

The telescope diameter was a fixed parameter ($D_0 = 42$ m), knowing that we could extend the results to the other 2 diameters (30 and 60 m) as well (see Section 13.5.2 hereafter) without re-running the ETC. The coating was a variable parameter (see Section 13.4.3.2 for more explanations on how we corrected the results for the change of coating from Al to Ag/Al). The observatory site was also a variable parameter. The seeing was set at 0.8″, and the airmass at 1.15 (zenith distance of 30deg).

For the instrument, the AO mode was set to seeing limited ($\mathrm{FWHM} = 0.8''$). The observing wavelength λ_{ref} was a variable parameter. The radius of the circular S/N reference area, the number of spectra on the detector and the spectral resolution were set at $R_{\mathrm{ref}} = 400$ mas, $N_{\mathrm{spec}} = 280$, and $R = 100\,000$ (see Section 13.4.3.2 for more details).

Table 13.6 lists the different simulation runs.

Table 13.6: ETC runs for our EMP stars.

Target	Telescope diameter [m]	Observatory site	Coating	λ_{ref} [nm]
Giant I	30, 42, 60	Paranal-like	Al	310, 380, 440, 670
	30, 42, 60	High & Dry	Al	
	42	Paranal-like	Ag/Al	
Giant II	30, 42, 60	Paranal-like	Al	310, 380, 440, 670
	30, 42, 60	High & Dry	Al	
	42	Paranal-like	Ag/Al	
Dwarf	30, 42, 60	Paranal-like	Al	310, 380, 440, 670
	30, 42, 60	High & Dry	Al	
	42	Paranal-like	Ag/Al	

13.5.2 Analysis

13.5.2.1 Final data products and synthetic spectral line fitting

The spectra were analysed using classical spectra line synthesis as explained in Section 13.4.2. Of course, the analysis is not thorough since we had set the atmospheric parameters and forced some of the abundance ratios from the beginning, but the idea was just to show in a visual way the kind of accuracy that one could get with sufficient spectral resolution and S/N. Figs. 13.7, 13.8 and 13.9 show typical regions of interest for this science case in different types of EMP stars. The dots represent the

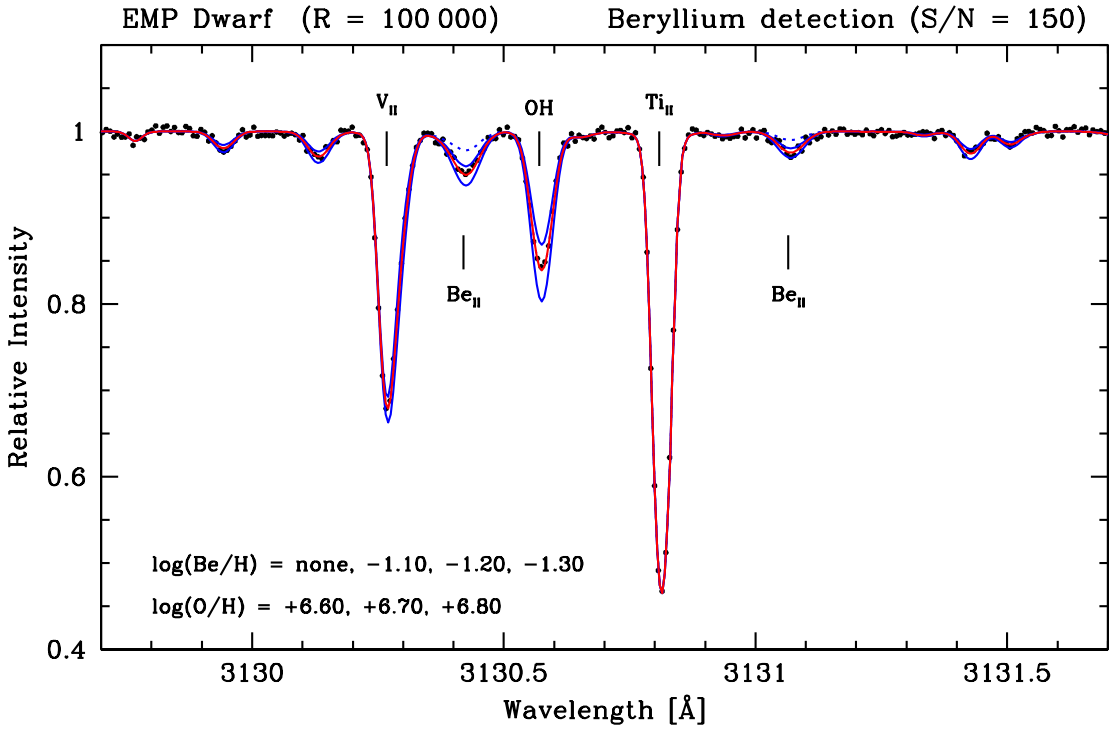

Figure 13.7: Simulated observed spectrum of the EMP "Dwarf" star in the region of the Beryllium lines and several OH bands.

simulated spectra observed at the resolution of the HRS ($R = 100\,000$) and with the S/N ratios indicated in Table 13.5.

In Fig. 13.7 we show a spectrum of the EMP "Dwarf" in the UV. The spectrum has a very high S/N of 150 in order to get a contrast high enough for the 2 weak Beryllium lines at 3130.420 and 3131.065 Å. Several OH bands are also clearly visible in this region. With this high quality spectrum, we can see that the accuracy on the Beryllium abundance is of the order of $\sigma_{\log(\mathrm{Be/H})} \sim 0.08$ dex or better, while the main visible OH line gives an even better result with $\sigma_{\log(\mathrm{O/H})} \sim 0.05$ dex. For the OH bands, there is no need of such high resolution and S/N, but this is in any case a byproduct that will come if one meets the conditions for a good observation of the Beryllium lines.

Fig. 13.8 shows a blue region of the EMP "Giant I" around the Uranium line at 3859.571 Å. This particular line can only be detected provided there is sufficient r-process enrichment ($[r/\mathrm{Fe}] \gg +1$ dex, typically +1.5–1.7 dex) as well as enough contrast compared to the CN band right next to it, meaning the star should not be too Carbon-rich or too hot: e.g., with $T_{\mathrm{eff}} > 5000$ K it becomes already hard to distinguish the line even in the case of an r-process enrichment. This spectrum needed thus to have an excellent quality. Hence we simulated $S/N = 200$, which appears to be more or less the minimum requirement for a really good measurement of the Uranium abundance. For this star, we did not simply play with the Uranium abundance, but more in general with the r-process enrichment, which is the key parameter for this detection. As a consequence, we see the variations of other r-process elements with transitions close to the Uranium line. Already at $S/N = 200$, the uncertainty is of the order of $\sigma_{\log(\mathrm{U/H})} \leq 0.08$ dex (no variations in the atmospheric parameters taken into account). This is good, even slightly better than what is seen today, but still not sufficient for the accuracy desired for cosmochronometry techniques (see Section 13.2): an uncertainty of no more than 0.2 Gyr on the age of a star requires $\sigma_{\log(\mathrm{U/H})} \sim 0.03$ dex, which translates to a S/N close to 1000 (at the resolution of the HRS in these simulations). This is "of course" only a matter of exposure time, meaning that with the E-ELT we will be able to derive the age of our Galactic halo with all the accuracy we want provided we aim at stars which are bright enough ($V \sim$ 12–13 mag). In fact the two currently known Uranium stars could be re-analysed in less than an hour of exposure time each and would yield this fantastic S/N, but it is unlikely that we will extend such

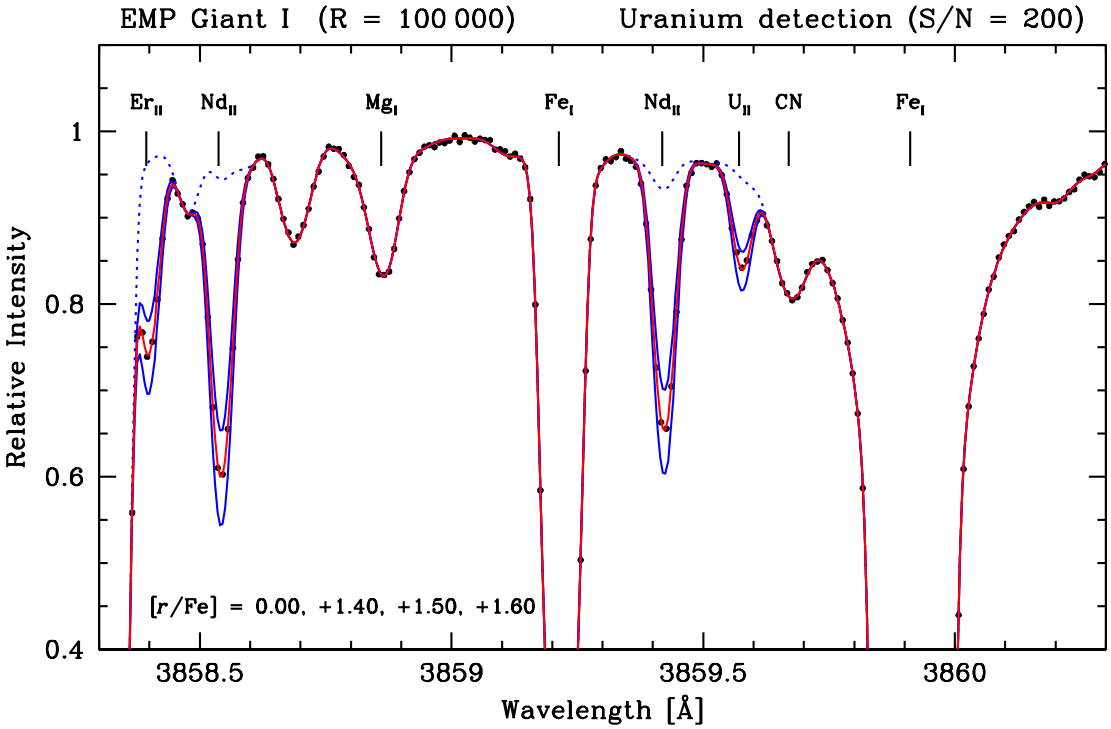

Figure 13.8: Simulated observed spectrum of the EMP "Giant I" star in the region of the Uranium line at 385.96 nm.

extremely accurate determinations to even very close satellites, where the brightest EMP giants have already magnitudes of $V \sim$ 14–16 mag (reaching $S/N = 1000$ for an EMP giant with $V = 15$ mag would require close to 15 hours of exposure time, as can be computed, e.g., from the results shown in Fig. 13.10, see below). The good point, as we will see later in this section, is that at least several Uranium detections and already very good measurements should become possible in some of the nearby satellites of the Milky Way.

In Fig. 13.9 we can see the red part of the spectrum of the EMP "Dwarf" in the region of the Lithium doublet (lines at 6707.761 and 6707.912 Å). The lines are isolated, hence there is no need of a very high S/N to reach an uncertainty of $\sigma_{\log(\mathrm{Li/H})} \sim 0.1$ dex. Here we produced a spectrum with $S/N = 60$ in this region, just to show the minimal requirements to get a good accuracy. However, a good determination of the atmospheric parameters needs a better S/N already in the bluer parts of the spectrum (see Sections 13.2 and 13.3), and since future spectrographs will probably take the full visual spectrum in one shot, the Lithium abundance will be determined using spectra of much higher quality than this one.

In a real future observation, we could expect a total uncertainty of $\pm$0.1 dex or better on all these abundance ratios provided that the S/N ratios given in the previous sections are reached, and if the atmospheric parameters are know with enough accuracy (e.g., $\sigma_{T_{\rm eff}} \leq 50$ K, $\sigma_{\log g} \leq 0.1$ and $\sigma_{v_{\rm t}} \leq 0.1$ km s^{-1}). Of course, this depends also on the methods used to determine these parameters – purely spectroscopic or including photometric data – and in the latter case, on the quality of ongoing and future EMP photometric surveys.

13.5.2.2 Variation of the S/N and DIT as a function of the apparent magnitude

As described in Sections 13.4.2 and 13.5.1, for each one of our EMP stars with a given reference magnitude, the ETC returned values for the S/N resulting from a DIT = 3600 s exposure with a 42 m E-ELT using a bare Al coating, combined with the HRS. This was repeated at each wavelength (310, 380, 440 and 670 nm) and for the two observing sites (Paranal-like and High & Dry). First, these ETC results were corrected with more realistic overall efficiencies depending on the wavelength (see Section 13.4.3.2) yield-

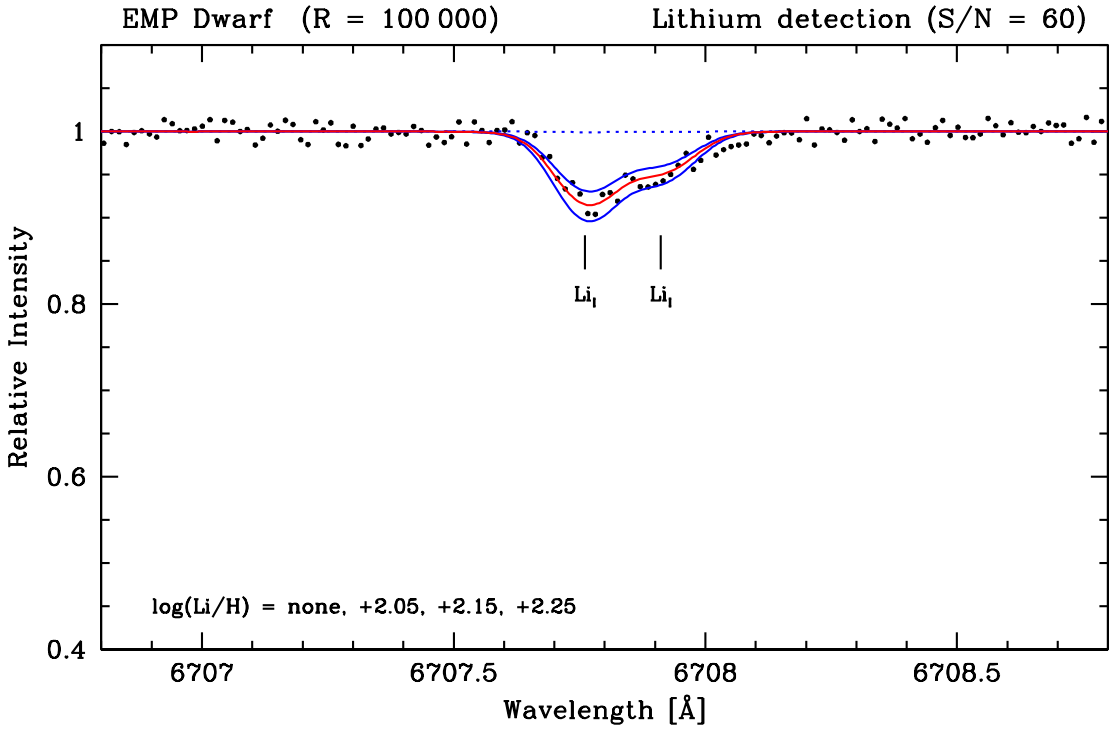

Figure 13.9: Simulated observed spectrum of the EMP "Dwarf" star in the region of the Lithium doublet.

ing a reference signal-to-noise ratio $(S/N)_0$ at the chosen central wavelength $\lambda_{\rm ref}$ for the observation of a V_0 magnitude EMP star, using a $D_0 = 42$ m telescope, with bare Al coating on the selected observing site, and with an exposure time of $t_0 = 3600$ s. Knowing the relations between S/N, flux (or apparent magnitude), exposure time, and telescope diameter

$$S/N \propto \sqrt{\mathcal{F}(m_V)} \times \sqrt{t} \times D \tag{13.3}$$

where

$$\frac{\mathcal{F}(m_V)}{\mathcal{F}(m_{V_0})} = 10^{-(m_V - m_{V_0})/2.5} \tag{13.4}$$

is the flux ratio between an observed star and a reference star of apparent magnitudes m_V and m_{V_0}, we computed new reference points for the other telescope diameters (D = 30 and 60 m), and finally we extended these results to all magnitudes.

The upper panels of Fig. 13.10 to 13.15 show these predicted S/N ratios as a function of the apparent magnitude for our three typical EMP stars (DIT = 3600 s) at different wavelengths. For each telescope diameter, the observing sites are shown as solid lines for the Paranal-like site and as short dashed lines for the High & Dry site. All curves show predictions using the bare Al coating, except the black solid line which stands for a 42 m telescope located on a Paranal-like site and using the Ag/Al coating. Each plot in the upper panels shows an indication of the reference signal-to-noise ratios $(S/N)_{\rm ref}$ needed to detect and analyse the chemical elements that can be found around 310 nm (Be, OH, NH), 380 nm (U), 440 nm (Fe-peak, α elements) and 670 nm (Li doublet).

The relations given above between S/N, DIT, flux and diameter, were used again to predict the variations of the exposure time (DIT) needed to reach the reference value $(S/N)_{\rm ref}$ (i.e. the S/N that allows a successful achievement of the corresponding science case), as a function of the apparent magnitude for a given type of EMP star. These variations are shown in the lower panels of Figs. 13.10 to 13.15.

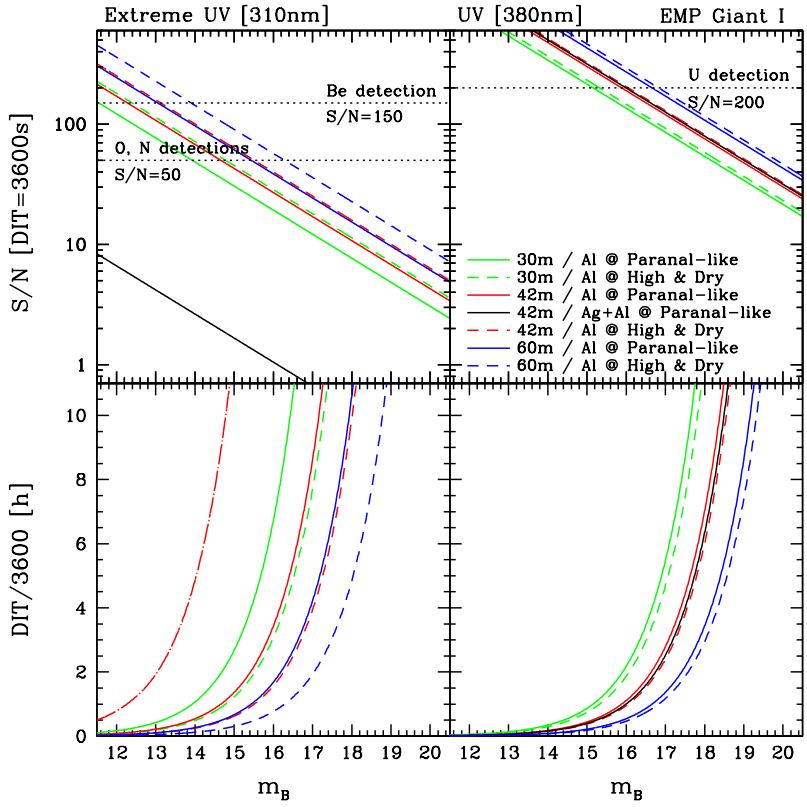

Figure 13.10: Upper panels: predicted S/N ratios for the EMP Giant I as a function of the apparent magnitude (DIT = 3600 s), in the UV near 310 nm and 380 nm. Lower panels: predicted exposure times needed to reach the reference S/N shown in the upper panels as a function of the apparent magnitude. N.B.: In the lower left panel, the red dots/long dashed line shows the DIT needed to reach $S/N = 150$ in the case of a 42 m telescope with a bare Al coating located on a Paranal-like site, while all the other curves show the DIT needed to reach $S/N = 50$ in order to detect O and N. In that panel, the plot for the 42 m using the Ag/Al coating falls out of the plotting area as a result of its extremely poor efficiency at extreme UV wavelengths.

13.5.3 Compliance with figures of merit

A quick inspection of Figs. 13.10 to 13.15 reveals that most objectives will be reached with the E-ELT provided the spectral coverage goes all the way to the extreme UV. A telescope of $D_0 = 42$ m located on a Paranal-like site, using a bare aluminium coating, and equipped with a high-resolution spectrograph like the one used in these simulations is clearly capable of producing the high-quality spectra needed for a successful achievement of this science case. Table 13.7 shows typical predictions that can be made using the previous figures.

Table 13.7: Typically predicted exposure times needed to reach a given S/N ratio at a given reference wavelength for different EMP stars at different apparent magnitudes.

EMP	$(S/N)_{\rm ref}$	$\lambda_{\rm ref}$	DIT for EMP type		
App. magnitude		[nm]	Giant I	Giant II	Dwarf
$B = 15.5$	150	310	> 9 h	≤ 7 h	≤ 3 h
$B = 17.5$	50	310	> 6 h	≤ 5 h	≤ 2 h
$B = 18.0$	200	380	≤ 7 h	≤ 6 h	≤ 5 h
$B = 20.0$	80	440	≤ 2 h 30	≤ 3 h	≤ 3 h
$V = 21.5$	60	670	≤ 3 h	≥ 3 h	≤ 4 h

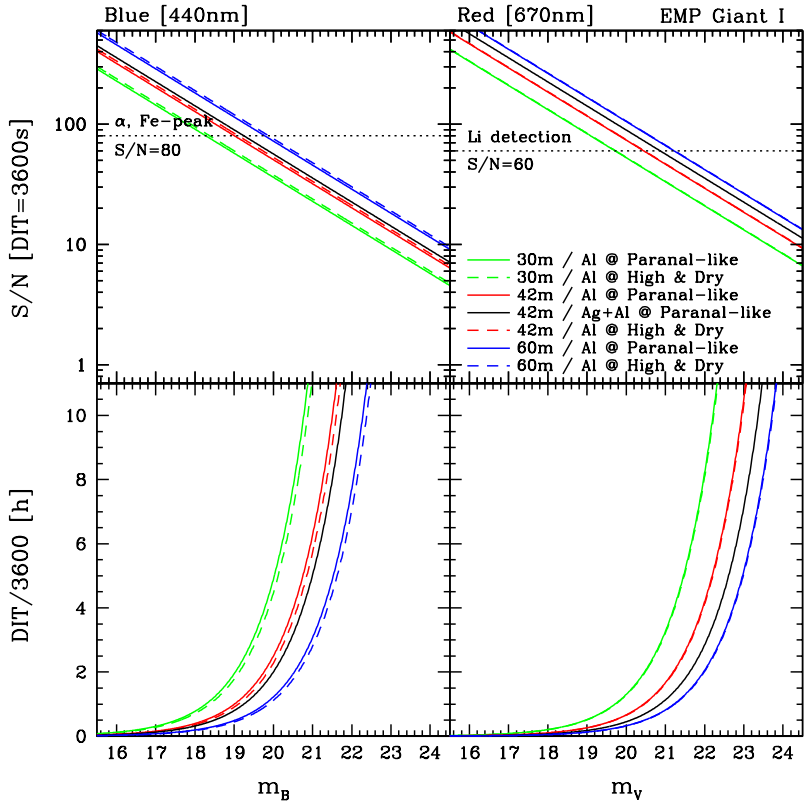

Figure 13.11: Same as Fig. 13.10, but for the regions near 440 nm (Fe-peak and α elements) and 670 nm (Li doublet).

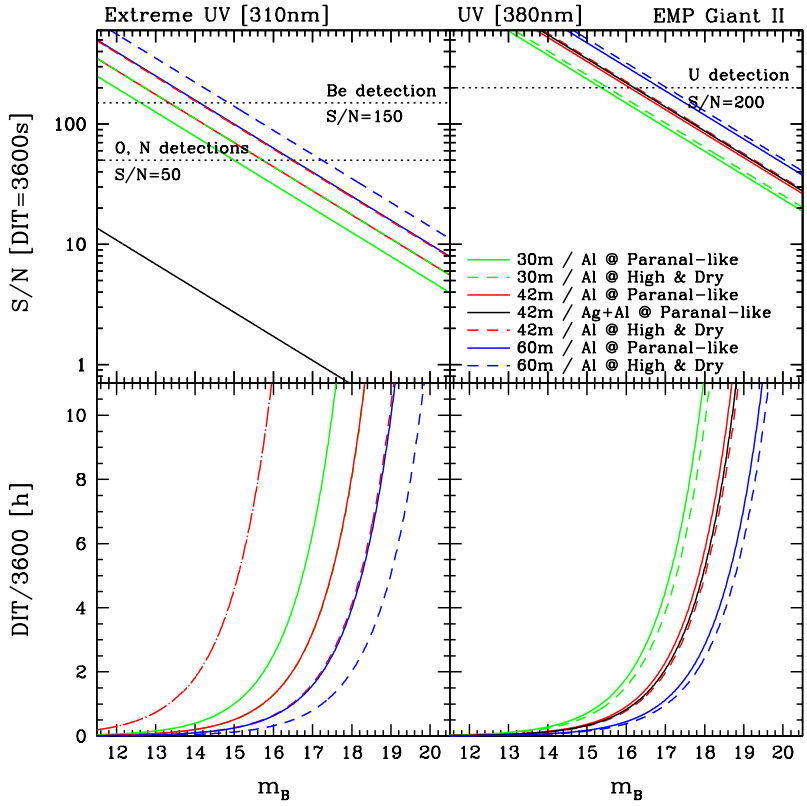

Figure 13.12: Same as Fig. 13.10, but for the case of the EMP Giant II.

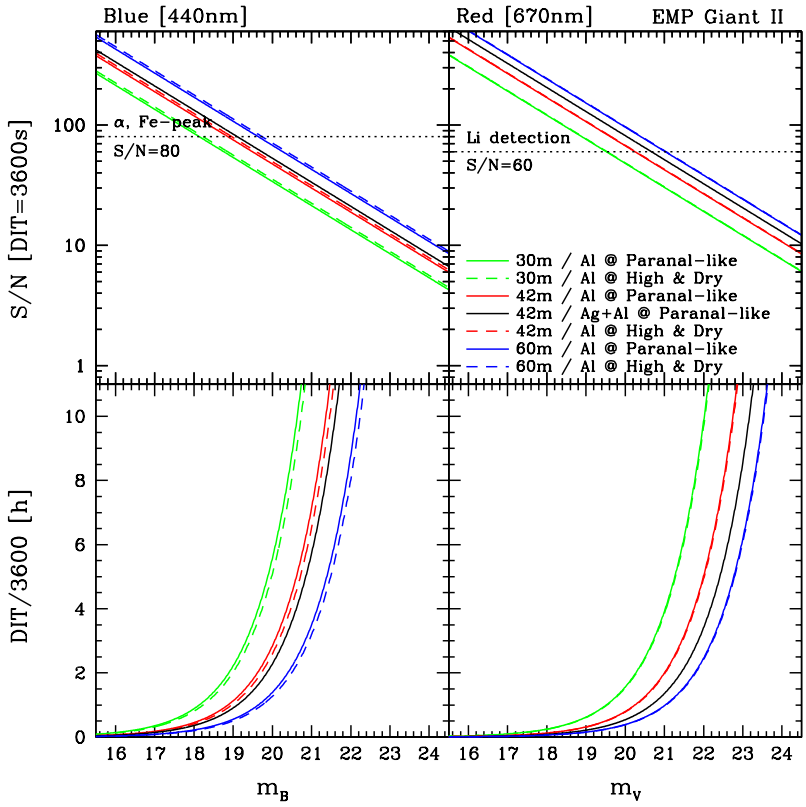

Figure 13.13: Same as Fig. 13.11, but for the case of the EMP Giant II.

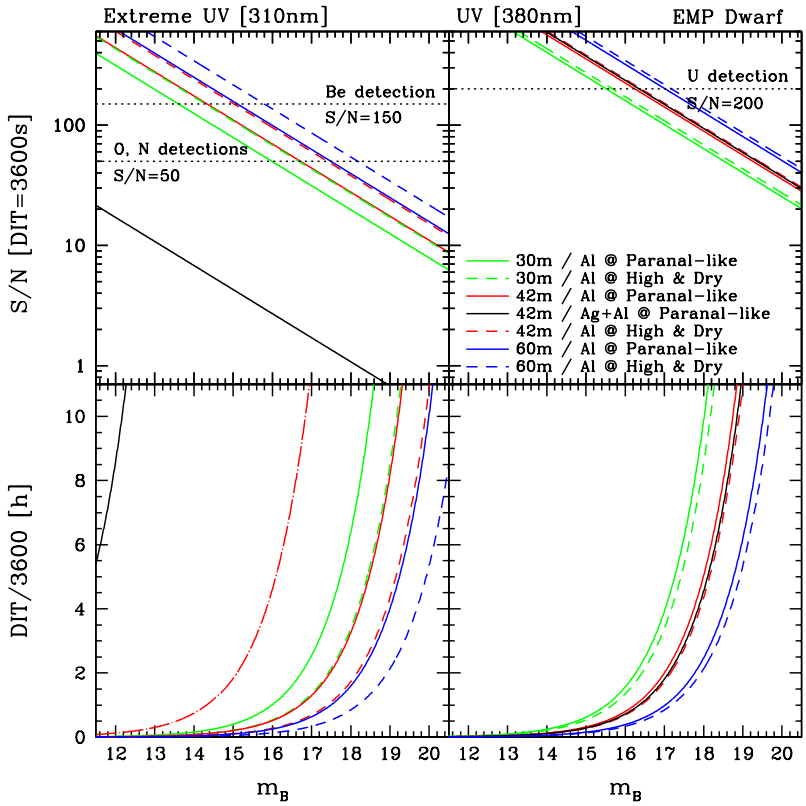

Figure 13.14: Same as Figs. 13.10 and 13.12, but for the case of the EMP Dwarf.

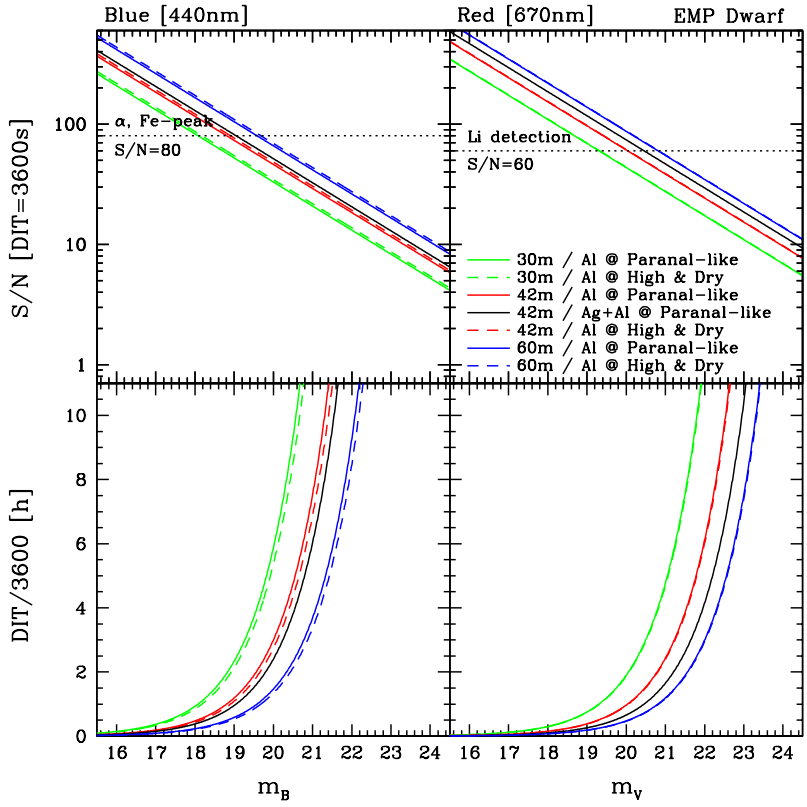

Figure 13.15: Same as Figs. 13.11 and 13.13, but for the case of the EMP Dwarf.

Another way to look at the compliance issue could be this short exercise: let us consider two possible campaigns aimed at observing a total of ∼200 EMP candidate stars scattered across the Milky Way and 4 nearby galaxies (e.g., Sagittarius, LMC, Sculptor and Fornax). The first campaign would concentrate on the distant Halo EMP stars in the Milky Way, while the second campaign would extend the accurate chemical analysis of EMP stars to 4 nearby satellite galaxies. We assume the case of a 42 m E-ELT on a Paranal-like site, using a classical aluminium coating, combined with an $R = 100\,000$ high-resolution spectrograph. What would be the total amount of exposure time (no overheads considered) needed, e.g., to detect Lithium in all our targets, and Uranium in all targets brighter than $V = 18$ mag?

Table 13.8 summarises the typical distances of nearby galaxies, as well as the typical apparent magnitudes expected for our target stars.

Table 13.8: Typical distances and magnitudes of nearby galaxies. For each galaxy, apparent magnitudes of typical EMP stars are given.

Target	d [kpc]	μ^a	Typical V-band apparent magnitudes Bright K Giants ($M_V \sim -3$	Lower-RGB +1	TO Dwarfs +4)
Sagittarius dSph	25	17.0	14.0	18.0	21.0
LMC	55	18.7	15.7	19.7	22.7
SMC	60	18.9	15.9	19.9	22.9
Sculptor dSph	80	19.5	16.5	20.5	23.5
Sextans dSph	90	19.8	16.8	20.8	23.8
Fornax dSph	130	20.6	17.6	21.6	24.6
Leo IV dSph	155	20.9	17.9	21.9	24.9
M31 (Andromeda)	725	24.3	21.3	25.3	28.3

[a] Distance mudulus: $\mu = m - M = -5 + 5 \log d$, where d is the distance in pc.

First of all, we have to keep in mind that whatever the goal, we also need to have accurate atmospheric parameters, and thus sufficient S/N to analyse iron lines in the blue and optical wavelength ranges ($S/N >$ 80–100 above 400 nm). This will be the limiting parameter even if the Lithium detection happens at redder wavelengths and requires a lower S/N than iron and iron-peak elements (and thus a shorter exposure time, as seen in Figs. 13.11, 13.13 and 13.15).

Turn-off stars will obviously not be observable within a reasonable exposure time outside the Milky Way further than the distance of the Sagittarius dwarf spheroidal galaxy.

Here is a hypothetical list of what could be observed:

Milky Way:	50 giants (along the RGB) + 50 dwarfs (all of them with $V \leq 15$ mag) could be observed in ∼100 hours.
Sagittarius:	20 giants ($V =$ 14–18 mag) + 10 TO dwarfs ($V \sim 21.5$ mag) would require close to 200 hours.
LMC:	20 bright giants ($V \sim 15.5$ mag) + 10 lower RGB ($V \sim 19.5$ mag) would need about 50 hours.
Sculptor:	15 bright giants ($V \sim 16.5$ mag) + 10 lower RGB ($V \sim 20$ mag) is feasible in less than 80 hours.
Fornax:	10 bright giants ($V \sim 17.5$ mag) + 5 lower RGB ($V \sim 21.5$ mag) should need less than 100 hours.

According to our results, detections and measurements in the Milky Way could be achieved even for the Uranium and Thorium lines within a total of about 100 hours ($\mathrm{DIT} \leq 1$ h per target), which could be done over a couple of years (in agreement with the length of current Large Programmes at the VLT). In the same way, we can estimate the total amount of time needed for the second campaign to about 400 hours, which is conceivable in a 4–5 years observing campaign. Of course these results might still be too optimistic (see discussion about the overall efficiency), but at least we get an idea of the order of magnitude of the duration of such observing campaigns.

13.5.4 Sensitivity to input parameters

Several input parameters are crucial for the success of this science case. Here we review them in detail:

Diameter: As we mentioned in Section 13.2, the main problem today is the photon collecting power: we simply need bigger than the current generation of large telescopes to get enough S/N for really accurate results in the Milky Way. We also want to go as deep as possible in magnitude and reach several of our nearest nearby satellites with a quality comparable to what we have today inside the Milky Way. The VLT+ESPRESSO could in principle do some of these observations and reach sufficient S/N provided the light path to the common focus was shorter allowing for a UV/blue arm. But this is unlikely. And even then, it would “only” be a 16 m telescope, which pales in comparison with the gigantic collecting power of the E-ELT. Figs. 13.10 to 13.15 show clearly that at all wavelengths, and for all coatings and whatever the observing site, for the largest size, we get the highest S/N, and we can go to deeper magnitudes.

Observing site: In the optical, the site selection has little influence on the evolution of the S/N (and hence exposure time) because the sky transmission does not really change much between the two sites. However, in the extreme UV, the sky transmission plays a very important role and is clearly better at the High & Dry site (see Figs. 13.5 and 13.6). At these very low wavelengths, we can see from Figs. 13.10, 13.12 and 13.14 that moving from a Paranal-like site to the High & Dry site is in fact equivalent to increasing the diameter of the telescope by about 40% (e.g., a 30 m at the High & Dry site is equivalent to a 42 m at the Paranal-like site, and a 42 m at the High & Dry site is equivalent to a 60 m at the Paranal-like site), whereas the influence of the site decreases at longer wavelengths and becomes negligible in the optical (see Figs. 13.11, 13.13 and 13.15).

Mirror coating: Figs. 13.5 and 13.6 clearly showed that the reflectivity is much higher at optical wavelengths when using the Ag/Al coating, but it drops dramatically below 370 nm. Table 13.3 shows the relative gain/loss in terms of magnitudes when going from one coating to the other, and we clearly see

that the Ag/Al coating becomes useless below 370 nm (loss of up to $\sim$ 7 mag in the extreme UV compared to the Al coating). On the other hand, it is obvious that the Ag/Al coating is far better at all optical and red wavelengths than the simple bare Al coating (gain of up to +0.4 mag), but the latter has the advantage of having a much more constant reflectivity at almost all wavelengths from the extreme UV to the end of the optical (see again Figs. 13.5 and 13.6), meaning it is by far the preferred coating if the E-ELT is supposed to be equipped with a HRS that can go all the way to the extreme UV. If this is not the case, and assuming the lower end of the wavelength coverage of the HRS is at $\sim$370 nm (meaning we drop the extreme UV cases, but can still observe Uranium and Thorium in the UV/blue), then the Ag/Al becomes clearly the preferred coating.

Instrument efficiency and wavelength coverage: A part of this science case relies strongly on the fact that the HRS onboard the E-ELT will be capable of working at wavelengths below 370 nm and if possible as extreme as 310 nm. It is obvious in that case that the choice of the detector will be crucial since the QE usually drops dramatically in the UV. Even if next generation detectors are more efficient that current ones, a simple rescaling like we did according to Table 13.4 might prove to be still too optimistic. This has of course less influence at optical and red wavelengths, where most optical detector work at their optimal regime.

Instrument spectral resolution: The resolving power $R = \lambda/\Delta\lambda$ is also an important factor for this science case. It should always be higher than 50 000, if possible reach 70 000 (and more if we want to leave a door opened for other topics). Higher resolutions are not a problem even if we lose some photons because the instrument might be less luminous. There are always ways to recover enough S/N and observe fainter objects. In any case, the resolution should not be lower than 50 000, otherwise too many parts of the science case become not feasible.

13.5.5 Calibration requirements

These are standard observations of individual halo stars in the Milky Way and some nearby satellite galaxies, and the target selection depends on the progress of current and future EMP surveys (see Section 13.1). The stars are expected to be well isolated (see Fig. 13.1) and easily observable with an E-ELT even in nearby dwarf spheroidal galaxies. There are no specific requirements on the spatial resolution: seeing limited (0.8″) observations are assumed in all our simulations.

Since we expect to push today's achievements in EMP abundance analysis to fainter magnitudes / larger distances, the determining factors for this science case are the photon collecting power (telescope size and coating) and the spectral coverage. The techniques for analysing absorption lines, determining atmospheric parameters ($T_{\rm eff}$, $\log g$, [Fe/H] and microturbulent velocity) and elemental abundances are all well-known and further simulations do not require more specific calibrations (see also Section 13.5.6). However, the real observations might need a few calibration stars (spectro-photometric / telluric / radial velocity standards), depending on the scientific goals and the accuracy that will need to be reached.

13.5.6 Limitations

We did not simulate a complete set of observations including data reduction, but preferred to work on final products. As a consequence, we were not confronted with problems such as background noise, sky subtraction or wavelength calibration. Sky subtraction and background noise could have been problematic if we were looking for extremely faint stars and were expecting at the same time very low S/N ratios, but the quality of the spectra needed for this science case are clearly beyond that. Radial velocity standards and/or a good wavelength calibration (hence stability of the instrument) are of course crucial, but will depend directly on future generations of high-resolution spectrographs, and are not an essential parameter here since we are mostly interested in line detection, but not necessarily on very accurate radial velocities. If the HRS is an ultra-stable spectrograph, then highly accurate velocities will be a welcome byproduct.

We ran the ETC simulations only for a seeing of 0.8″ and an airmass of 1.15, but obviously the observing conditions can vary a lot during the year, let alone within a single night, and the airmass of a given

target depends on the moment of the observation and can change a lot during long exposures. Our goal was to see wether the proposal was feasible, and not necessarily to study the impact of the weather conditions on such observations. Concerning the seeing, we simply assumed that the E-ELT will be on a high-quality site, meaning the average yearly seeing will not be worse than the value taken here. We chose to restrain ourselves to nearly optimal observing conditions considering all these assumptions. Of course, the simulations could be repeated for very bad observing conditions (very bad seeing, very high airmass), but if such conditions happened, the observer would simply go to the brightest targets among the observable EMP stars, hence the conclusion that a more thorough analysis of the variations of the S/N as a function of the seeing or the airmass is not required at this stage.

Another point is the sensitivity to the atmospheric parameters. We could redo the simulations using randomly changed parameters in order to have a better determination of the final uncertainties in the abundances, and make robust predictions on these sensitivities. However this depends, e.g., on how one chooses to determine the atmospheric parameters when they are not known (e.g., spectroscopic versus photometric methods for the $T_{\rm eff}$, ionisation equilibrium versus isochrones for the $\log g$). Since this is a quite well-known issue already, and the problem is regularly tackled in papers on abundance determinations, we chose not to focus on that and preferred to show directly the raw quality that we expect from an E-ELT. A thorough analysis of the error budget could also improve the confidence on our simulations, but the impact of resolution and S/N on the abundance analysis is also a well-known parameter. As mentioned above, we preferred to focus on final products and see if and how we could reach the desired quality for a proper abundance analysis, and if yes in which amount of observing time. The idea behind this being to give a good feeling of the (almost) obvious feasibility of this science case.

13.6 Concluding remarks

In the previous sections, we have been able to realize the impact of the photon collecting power on this science case. We have seen how the telescope size, the observing site, the mirror coating, as well as the choice of a performant high-resolution spectrograph equipped with an efficient detector are crucial, in particular for observations in the extreme UV. We have noticed that the current concepts for a CODEX-like spectrograph do not include the possibility to go below 370 nm, ruling out by default the observations of OH and NH bands, but also making it impossible to observe Beryllium, which will not allow further improvements in the understanding of themes such as cosmic ray spallation. It is also interesting to note (and keep in mind) that a classical bare aluminium coating could allow future upgrades of first-generation high-resolution spectrographs not yet equipped to reach UV wavelengths, whereas an Ag/Al coating would be excellent in the visual and near-IR domains, but would be a real problem in the UV and would prevent any further upgrades to reach these wavelengths.

References

Aoki W., et al., 2009, A&A, 502, 569
Belokurov V., et al., 2007, ApJ, 654, 897
Bonifacio P., et al., 2007, A&A, 462, 851
Bonifacio P., et al., 2009, A&A, 501, 519
Cayrel R., 2006, RPPh, 69, 2823
Cayrel R., et al., 2004, A&A, 416, 1117
Christlieb N., Schörck T., Frebel A., Beers T.C., Wisotzki L., Reimers D., 2008, A&A, 484, 721
Cohen J.G., Huang W., 2009, ApJ, 701, 1053
Feltzing S., Eriksson K., Kleyna J., Wilkinson M.I., 2009, A&A, 508, L1
François P., et al., 2007, A&A, 476, 935
Frebel A., Christlieb N., Norris J.E., Thom C., Beers T.C., Rhee J., 2007, ApJ, 660, L117
Frebel A., Simon J.D., Geha M., Willman B., 2010, ApJ, 708, 560
Fulbright J.P., Rich R.M., Castro S., 2004, ApJ, 612, 447
Gratton R.G., Bragaglia A., Carretta E., Clementini G., Desidera S., et al., 2003, A&A, 408, 529
Helmi A., et al., 2006, ApJ, 651, L121

Hill V., et al., 2002, A&A, 387, 560
Koch A., McWilliam A., Grebel E.K., Zucker D.B., Belokurov V., 2008, ApJ, 688, L13
Norris J.E., Yong D., Gilmore G., Wyse R.F.G., 2010, ApJ, 711, 350
Rich J. A., Boesgaard A. M., 2009, ApJ, 701, 1519
Schörck T., et al., 2009, A&A, 507, 817
Simon J.D., Frebel A., McWilliam A., Kirby E.N., Thompson I.B., 2010, ApJ, 716, 446
Smiljanic R., Pasquini L., Bonifacio P., Galli D., Gratton R.G., Randich S., Wolff B., 2009, A&A, 499, 103
Spite M., Spite F., 1982, Natur, 297, 483
Spite F., Spite M., 1982, A&A, 115, 357
Spite M., et al., 2005, A&A, 430, 655

14 G9: A survey of Black Holes in different environments

Authors: A. Küpcü Yoldaş, W. Freudling, M. Cappellari

14.1 The science case

The goal of this DRM case is to obtain spatially resolved spectroscopy of the sphere of influence of suspected black holes (BHs) in elliptical and early-type galaxies to investigate the low and high-mass end of the $M_{\rm BH}$-σ relation. The sample includes galaxies in a broad range of clusters as well as field galaxies. The centres of all galaxies in our sample will first be imaged e.g. with HST and/or JWST, and expected BH masses will be estimated based on indicators such as the cusp brightness and velocity dispersion. The survey intends to resolve for the first time the sphere of influence of BHs with masses $M_{\rm BH}$ of around $10^6\ M_\odot$ outside of the local Universe, and at the same time yield spectroscopic data of a significant sample of the most massive BHs currently known. The survey also might discover supermassive BHs with masses of up to about $10^{10}\ M_\odot$.

14.2 Goals of the DRM simulations

- What is the lower mass limit for the signature of BHs to be detectable at the distance of Virgo?
- What is the upper distance limit for the signature of a supermassive BH to be detectable?
- What are the necessary exposure times?
- What is the most efficient spectral region to observe in?
- What is the impact of the PSF on the detection of the signature of BHs?

14.3 Metrics / figures of merit

The usual way to determine the mass of a SMBH is to model the dynamics based on Schwarzschild (1979) orbit-superposition method (e.g. Cappellari et al. 2009; Nowak et al. 2007). The significance of the detection of SMBH is determined by fitting these models and calculating the confidence levels of the BH mass. This is a long and complex procedure where each galaxy should be separately modelled in detail and fit. Therefore instead we developed an alternative way to determine the significance of the detection of the SMBH.

The simulation should produce realistic observations, which we then analyse with existing programs to measure the kinematics of the galaxy. We do this for two different cases for the same galaxy: with an SMBH and without an SMBH.

The figure of merit to determine the significance of the detection of the BH, is the significance of the difference of kinematics measurements between the case with BH minus the maps of the case without BH.

We apply two different methods to determine the significance:

1. The measured difference in kinematics between the case with BH and without BH is compared to the input difference in kinematics by fitting a straight line (see Fig. 14.1). In the ideal case, the slope should always be one. This fitting is done for multiple realisations (20–100) for each simulation. The significance of the detection of the BH for that simulation is the mean divided by the standard deviation of the fit slopes.

2. The input and the measured difference in kinematics between the case with BH and without BH are examined as a function of the radius from the centre of the galaxy to the outskirts. The significance of the detection of the BH for a simulation is determined by the input difference kinematics divided by

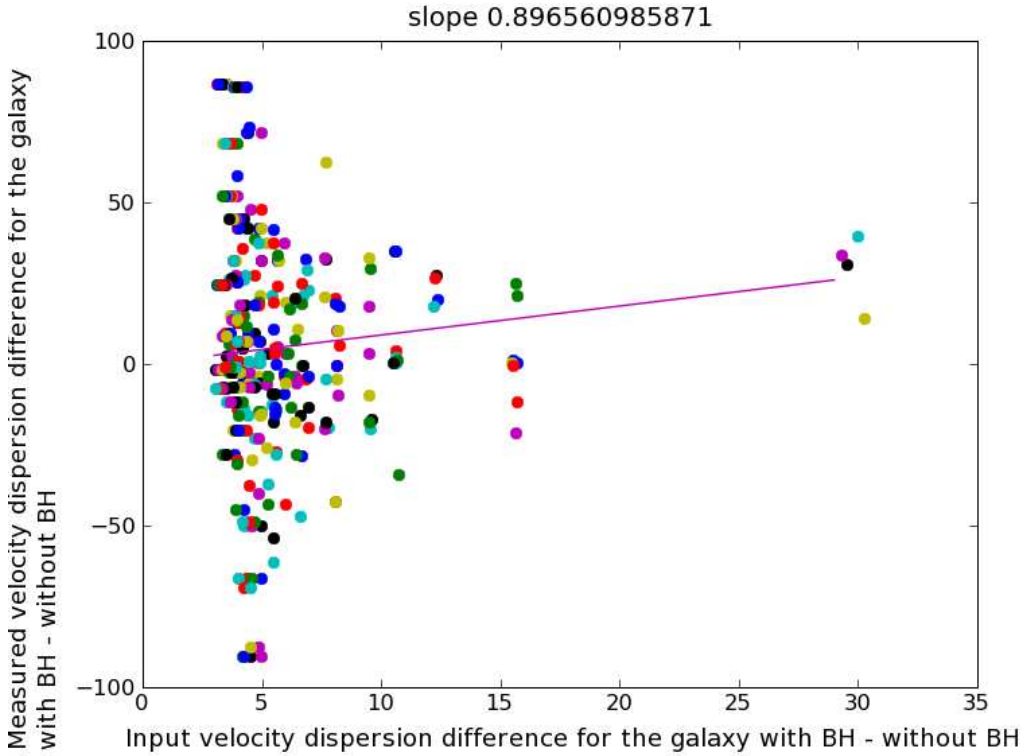

Figure 14.1: An example of fitting the input versus measured velocity dispersion difference with BH minus without BH, for observation simulations of M87 at z=0.1. The best fit slope for this realisation is ∼0.89.

the standard deviation of the measured kinematics at that spaxel, using multiple realisations of that simulation. The spaxels are added starting from the centre towards the outskirts until a maximum significance is reached.

We have calibrate this method with known observations for which the full analysis is available (see Section 14.5.1.1).

14.4 DRM simulations

14.4.1 Methodology

The methodology we follow to study the feasibility of the goals of this science case is to simulate observations (Section 14.4.2) and then to analyse them kinematically (Section 14.5.2).

The existence and mass of the BH affects the kinematics of the central part of the galaxies. Therefore kinematics is the main ingredient in our simulations. Detection of these signatures requires high spatial-resolution and hence the ideal instrument for this kind of observation is an Integral Field Unit (IFU) spectrograph. In order to simulate the IFU observations, we start with a model of a galaxy with a SMBH. The model consists of kinematics, surface brightness distribution and spectrum. This model is on a much finer grid than what we can observe. The galaxy kinematics with a high-resolution spatial scaling is calculated for different BH masses using the Jeans Anisotropic Multi-Gaussian expansion (JAM) dynamical models of stellar kinematics of axisymmetric galaxies (Cappellari 2008; Emsellem et al. 2004a) (see Section 14.4.3.1 for details). These models provide us with the intensity, velocity and velocity dispersion maps which are then convolved with the relevant E-ELT PSF and fed to the pipeline (see Section 14.4.2).

The goals of this study include determining the most efficient spectral region for observations and hence require simulating not only a single line but the full realistic spectrum with medium spectral resolution. This implies the usage of realistic composite stellar spectra as input as well as using realistic atmospheric emission and extinction spectra, and the throughput of the telescope as a function of wavelength. Finally, our simulations take into account the IFU properties such as the spectral resolution and coverage, spaxel size, and noise characteristics. This provides us the "observed" datacube on which we perform the data analysis.

14.4.2 Pipeline

In this section we describe how the observations are simulated. The end products of the pipeline are datacubes in FITS format as in the case of real IFU data.

Each spaxel of the IFU mock data is simulated as a full spectrum with the kinematics and intensity values corresponding to its spatial coordinates in the field-of-view (FoV). Afterwards, these spaxel spectra are combined and written into 3D FITS files.

To be able to provide the kinematics and intensity values at a given spaxel, the input kinematics and intensity maps are first binned to match the spatial pixel scale of the E-ELT PSFs and then convolved with the corresponding PSF using IRAF's FCONVOLVE task. The convolution is luminosity weighted, applied as follows:

$$I_{\rm conv} = I * {\rm PSF} \tag{14.1}$$

$$V_{\rm conv} = \frac{(I \times V) * {\rm PSF}}{I_{\rm conv}} \tag{14.2}$$

$$\sigma_{\rm conv} = \sqrt{\frac{I \times (V^2 + \sigma^2) * {\rm PSF}}{I_{\rm conv}} - V_{\rm conv}^2} \tag{14.3}$$

where I is the stellar surface brightness, V the velocity and σ the velocity dispersion. "conv" denotes the convolved values of these three parameters and "* PSF" denotes the convolution with the PSF. The PSF convolved maps are then binned to match the spaxel size of the IFU.

We made use of the IRAF SYNPHOT package to simulate the realistic spectra. It is a simulation package originally developed for HST observation simulations. Nevertheless, most of its tasks allow generic usage which makes the SYNPHOT package ideal for our purposes.

The spectrum at each spaxel initially consists of the raw input science spectrum (synthetic composite stellar spectra). We assume that the galaxy is symmetric in the sense that the input raw science spectrum in each spaxel is identical. We also assume that the line-of-sight velocity distribution of the stars is well approximated by a Gaussian, i.e. the higher order coefficients of the Gauss-Hermite polynomial are zero. Therefore the input spectrum is convolved with a Gaussian whose sigma is the velocity dispersion of the stars at that spaxel, using the IRAF GAUSS task. The convolved spectrum is then shifted to the velocity of the stars at that spaxel and redshifted to the desired redshift, using the IRAF CALCSPEC task. Afterwards, it is normalised to the intensity at that spaxel using the value from the input surface brightness maps and converted to apparent flux with the relevant luminosity distance. The resultant spectrum is the incoming spectrum of the galaxy at that spaxel.

To have realistic simulations we need to add the effect of the atmosphere, telescope and the instrument. The atmospheric emission spectrum and the OH lines are first convolved with a Gaussian to match the desired spectral resolution. Then, the incoming spectrum flux is summed with the atmospheric emission spectrum, the OH lines and the constant atmospheric continuum and the atmospheric extinction is applied. After these the telescope throughput and the quantum efficiency of the detector is applied. The flux of the resultant spectrum is then converted into counts using the collective area of the E-ELT and multiplied with the exposure time. All of these steps are done in a single IRAF CALCSPEC task. The collective area of the telescope is set accordingly at the IRAF REFDATA package.

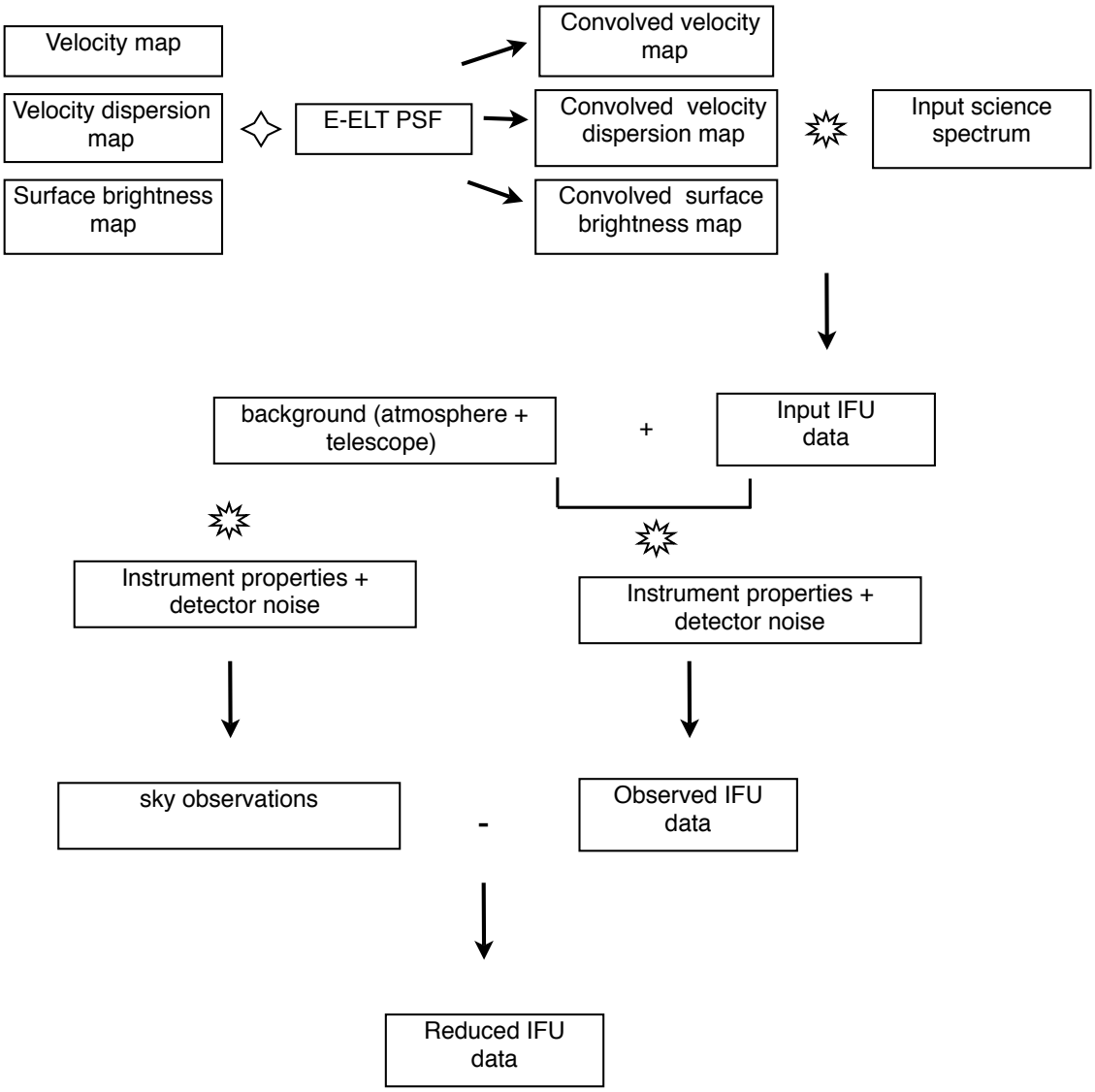

Figure 14.2: The flow-chart of the observation simulation pipeline.

Finally the dark current and read-out noise of the detector and the Poisson noise are added to obtain a realistic observed spectrum of a single exposure at a spaxel, using the IRAF MKNOISE task, as follows:

$$\text{Output spectrum} = P(\text{input spectrum} + \text{dark current}) + G(\text{RON}) \tag{14.4}$$

where $P(x)$ is the Poisson deviate with mean x, $G(x)$ the Gaussian deviate with mean 0 and standard deviation x, and RON the read-out noise.

Due to the high infrared sky emission, the near-infrared (near-IR) observations are restricted in exposure time in order not to saturate the detector. The usual observation strategy in the near-IR is to acquire multiple observations with short exposure times to reach a high S/N without saturating the detector. The individual observations are sky subtracted using IRAF IMARITH task. For the K-band observation simulations were also divided by the individual spectra with the simulated source observations normalised to one, in order to remove the telluric lines. The individual spectra are then combined using the IRAF SCOMBINE task. We followed the same strategy and produced multiple spectra for a single spaxel. The pipeline generates NDIT datacubes, all with the same individual exposure time. The sky spectra of the same exposure time and are generated separately following the above described procedure and then subtracted from each individual single exposure science spectrum. The sky subtracted (and telluric correction applied) spectra are then combined into a single spectrum.

The procedure described up to now generates a sky subtracted (and telluric absorption corrected) spectrum at a single spaxel. The spectra of the whole FoV are simulated in this way and combined into a single datacube corresponding to a reduced and combined IFU observation. The pipeline produced datacubes do not include any flat-field or wavelength distortion effects as we assume a perfect data reduction.

14.4.3 Inputs

14.4.3.1 Scientific data

The input scientific data consists of two parts: the kinematics and surface brightness maps, and the input stellar spectra.

The kinematics and surface brightness maps were provided by one of the authors, using the Jeans Anisotropic Multi-Gaussian Expansion (JAM) dynamical modelling of stellar kinematics of axisymmetric galaxies (Cappellari 2008; Emsellem et al. 2004a). The well-tested JAM models appear to provide good descriptions (Cappellari 2008; Scott et al. 2009) of the stellar kinematics of real early-type galaxies from state-of-the-art integral-field observations taken by the SAURON project (Emsellem et al. 2004b). For this reason JAM models are a useful tool to simulate the kinematics we need for our tests of BH detection. We used the JAM modelled kinematics and intensity maps of three types of galaxies at different distances:

- NGC 4468a-like dwarf elliptical galaxy with a BH of mass 1.25×10^7 $M_\odot$, at the Virgo distance;
- M87-like giant elliptical galaxy with a BH of mass 3×10^9 $M_\odot$, at redshift of z = 0.05, 0.1 and 0.2;
- NGC 5308-like flat S0 galaxy with a BH of mass 2×10^8 $M_\odot$, at redshift of z = 0.05, 0.1 and 0.2.

The masses of the SMBHs in NGC 4486a and M87 are adapted from the measured values (Nowak et al. 2007; Macchetto et al. 1997), while the mass of the SMBH in NGC 5308 is a realistic estimate from the $M_{\rm BH} - \sigma$ relation. The luminosity-weighted second moment within one effective radius of our three galaxies is: $\sigma_e \sim$ 100 km s^{-1} for NGC 4486a, 200 km s^{-1} for NGC 5308 and 300 km s^{-1} for M87. These three cases are selected to be representative of the full range of σ_e values in early-type galaxies.

We also obtained maps of the same galaxy models with no supermassive BH at the centre, to compare and determine the significance of the detection and to determine the BH masses.

We used synthetic spectra from the Marcs stellar library (Gustafsson et al. 2008) as the input stellar spectrum. The spectra in the Marcs stellar library have a spectral resolution of R$\sim$20 000, hence we first convolved them to match the spectral resolution used in our simulations. The pipeline is designed such that it can combine many spectra with different weights into a single spectrum to be used as input. We used stellar templates of cool giant stars for the simulations since both at near-IR wavelengths (especially

at the CO absorption) and in the 8500 Å Calcium triplet region the light of early-type galaxies is dominated by these evolved giant stars. We used the input spectra for the kinematic analysis fits assuming perfect knowledge of the underlying galaxy spectrum. Thus, the input spectrum does not need to be perfectly realistic but should contain the dominant components of the observed galaxy spectra in the relevant wavelength regions. For simplicity, we mostly used a single spectrum of a cool giant star as the input spectrum.

14.4.3.2 Technical data

The technical data used for these simulations were mainly taken from [RD1].

Site and background: As described in [RD1]. We make use of the background emission and atmospheric transmission models for both the Paranal-like and High & Dry sites.

Telescope: As described in [RD1]. We use both the bare aluminium and protected silver/aluminium coatings.

Instrument: We assumed the instrument to be a single-field IFU with either 5, 10 or 40 mas sized spaxels, a FoV of 0.8″×0.8″, a spectral resolution of 5000, and with a wavelength coverage from I to K, but only single band coverage in a single exposure. We assumed an instrument transmission of 1 and a QE and read-out noise as in [RD1].

Adaptive Optics and PSFs: Since this science case requires very high spatial resolution in order to resolve the sphere of influence of the central BH, we used the Laser Tomography Adaptive Optics (LTAO) PSFs described in [RD1]. For each wavelength we only used the on-axis PSF, assuming that it does not vary significantly across the FoV of 0.8″.

14.4.4 Outputs

The output of the simulation pipeline consists of three FITS files:

- the input template spectrum, i.e. if more than one stellar spectrum is given as input, they are combined with appropriate weights and written to a FITS file;
- the full sky-subtracted datacube (run through all the processes described above);
- the sky-subtracted datacube without dark current, read-out, and photon noise (for this noiseless case the sky spectrum used for subtraction also has no noise).

The FITS headers include the values or the file names of the assumed input scientific and technical data.

14.5 Results of simulations

14.5.1 Simulation runs

14.5.1.1 Test case: SINFONI simulations

We first ran a test case to check whether our simulations match the real observations. For this we used the model of NGC 4486a, a dwarf galaxy at the distance of the Virgo cluster (16 Mpc). As the mass of the BH for JAM maps, we used 1.25×10^7 $M_\odot$ which is the value obtained by Nowak et al. (2007) using observations obtained with SINFONI at the VLT. For this run we used the real SINFONI PSF obtained during the SINFONI observations of NGC 4486a (Nowak; priv. comm.). We simulated the datacube with the same exposure time (14×600 sec), spaxel size (50 mas) and field-of-view (3″×3″) as the real SINFONI observations. We used the site parameters of Paranal and the total efficiency of the telescope + instrument is assumed to be 30%. We combined the 14×600 sec observations to obtain a better S/N. The counts and S/N of the real and simulated data were found to be in agreement.

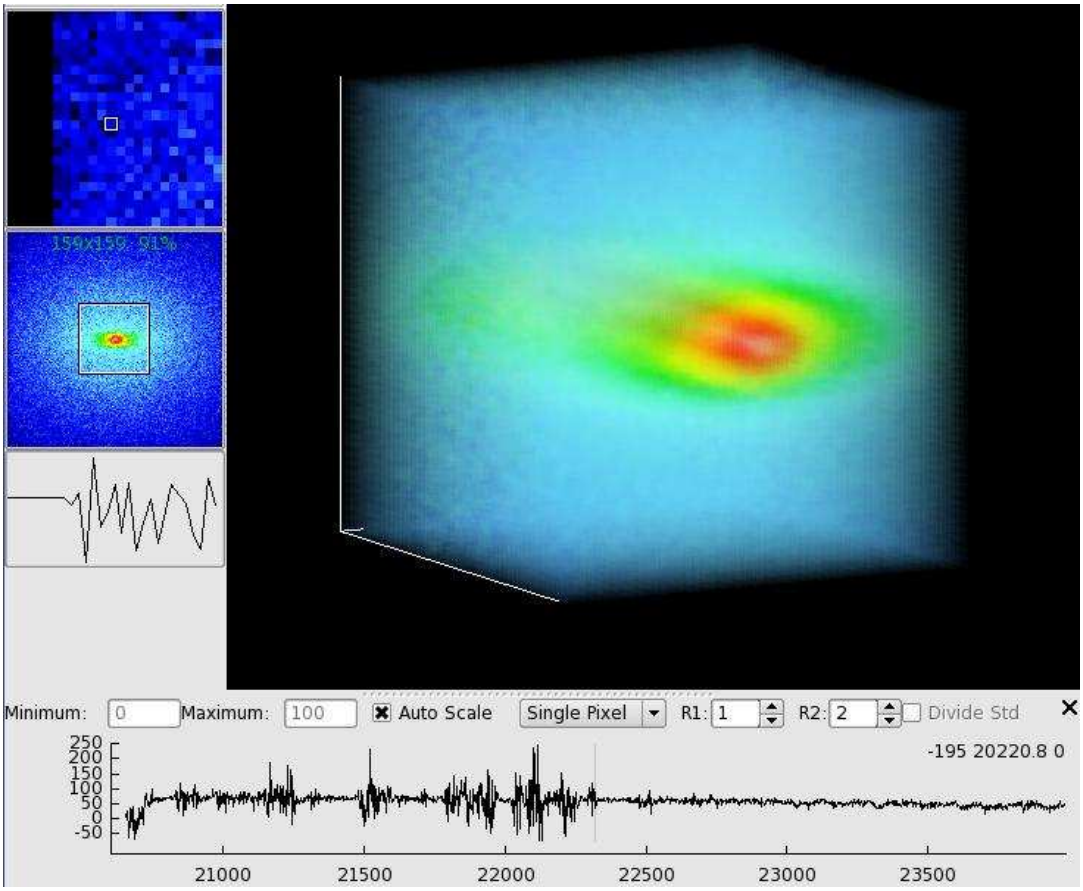

Figure 14.3: The 3D view of one realisation of simulated E-ELT observations of NGC 4486a.

After simulating the SINFONI datacubes we ran the kinematic analysis using pPXF (see Section 14.5.2 for details) around the CO absorption line region in the K-band. The S/N of the combined datacube, with 14$\times$600 = 8400 s exposure time, is not enough to obtain a decent fit. Therefore we binned the spaxels using the Voronoi binning method (Cappellari & Copin 2003) to match S/N$\sim$30 as was also done by Nowak et al. (2007) for the real SINFONI observations. The velocity and velocity dispersion maps of the real and simulated observations are shown in Figs. 14.4 and 14.5 for comparison. The general pattern of the maps are the same for the real and simulated observations. The lowest and highest values of the velocity and velocity dispersion are slightly larger for the real SINFONI observations. This is most likely due to the fact that the JAM model of NGC 4486a was not a best fit to the observations. The exact velocities depend on the M/L ratio and on the assumed tangential anisotropy. Our JAM model of NGC 4486a assumes isotropy. A velocity larger than the isotropic one is obtained in the model when the stars have little dispersion around the rotation axis and tend to rotate along circular orbits. This is a likely possibility for a thin disk as in NGC 4486a, where the stars may be still rotating with a kinematics similar to that of the gas disk from which they likely originated.

We also plotted the input difference maps (JAM maps convolved with SINFONI K-band PSF) in Fig. 14.6, to note the very small difference between the BH and no BH cases.

Nowak et al. (2007) calculated the mass of the SMBH at the centre of NGC 4486a using Schwarzschild modelling and reported $(1.25^{+0.75}_{-0.79}) \times 10^7\ M_\odot$ with 90 % confidence. Our figure of merit calculations reveal that the BH can be detected with a significance of 1.18σ when the measured and input velocity dispersion difference maps are compared and fitted (see Fig. 14.7 for the histogram of the fit slopes), whereas our other method yields a significance of 1.46σ. These values are in agreement with the significance of the BH mass calculation of Nowak et al. (2007).

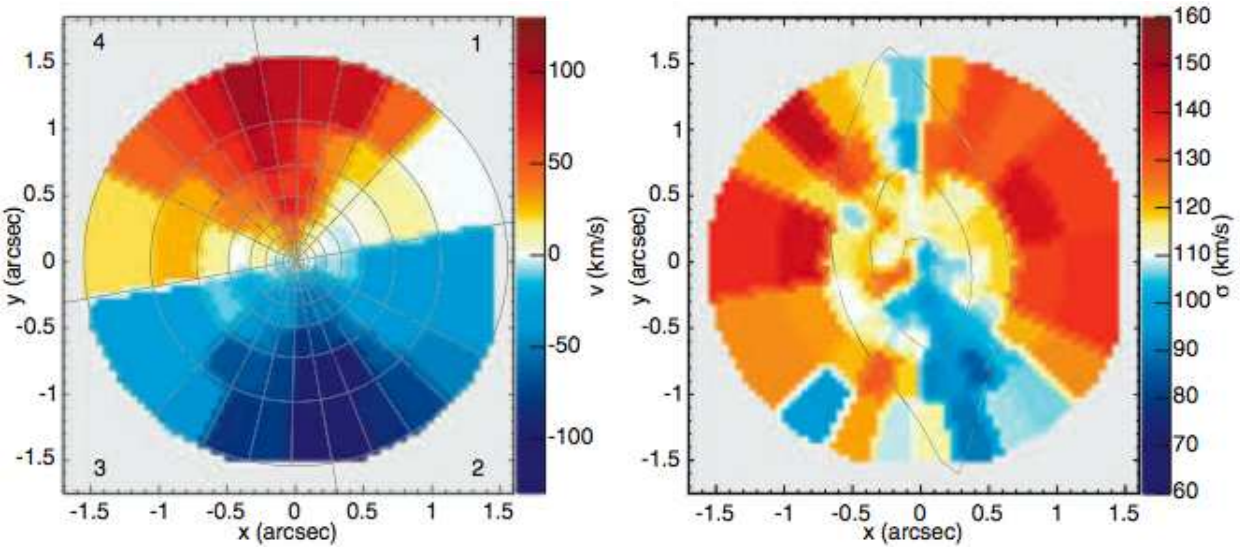

Figure 14.4: Velocity and velocity dispersion maps derived by Nowak et al. (2007) based on the real SINFONI observation of NGC 4486a.

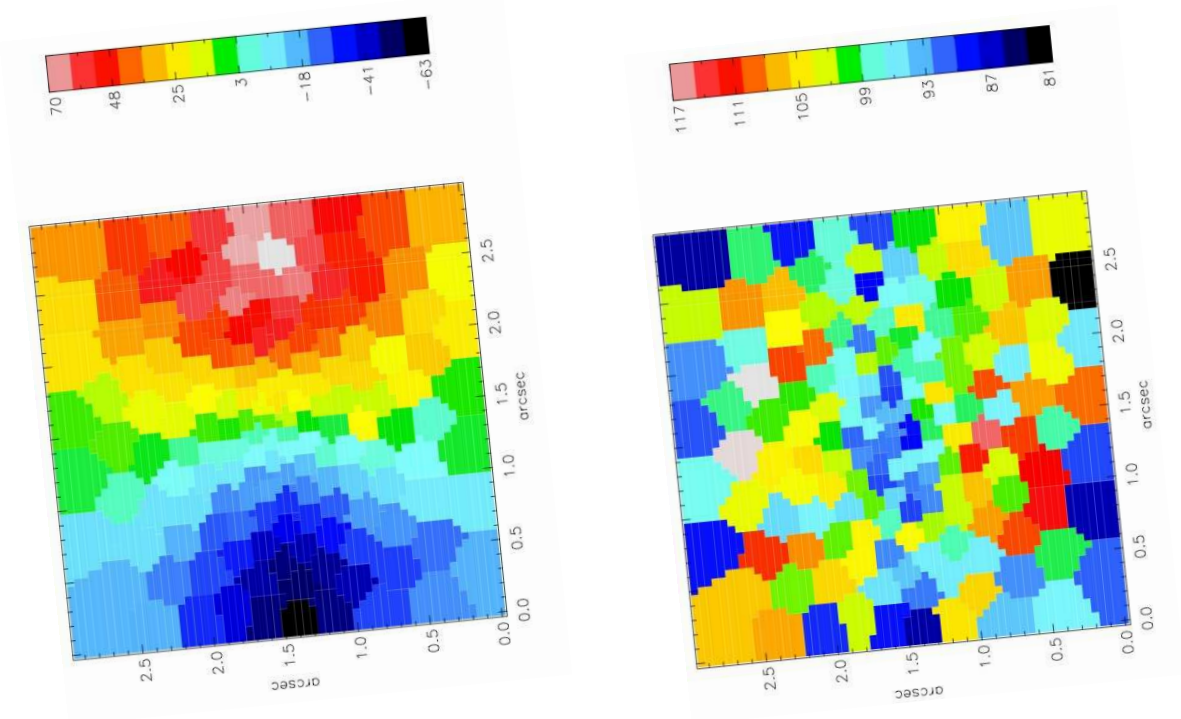

Figure 14.5: Velocity and velocity dispersion maps derived from the simulated SINFONI observation of NGC 4486a.

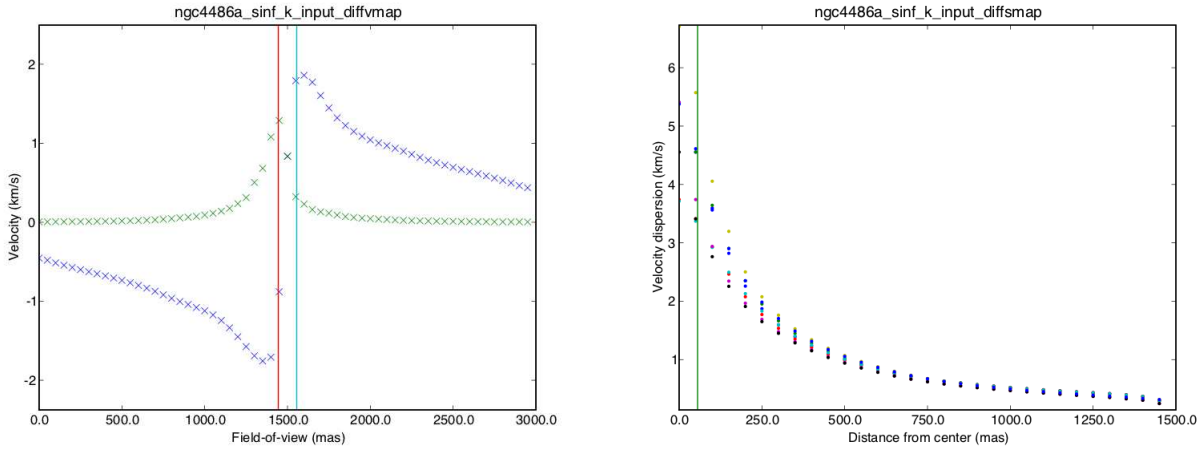

Figure 14.6: Velocity and velocity dispersion difference maps (convolved with the SINFONI K-band PSF) as input to the SINFONI observation simulations of NGC 4486a.

14.5.1.2 E-ELT simulations

We performed multiple simulation runs of E-ELT observations, with the main input parameters listed in Table 14.1. For all of these runs we also simulated the same galaxies with no SMBH at their centres.

All simulation runs have spectral resolution R $\sim$ 5000 and are performed for (i) a Paranal-like site, (ii) a High & Dry site, and also for (iii) an Al mirror coating and (iv) an Ag/Al mirror coating. We used an LTAO PSF, 5 mas spaxel size for low-redshift cases and 10 or 40 mas/spaxel for higher redshift cases. The FoV is $0.8'' \times 0.8''$ for all simulations.

Table 14.1: Main parameters for different simulation runs.

Galaxy template	BH mass	Distance	Wavelength	Sphere of influence
NGC 4486a	$1.25\times10^7\ M_\odot$	16 Mpc (Virgo)	K-band	55.8 mas
M87	$3\times10^9\ M_\odot$	z = 0.2	z, J and K-band	105 mas
NGC 5308	$2\times10^8\ M_\odot$	z = 0.2	z, J and K-band	12.65 mas
M87	$3\times10^9\ M_\odot$	z = 0.1	I and K-band	188.8 mas
NGC 5308	$2\times10^8\ M_\odot$	z = 0.1	I and K-band	22.6 mas
M87	$3\times10^9\ M_\odot$	z = 0.05	I and K-band	356.5 mas
NGC 5308	$2\times10^8\ M_\odot$	z = 0.05	I and K-band	42.8 mas

14.5.2 Analysis

The extraction of the stellar kinematics was performed using the penalised pixel-fitting method (pPXF) (Cappellari & Emsellem 2004), which fits the logarithmically re-binned spectra with a template after convolving it with a line-of-sight velocity distribution (LOSVD) described by a Gauss–Hermite expansion (Gerhard 1993; van der Marel & Franx 1993). The pPXF allows the template to be carefully optimised during the kinematics fit and includes a penalty criterion to deal with low S/N or insufficient resolution in the spectra. The pPXF code is publicly available at this web page.

To fit the IFU data we used the same stellar template as we used for the scientific input, in order to prevent any errors due to template mismatch. Furthermore, we used the BIAS parameter such that the LOSVD tend to be a Gaussian rather than a Gauss-Hermite polynomial. This is done by biasing the Gauss-Hermite parameters h3 and h4 towards zero.

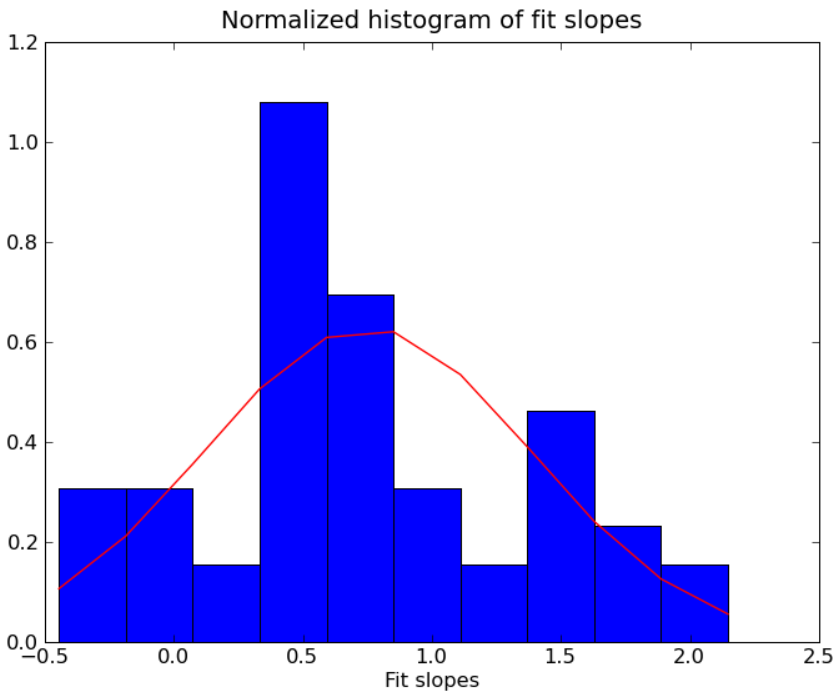

Figure 14.7: Normalised histogram of fit slopes for the input versus measured velocity dispersion difference with BH minus without BH, for NGC 4486a observation simulations for SINFONI. The measurements are done on 50 realisations of the SINFONI simulation for each case: with BH and without BH.

We ran the datacube pipeline multiple times with the same parameters and analysed them kinematically to evaluate and compare the results of the various simulation runs by applying the metrics and figures of merit from Section 14.3.

NGC 4486a: We simulated observations of NGC 4486a for E-ELT using the same galaxy kinematics, surface brightness and BH mass parameters that we used for the SINFONI simulations.

We used the Voronoi binning method (Cappellari & Copin 2003) to bin the spaxels to match S/N∼50 in order to obtain a decent fit. Then we fitted the CO region using pPXF (see Fig. 14.8). As expected, the observations using the same exposure time result in a much better S/N and the derived kinematics maps are much more precise (see Fig. 14.9) than in the SINFONI case. Note that the FoV is 0.8″×0.8″ and the spaxel size is 5 mas compared to the 3″×3″ FoV and 50 mas spaxel size of the SINFONI observations.

We also plot the input velocity and velocity dispersion difference maps (convolved with the E-ELT LTAO K-band PSF) and the output maps as measured using the simulated E-ELT observations of NGC 4486a. As can be seen from Fig. 14.11, the effect of the SMBH is quite significant.

14.5.2.1 z = 0.2

NGC 5308: The galaxy is not bright enough to be detected with 45 h exposure time using 5 mas/spaxel or with larger binning. We did not simulate spaxel sizes larger than 10 mas since the radius of the sphere of influence of the BH corresponds to 12.65 mas at this redshift.

M87: We simulated it in z–J band where the Calcium triplet lines fall at this redshift, using 5 mas/spaxel. However, there was not enough flux even when exposed for 45 h.

We also simulated observations of M87 at z=0.2 for E-ELT in the K-band with 5, 10 and 40 mas/spaxel.The best S/N in terms of flux is obtained with the spaxel size of 40 mas. The exposure time is 15 $\times$ 3600 s.

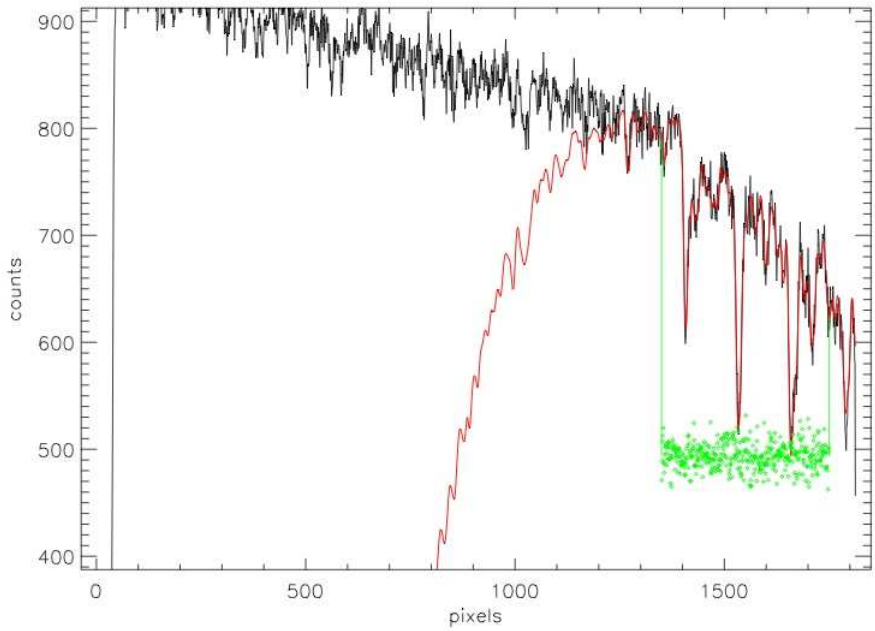

Figure 14.8: One of the binned spectra of simulated NGC 4486a observations with E-ELT, fitted using pPXF.

Note that we can still resolve the sphere of influence of the SMBH since this spaxel size is less than half of the radius of the BH's sphere of influence. We only simulated the worst case of the Paranal-like site with the bare aluminium coating. The S/N of the detection based on fit slopes of 50 realisations is 1.18 (see Fig. 14.12 for the histogram of slopes) whereas it is 1.15 using the second method.

These numbers are very similar to the significance of the detection of the BH in the case of SINFONI observations of NGC 4486a. Hence, using the E-ELT we will be able to detect a 3×10^9 $M_\odot$ black hole in an M87-like galaxy at a redshift of z = 0.2 with the same significance that we can detect the SMBH of NGC 4486a using SINFONI today.

14.5.2.2 z = 0.1

NGC 5308: The difference between BH and no BH is too small ($<$ 3 km s^{-1} in velocity dispersion, $<$ 1.5 km s^{-1} for velocity) in the I-band even though we used the best available PSF, i.e. LTAO I-band PSF.

We also simulated the observations in the K-band. Although the K-band LTAO PSF has a high Strehl ratio, there are several limitations. First of all the S/N of flux of the simulated observations is very low outside of the central $\sim$50 mas even with an exposure time of 45 hours. Furthermore, the input difference between velocity dispersions for BH and no BH cases is quite low ($<$ 12 km s^{-1}) when convolved with the LTAO K-band PSF, and there are no prominent lines (like the CO bandheads) that fall into the K-band at z=0.1.

M87: We simulated observations of M87 at z=0.1 for E-ELT in the K-band with 5, 10 and 40 mas/spaxel. The best S/N in terms of flux is obtained with the spaxel size of 40 mas. Table 14.2 shows the significance of the detection of the BH for four different cases using 40 mas/spaxel. The S/N is calculated using 50 realisations with BH and without BH each, for each case. We also did the fits and calculated the S/N using 100 realisations for the Paranal-like Ag/Al coating case. The normalised histogram is shown in Fig. 14.13. The S/N values (1.70 from fit slopes) are in agreement with the ones calculated using 50 realisations (1.61 from fit slopes).

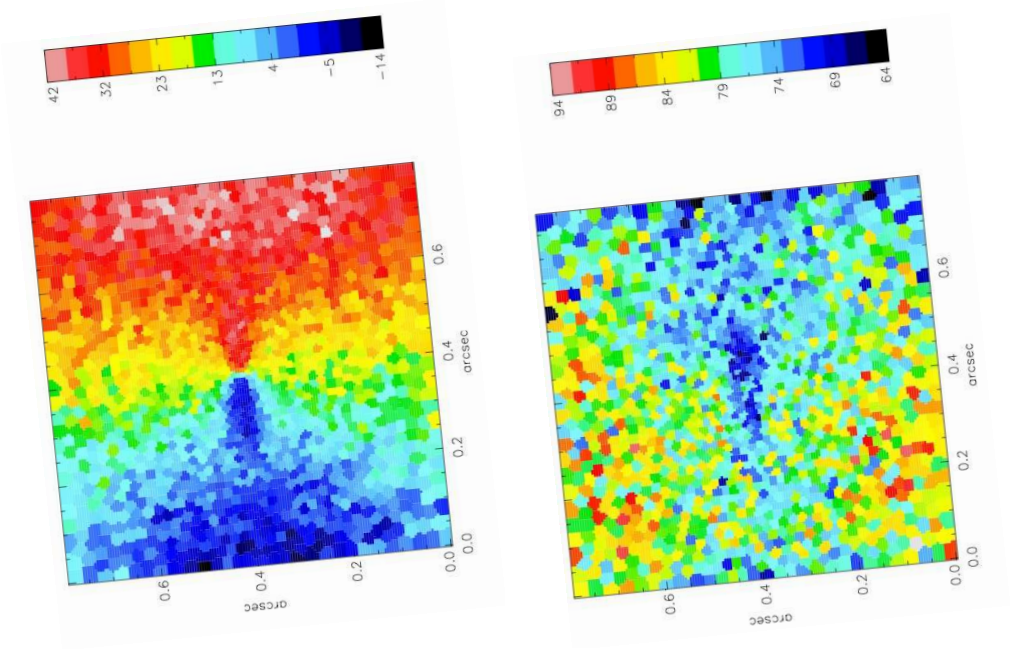

Figure 14.9: Velocity and velocity dispersion maps derived based on the simulated E-ELT K-band observations of NGC 4486a.

Table 14.2: Significance of the detection of the BH for M87 at z = 0.1.

Site	Mirror coating	S/N 1	S/N 2
Paranal-like	Al	0.47	1.89
Paranal-like	Ag/Al	1.70	1.88
High & Dry	Al	2.00	2.15
High & Dry	Ag/Al	2.35	2.49

Note that the S/N values obtained with both methods are in very good agreement except for the case with the Paranal-like site anf the bare Al coating (see Table 14.2). The difference for this case is due to a few vast outliers in the fits. It completely disappears if we give more weight to the spaxels at the centre of the galaxy (which are the most important ones since they are inside the sphere of influence of the BH).

14.5.2.3 z = 0.05

NGC 5308: We simulated observations of NGC 5308 at z=0.05 in the I-band to cover the 8500 Å Calcium triplet region. We used 5 mas/spaxel and an exposure time of 10 $\times$ 3600 s. We only simulated multiple instances for the case of the High & Dry site with the Ag/Al coating. Using 30 realisations of this case, we find that the BH is detected with a S/N of 3.24 using the first method, and 1.19 using the second method. The difference in S/N is most probably due to the fact that the input difference in velocity dispersion with BH minus without BH is less than 5 km s^{-1} which is on the order of the error of kinematics measurements using pPXF.

M87: We simulated observations of M87 at z=0.05 in the I-band which covers the 8500 Å Calcium triplet region. We used 5 mas/spaxel and an exposure time of 10 $\times$ 3600 s. We only simulated multiple instances for the case aof the Paranal-like site with the Ag/Al coating. Based on 20 realisations, we find that the BH is detected with a S/N of 25.6 using the first method, and 4.8 using the second method.

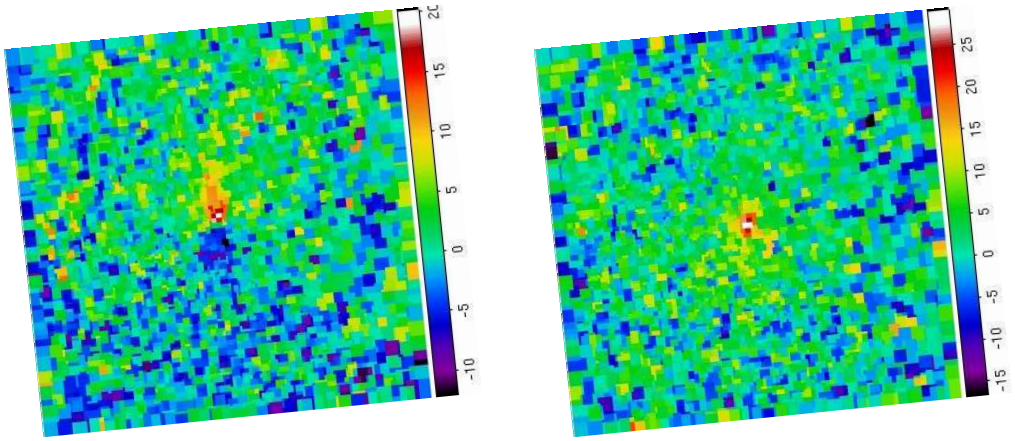

Figure 14.10: The difference of velocity and velocity dispersion maps with BH and no BH cases derived based on the simulated E-ELT K-band observation of NGC 4486a.

14.5.3 Compliance with figures of merit

Small mass BH at the Virgo distance: Black holes with 10^7 $M_\odot$ can be easily detected by the E-ELT (see Fig. 14.10). Extrapolating our simulations, we conclude that lower mass BHs like the one at the centre of our own Galaxy can be studied using E-ELT IFU with LTAO observations at the distance of the Virgo cluster.

Upper distance limit for BH detection: Our analysis showed that M87 at a redshift of z = 0.2 can be detected with the same significance as the furthest SMBH detection today using the best available facilities. Based on this, we conclude that black holes with $\sim 10^{10}$ $M_\odot$ can be detected out to higher redshifts of around z$\sim$0.3.

Necessary exposure times: The highest redshift simulations (M87 at z=0.2) for which one can achieve a similar detection significance as for the furthest SMBH detection with the current best facilities requires 15 hours of integration (15 $\times$ 3600 s). The low mass SMBH at the Virgo distance requires an hour of integration.

Most efficient spectral regions to observe: For the local Universe, up to the distance of Virgo, the K-band is better since the K-band LTAO PSF has a much better Strehl ratio than the other investigated bands and its wavelength range covers the prominent CO bandheads. The I-band region is better for redshifts z $\leq$ 0.05 since it covers the prominent Calcium triplet lines which are also used as metallicity indicators. For higher redshifts, up to z$\sim$0.3 the Ca triplet lines fall inbetween the z and J bands and there are no prominent lines falling into the wavelength range of the K-band. Nonetheless, the K-band observations provide the best data, as the line detection is improved by telluric correction and the high Strehl ratio of the K-band PSF ensures that the influence of the BH on the kinematics is within detectable limits with a moderate spectral resolution (R $\sim$ 5000).

Impact of PSF on the detection of BH signature: We have studied the effect of the PSF both by comparing the observation simulations in different wavelength ranges and also by comparing the simulations done with the LTAO I-band PSF and a modified I-band PSF which has a higher Strehl ratio (see Fig. 14.14).

We simulated multiple instances for the case of M87 at z=0.05 at the Paranal-like site with the Ag/Al coating. We used 20 realisations each with the original LTAO I-band PSF and the modified LTAO I-band PSF to determine the significance of the BH detection. For the original I-band PSF, we find that the BH is detected with a S/N of 25.6 using the first method, and 4.8 using the second method. For the modified I-band PSF with higher Strehl ratio, we find that the BH is detected with a S/N of 35.2 using the first

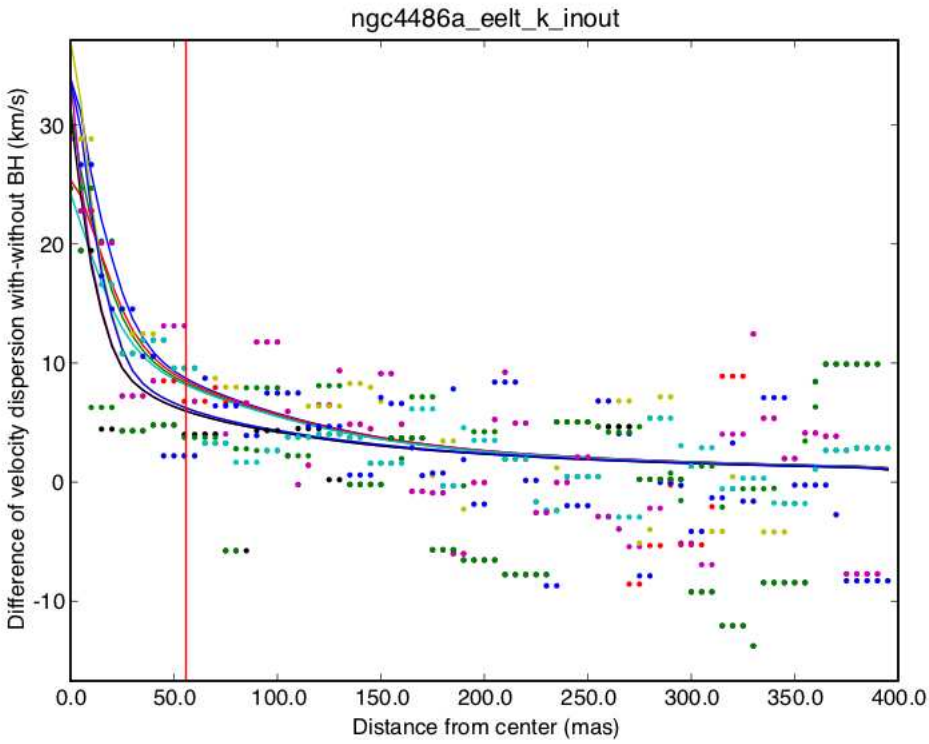

Figure 14.11: Velocity dispersion difference maps both input and as measured based on the simulated E-ELT K-band observations of NGC 4486a.

method, and 30.1 using the second method. Therefore, we conclude that the PSF, i.e. the Strehl ratio, has a strong impact on the detection of the SMBH.

14.5.4 Sensitivity to input parameters

We conducted the simulations of M87 at z = 0.1 in the K-band with an LTAO PSF for four different cases: (i) the Paranal-like site with the bare Al mirror coating; (ii) the Paranal-like site with the protected Ag/Al coating; (iii) the High & Dry site with the bare Al coating; and (iv) the High & Dry site with the Ag/Al coating. As shown in Table 14.2 using the High & Dry site parameters and to a lower extent using the Ag/Al coating improves the significance of the detection of the SMBH.

The mirror coating also plays a role for the I-band observations.

14.5.5 Calibration requirements

The calibration requirements are the same as with those of current near-IR IFU observations. Namely, a dither pattern like object-sky-object is needed to subtract the infrared sky, and observations of a star are needed to determine the PSF and the telluric correction.

14.5.6 Limitations

The main limitation of these simulations is that we do not determine the significance of SMBH detection by fitting the dynamics using complex modelling based on Schwarzschild orbit-superposition method. Instead we compare the kinematics of the galaxy with the case in which the same galaxy has no SMBH (see Section 14.3 for details).

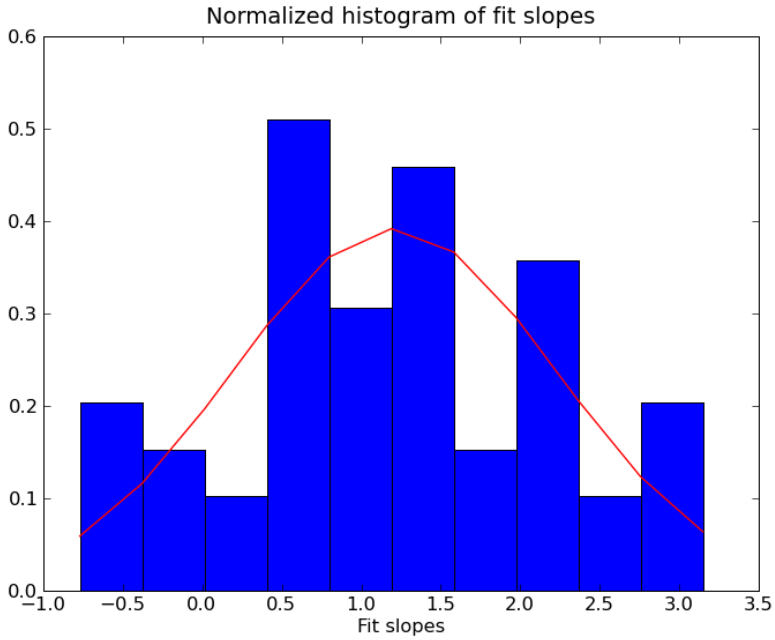

Figure 14.12: Normalised histogram of fit slopes for the input versus measured velocity dispersion difference with BH minus without BH, for M87 at z = 0.2 using the Paranal-like site and the bare Al mirror coating. The measurements are done on 50 realisations for each case: with BH and without BH.

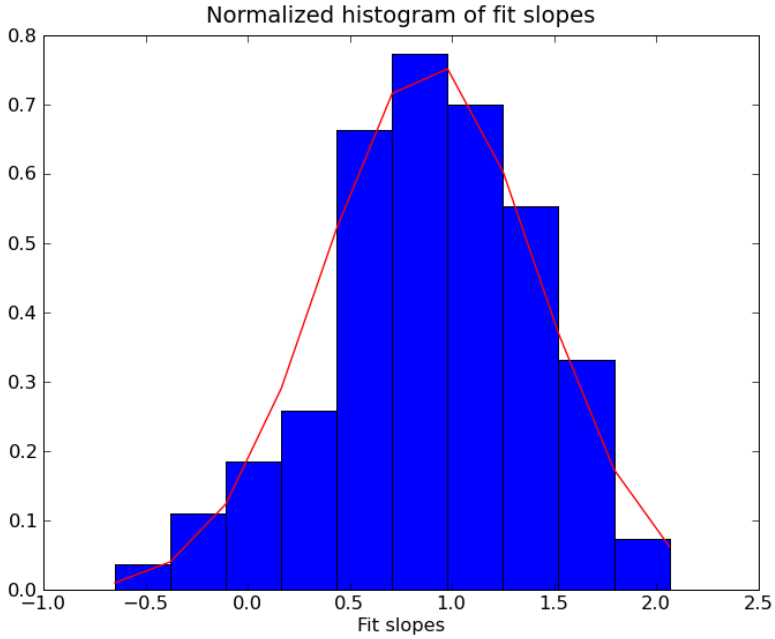

Figure 14.13: Normalised histogram of fit slopes for the input versus measured velocity dispersion difference with BH minus without BH, for M87 at z = 0.1 using the Paranal-like site and the Ag/Al mirror coating. The measurements are done on 100 realisations for each: with BH and without BH.

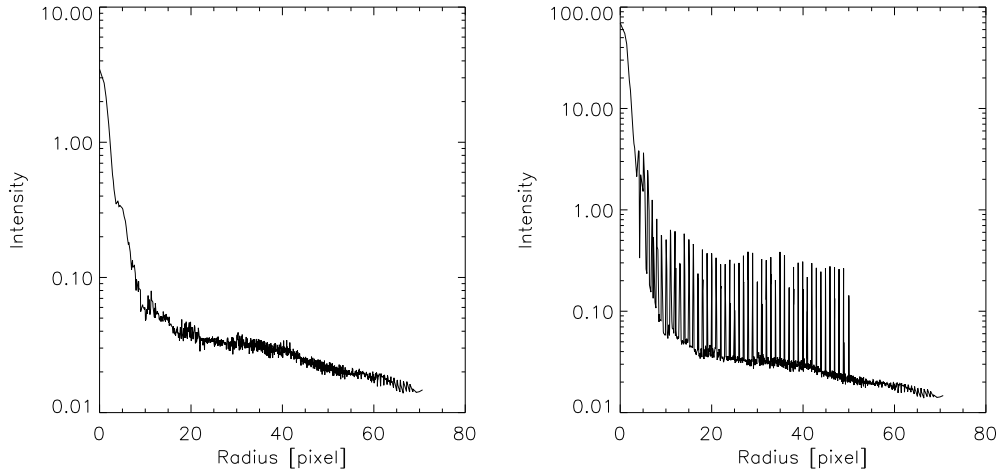

Figure 14.14: Azimuthally averaged profile of the original E-ELT LTAO I-band PSF (left) and of the modified LTAO I-band PSF which has a higher Strehl ratio (right).

Further possible limitations of the simulations are the following:

- The galaxy is assumed to be symmetric, i.e. the input raw science spectrum in each spaxel is identical.
- A perfect knowledge of the underlying galaxy spectrum is assumed, i.e. the template spectrum used for kinematics fits is same with the input raw science spectrum.
- The velocity dispersion of stars is assumed to be Gaussian during the simulations of the datacubes.
- A perfect data reduction, i.e. bias, dark, flat correction, is assumed.
- A perfect filter curve, i.e. with 100% transmission everywhere, is assumed.
- The effect of airmass is not considered.
- A good knowledge of the PSF is assumed when determining the significance of SMBH detection, i.e. the measured kinematics are compared with the input values convolved with the relevant E-ELT PSF.

The following can be improved in terms of analysis and figures:

- The BH mass can be determined via more complex methods like Schwarzschild modelling as a measure of the detection significance.
- A more sophisticated data analysis technique could be adopted for the I-band simulations, such as a weighted summation of the flux, given that 94 % of the flux in I-band falls outside the spaxel.

14.6 Concluding remarks

We have conducted simulations for the DRM science case G9 “A survey of Black Holes in different environments” exploring a wide range of observational and physical parameters, for SMBH masses of $\sim 10^7$ – 10^9 $M_\odot$ at distances from 16 Mpc to z=0.2. The goals of the science case defined in Section 14.2 are addressed using figures of merit as described in Section 14.3. The details can be found at Section 14.5.3.

Up to a redshift of z$\sim$0.2, the E-ELT can resolve the BH sphere of influence with IFU kinematics for a large number of targets, dramatically reducing the measurement errors on BH masses. This would allow us to clarify the long debated issue of the scatter of the BH-host galaxy correlations such as the $M_{\rm BH}$-σ relation.

Acknowledgements

Aybüke Küpcü Yoldaş is grateful to Nina Nowak for providing the SINFONI PSF of NGC 4486a observations and useful suggestions at the beginning of this work. AKY also thanks Nadine Neumayer and Mariya Lyubenova for their help in getting started with IDL and pPXF fitting.

References

Cappellari M., 2008, MNRAS, 390, 71
Cappellari M., Copin Y., 2003, MNRAS, 342, 345
Cappellari M., Emsellem E., 2004, PASP, 116, 138
Cappellari M., Neumayer N., Reunanen J., et al., 2009, MNRAS, 394, 660
Emsellem E., Monnet G., Bacon R., 1994, A&A, 285, 723
Emsellem E., Cappellari M., Peletier, R.F., et al., 2004, MNRAS, 352, 721
Gerhard O.E., 1993, MNRAS, 265, 213
Gustafsson B., Edvardsson B., Eriksson K., Jørgensen U.G., Nordlund A., Plez B., 2008, A&A, 486, 951
Macchetto F., Marconi A., Axon D., Capetti A., Sparks W., Crane P., 1997, ApJ, 489, 579
Nowak N., Saglia R.P., Thomas J., et al., 2007, MNRAS, 379, 909
Schwarzschild M., 1979, ApJ, 232, 236
Scott N., Cappellari M., Davies R.L., et al., 2009, MNRAS, 398, 1835
van der Marel R.P., Franx M., 1993, ApJ, 407, 525

15 C10-1: The physics and mass assembly of galaxies out to z∼6

Authors: M. Puech, P. Rosati, S. Toft, A. Cimatti

15.1 The science case

We propose to obtain ELT spatially resolved spectroscopy of a sample of a thousand massive galaxies at $2 \leq z \leq 6$, selected from future large area optical-near IR surveys. These observations will yield direct kinematics of stars and gas in the first generation of massive galaxies (in the range $0.1 \leq M_{\rm stellar} \leq 5 \times 10^{11}\ M_{\odot}$), as well as their stellar population properties. One will be able to derive dynamical masses, ages, metallicities, star-formation rates, dust extinction maps, to investigate the presence of disk and spheroidal components and the importance of dynamical processes (e.g. merging, in/outflows) which govern galaxy evolution. These data will also allow one to study the onset of well known scaling relations at low redshifts, and to witness the gradual shift of star formation from the most massive galaxies in the highest density regions to less massive galaxies in the field. The whole program is designed to provide the ultimate test of galaxy formation theories.

15.2 Goals of the DRM simulations

DRM simulations for the present science case shall focus on the sub-sample of distant emission line galaxies. Due to signal-to-noise ratio (SNR) limitations, absorption line galaxies at $z \gtrsim 1.5$ will more likely be studied using integrated spectroscopy (see Science Case C10-3 "ELT integrated spectroscopy of early-type galaxies at $z > 1$"). Therefore, in the following, only emission line galaxies will be considered and for convenience, we will sometimes use the single word "galaxies" to refer to "emission line galaxies".

Resolved kinematics of such galaxies provides a useful testbed for 3D spectrographs, because it drives the stringent requirements on the SNR: while flux is the zero-order moment of an emission line, the velocity and velocity dispersion are derived from the position and the width of emission lines, which are their first and second moments, and higher order moments always have more relative measurement uncertainty. Hence, DRM simulations shall assess the principal scientific goal of spatially-resolved spectroscopy of distant emission line galaxies, focusing on kinematics. Different objectives can be defined depending on the level of accuracy and/or spatial scale one wants to probe; we have defined the following objectives (dubbed as "steps"), ordered by increasing complexity/difficulty:

1. Simple (3D) detection of emission line galaxies: what stellar mass can we reach at a given SNR as a function of redshift, AO system, environmental conditions, etc.?
2. Recovery of large scale motions (see, e.g., Flores et al. 2006): what are the conditions under which it is possible to recover the dynamical state of distant galaxies (e.g., relaxed rotating disks vs. non-relaxed major mergers)?
3. Recovery of rotation curves: what are the conditions under which it is possible to recover the rotation velocity $V_{\rm rot}$ (to derive, dynamical masses or the Tully-Fisher relation, see, e.g., Puech et al. 2006a, 2008) or the whole shape of the RC (for derivation and decomposition in mass profiles, see, e.g., Blais-Ouellette et al. 2001)?
4. Recovery of the detailed kinematics: what are the conditions under which it is possible to detect internal structures in distant galaxies, like clumps in distant disks (see, e.g., Bournaud et al. 2008)?

15.3 Metrics / figures of merit

As a general metric we will use the total observation time of the survey $T_{\rm intg}$ required to achieve the proposal goals, i.e., observing ∼1000 galaxies more massive than $M_{\rm stellar} = 10^{10}\ M_{\odot}$ in the redshift range 2–6. This high number of galaxies is required to allow us to derive statistics over the morpho-kinematic types of galaxies in several redshift and stellar mass bins (see Section 15.5.3). Given the

very ambitious goals of the proposal, they will be considered as met if $T_{\rm intg} \leq 100$ nights. This roughly corresponds to the total time allocated to Large Programs at the VLT per year (30% of the available time). Therefore, such a survey could reasonably be implemented as a several years effort. Finally, this total time should correspond to SNR levels that guarantee to reach at least step 2 for most of galaxies in the survey, and step 4 for a more limited sub-sample (to be defined).

15.4 DRM simulations

15.4.1 Methodology

All input parameters can be separated into two broad categories, namely the "physical" parameter space, which includes all parameters defining the distant galaxy (i.e., redshift z, galaxy diameter, continuum AB magnitude $m_{\rm AB}$, rest-frame emission line equivalent width EW_0, velocity gradient, morpho-kinematical type) and the "observational" parameter space, which includes the telescope (M1 and M2 sizes, temperature $T_{\rm tel}$, emissivity $\epsilon_{\rm tel}$), the instrument (spectral resolution R, IFU pixel size $\Delta_{\rm pix}$, detector integration time dit, number of exposures $ndit$, temperatures $T_{\rm instr}$ and emissivities $\epsilon_{\rm instr}$, AO correction, global transmission $t_{\rm trans}$) and the site (seeing, C_n^2 profile, outer scale of the turbulence L_0, sky brightness and atmospheric transmission).

Given the very large number of parameters to be investigated, as well as the very large range of values to be explored, it is useful to define a "reference case", around which the parameter space can be explored and compared with (see Fig. 15.1). As such a reference case, we took an M_* galaxy at z=4 observed using MOAO with a seeing of 0.8 arcsec (see Section 15.4.3.1). At this redshift, the [OII] emission line is observed in the H-band, where the influence of the thermal background is minimised in comparison with the K-band. This makes this reference case as independent as possible of the telescope design (e.g., number of mirrors), environmental conditions (site selection) and instrument characteristics (e.g., number of warm mirrors), which are not all fully known at present.

This reference case will be used to assess separately the influence of:

- the AO correction for a given set of other technical and physical parameters;
- the observational parameters for a given set of physical parameters and AO correction;
- the physical parameters for a given set of observational parameters (AO correction included).

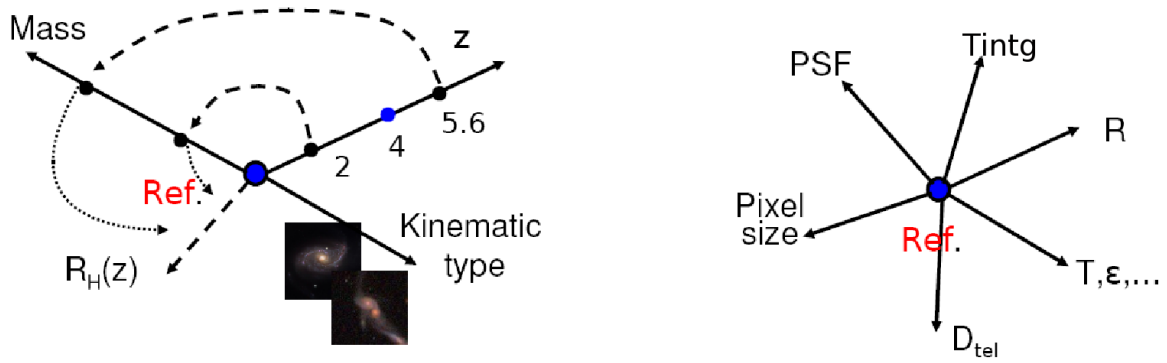

Figure 15.1: Illustration of the strategy used for exploring the parameter space. A reference case is defined, which is an M_* galaxy at z=4 observed using MOAO with a seeing of 0.8 arcsec (see Section 15.4.3.1). Then, the physical parameter space (on the left) and the observational parameter space (on the right) can be explored relatively to this reference case.

15.4.2 Pipeline

The end-products of 3D spectrographs are data-cubes in FITS format. Hence, we have developed a simulation pipeline that produces such mock data. These fake data are produced assuming a perfect data reduction process, whose impact has therefore been neglected.

15.4.2.1 General description

Data-cubes are simulated using a backward approach of the usual analysis of 3D data. During such a process, each spaxel is analysed separately in order to extract from each spectrum the continuum and the first moments of emission lines. One ends up with a set of maps describing the spatial distribution of the continuum and line emissions (zeroth order moment of the emission line), as well as the gas velocity field (first order moment) and velocity dispersion map (second order moment). Such maps are routinely derived from Fabry-Perot observations of local galaxies (see, e.g, Epinat et al. 2008), or can be produced as by-products of hydrodynamical simulations of local galaxies (e.g., Cox et al. 2006). Using simple rules and empirical relations, these maps can be rescaled (in terms of size and total flux) to provide a realistic description of distant galaxies (see Section 15.4.3.1). Assuming a Gaussian shape for all spectra, it is then straightforward to reconstruct a data-cube from such a set of four maps, by "reversing" the usual analysis of 3D data. This allows us to produce a data-cube at high spatial resolution, which can then be degraded at the spatial resolution (AO PSF) and sampling (IFU pixel scale) required for simulating data produced by a given 3D spectrograph. A more detailed description of the process is given below.

15.4.2.2 High resolution data-cube

At very high spatial resolution, emission lines with kinematics driven only by gravitational motions are well described by simple Gaussians (Beauvais & Bothun 1999). Under this assumption, within each pixel of the data-cube, only four parameters are needed to fully define a spectral line. The first three parameters are the position in wavelength, the width and the area (or, equivalently the height) of the emission line. The current version of the software only models emission lines and does not take into account the detailed shape of the continuum in galaxies - it is simply modelled as a constant in f_λ. So only one parameter is required to set the level of this pseudo-continuum around the emission line. Each spectrum is generated in the observational frame, at a given spectral sampling of $\lambda_{\rm obs}/2R$, where R is the spectroscopic power of resolution of the instrument, and $\lambda_{\rm obs} = (1+z)\lambda_{\rm em}$, $\lambda_{\rm em}$ being the rest-frame wavelength of the emission line, and z the redshift of the simulated object. During this process, the rest-frame line width is multiplied by $(1+z)$, as one needs to take into account the widening of emission lines with increasing redshift.

All these four parameters can be extracted from observations of local galaxies. In the following, we use Fabry-Perot (FP) observations of the Hα emission distributions of nearby galaxies obtained as part of the GHASP survey (Amram et al. 2002). From these data, we can extract four parameters, wavelength, width, area and continuum level to construct the velocity field, the velocity dispersion map, the flux map of the Hα emitting gas and the continuum map of the galaxy. The software first re-scales all these maps at a given angular size (in arcsec) provided by the user. These maps are then interpolated at a spatial sampling of $\lambda_{\rm obs}/2D$. This spatial sampling is motivated by the fact that the AO PSFs used have been simulated at this sampling rate (see below). The software then re-scales the overall amplitude of the continuum map at a given integrated number of photons using an integrated magnitude $m_{\rm AB}$ directly provided by the user. This magnitude is converted to the number of photons per spectral pixel depending on the telescope diameter D, the total integration time $T_{\rm intg}$ and the global transmission $t_{\rm trans}$ of the system (atmosphere excluded). The overall amplitude of the Hα map is also rescaled at a given integrated number of photons, derived from the integrated continuum value and a rest-frame equivalent width EW_0 provided by the user, the latter being re-scaled in the observed-frame by multiplying by $(1+z)$. All parameters can also be extracted from outputs of hydro-dynamical simulations (Cox et al. 2006). In this case, the two last parameters (area and continuum level) can respectively be extracted from total gas and stellar surface density maps.

15.4.2.3 Modelling the IFU and the detector

Each monochromatic slice of the high resolution data-cube is convolved by a PSF with matching spatial sampling. This PSF must be representative of all elements along the optical path, from the atmosphere to the output of an (optional) Adaptive Optic system. Because the isoplanatic patch (Fried 1981; the median value at Paranal is ∼2.4 arcsec at $\lambda \sim 0.5$ μm, which leads, e.g., to ∼10 arcsec at $\lambda \sim 1.6$ μm) is larger than the individual FoV of the IFU (typically a few arcsec, depending on the size of objects at a given redshift), the same PSF can be used to convolve the data-cube regardless of position within the IFU. We also neglected the variation of the PSF with wavelength (the FWHM varies as $\lambda^{-1/5}$ in a Kolmogorov model of the atmospheric turbulence, see, e.g., Roddier 1981), as we are only interested in the narrow spectral range around a single emission line.

The next step is to reduce the spatial sampling of the data-cube. This is done by re-binning each monochromatic channel of the data-cube at the pixel size $\Delta_{\rm pix}$ of the simulated IFU. A wavelength dependent atmospheric absorption curve is then applied to each spectrum of the data-cube. Sky continuum, detector dark level, bias, and thermal background are then added to the spectra. Thermal backgrounds are modelled using grey bodies parametrised by a temperature T and an emissivity ϵ, i.e., a black body of temperature T multiplied by a constant emissivity ϵ. Photon and detector noise are then added. In the NIR, the detector noise is due to the dark current ($dark$) and read-out noise (RON).

Finally, the simulation pipeline generates $ndit$ data-cubes with individual exposure time of dit, which are combined by estimating the median of each pixel to simulate several individual realistic exposures. Since we have only included random noise, it is similar to having dithered all of individual exposures and combining them after aligning them spatially and spectrally. Sky frames are evaluated separately (i.e., with a different noise realisation), and then subtracted from each individual science frame, reproducing the usual procedure in both optical and NIR spectroscopy.

15.4.2.4 Post-treatment of simulated data-cubes

The simulated data-cubes are analysed using an automatic data analysis pipeline similar to those generally used to analyse data of high redshift galaxies. During this process, each spatial pixel of the simulated data-cube is fitted with a Gaussian in wavelength, whose position and width correspond to the velocity and velocity dispersion of the gas in this spatial pixel, respectively. Because the accuracy on these measurements is driven by the flux within the emission line relative to the noise on the continuum, one has to define a kinematic signal-to-noise ratio $\mathrm{SNR}_{\rm kin}$ specific to this measurement. We chose to rely on the definition of Flores et al. (2006), who defined $\mathrm{SNR}_{\rm kin}$ as the total flux in the emission line divided by the noise on the continuum $\sigma_{\rm continuum}$ and $\sqrt{N_{\rm pix}}$, the number of pixels within the emission line. We emphasise that this $\mathrm{SNR}_{\rm kin}$ is different from the classical SNR: the latter quantifies the signal-to-noise ratio relative to sky and detector noise (detectability of the emission line), while the former characterises the capability of measuring the first and second order moments of the emission line, relative to the noise in the continuum (Sarzi et al. 2006). Although they are different in nature, we find that both SNRs roughly correlate provided that SNR > 3 (corresponding the $\mathrm{SNR}_{\rm kin} \sim$ 3–4). Tests show that pixels with $\mathrm{SNR}_{\rm kin} < 3$ have very large errors mostly due to misidentification of noise peaks as emission lines. Hence, only pixels having $\mathrm{SNR}_{\rm kin} \geq 3$ have been considered for further analysis.

15.4.2.5 Validation experiment

We conducted a special run to compare the simulation pipeline with real 3D observations on the VLT of a galaxy at z∼2.4 observed by R. Genzel's group using SINFONI (Genzel et al. 2006). The goal of this run is to use input parameters corresponding to a real observed case and assess whether or not the simulation pipeline is able to produce a data-cube with the same quality, quantified using the median $\mathrm{SNR}_{\rm kin}$ (see definition in the previous section). For the z∼2.4 galaxy observed with SINFONI, the corresponding input parameters are: size of 0.8 arcsec, $K = 21.47$, $\mathrm{EW}_0(\mathrm{H}\alpha) = 140$ Å, $T_{\rm intg} = 6$ hr, pixels of 50×100 mas^2, PSF with FWHM = 150 mas, temperature of the telescope and instrument of 287 K (VLT), emissivity of the telescope of 6% (Cassegrain focus), emissivity of SINFONI of 15%. We took care of mimicking the SINFONI data reduction procedure by interpolating pixels from the physical 50×100 mas^2 spatial

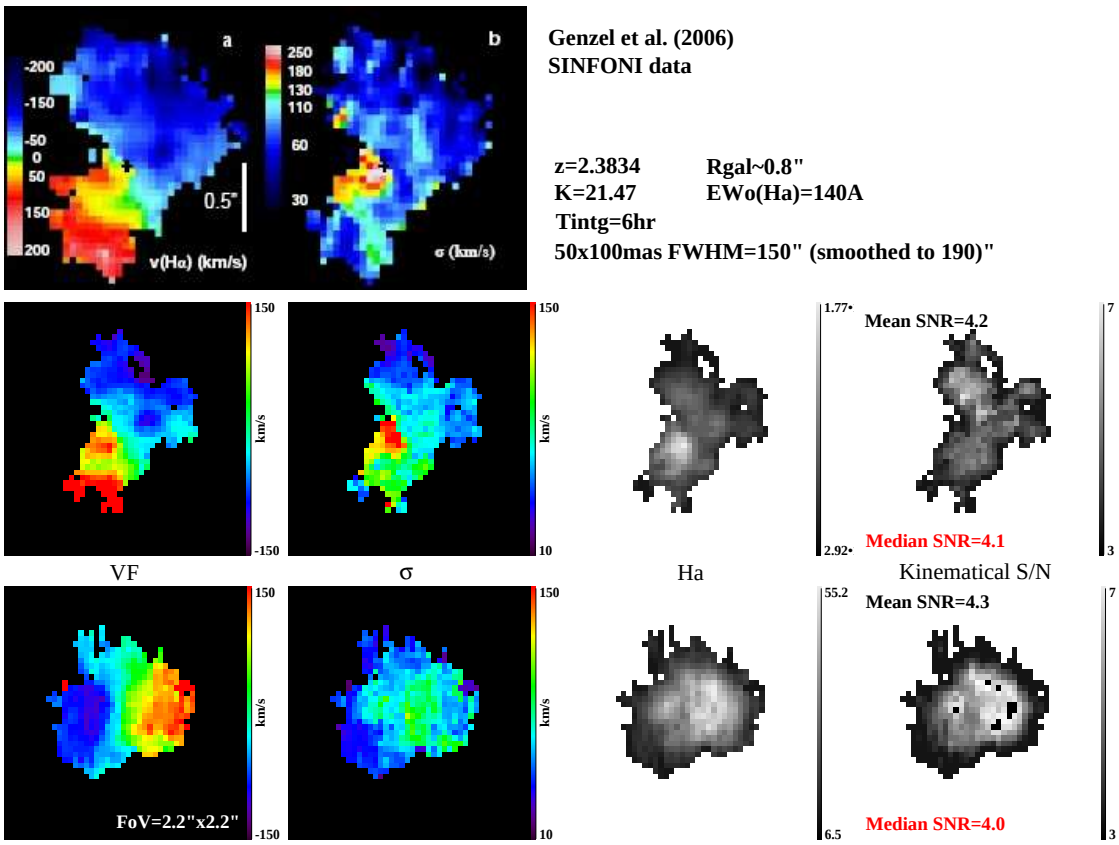

Figure 15.2: Comparison of SINFONI data with simulations of a rotating disk (UGC5253 template, see Section 15.4.3.1) using the same set of observational and physical parameters (see text). First line: SINFONI observations of a z$\sim$2.4 galaxy (left: velocity field; right: velocity dispersion map). Second line: Analysis of the reduced SINFONI data-cube (courtesy of N. Förster-Schreiber). From left to right: velocity field, velocity dispersion map, Hα map and $\mathrm{SNR_{kin}}$ map. Third line: simulations of a rotating disk using the technical and physical parameters corresponding to the real SINFONI observations. The median $\mathrm{SNR_{kin}}$ is found to be very similar in both cases.

scale to the 50$\times$50 mas^2 final scale and smoothing down the data-cube to a resolution of 190 arcsec. The results of this validation run are shown in Fig. 15.2 and demonstrate that, given a complete set of technical and physical parameters, simulations can produce data-cubes with a quality similar to that of real observations. Moreover, the similarity between the observed velocity field and velocity dispersion map (see first line in Fig. 15.2) and those produced by the automatic analysis pipeline (see second line) shows that the automatic analysis pipeline can be used safely.

15.4.3 Inputs

All magnitudes are given in the AB system.

15.4.3.1 Scientific data

Scientific data are needed in order to (1) provide the simulation pipeline with morpho-kinematical templates to be used as inputs and (2) rescale these templates in terms of realistic size, flux and velocity gradient compared to distant galaxies. A global flowchart of the rescaling procedure is shown in Fig. 15.3, while specific details are given below.

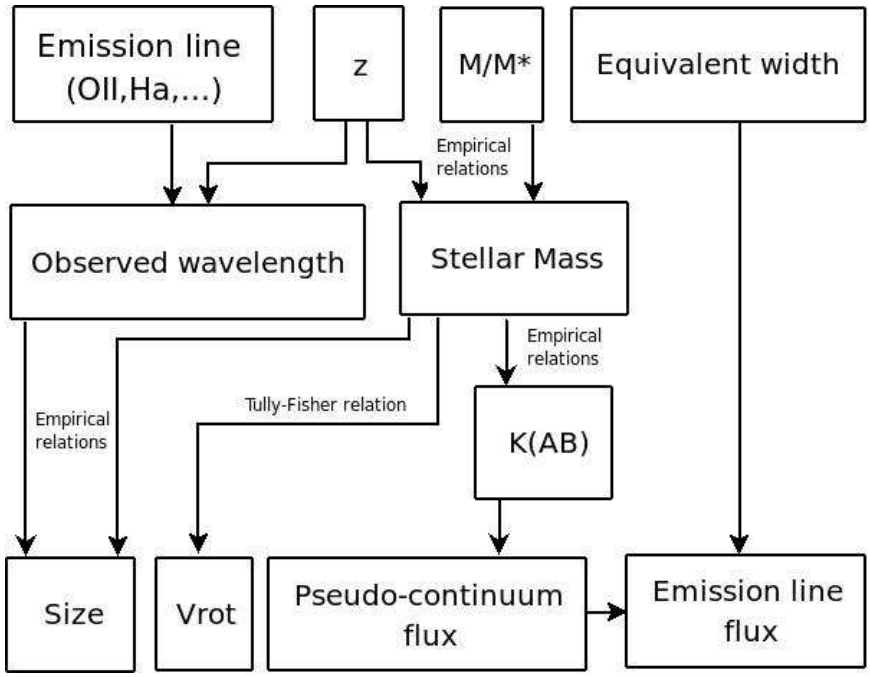

Figure 15.3: Schematic illustration of the method used to rescale the morpho-kinematical templates in order to make them match the size and flux of distant galaxies. Also, the velocity dispersion (not shown here) can be rescaled using a multiplicative value.

Morpho-kinematical templates: In Figs. 15.4 and 15.5 we show the high resolution templates used for the simulations and we summarise their main properties in Table 15.1.

Redshift: Given the objective of the survey, three redshift "bins" have been considered for simulations. Table 15.2 summarises these redshifts with the "observed" emission line. For simplicity, the [OII] emission line was assumed to be a single line centered at 3727 Å; this allows us to avoid introducing an additional parameter, i.e., the line ratio between the lines of the doublet, which depends mostly on the electron density in the medium (Puech et al. 2006a). This does not influence the integrated SNR over the emission line but leads to overestimate the maximal spectroscopic SNR in the emission line. The impact of this assumption will be investigated in Section 15.5.5. Note that the highest redshift considered (z=5.6) corresponds to the limit where [OII] falls within the K-band.

Stellar mass and flux: We used the MUSIC compilation of public spectro-photometric data in the GOODS field (Grazian et al. 2006) to derive empirical relations between redshift, observed K-band magnitude and stellar-mass $M_{\rm stellar}$ (see Fig. 15.6). The latter quantity was expressed as a fraction of the characteristic stellar-mass $M_*(z)$ at a given redshift, which describes the knee of the Galaxy Stellar Mass Function (GSMF) at this redshift, according to a Schechter function (see inset in Fig. 15.6). In other words, at a given z, galaxies with stellar mass $M_{\rm stellar} = M_*(z)$ are those that contribute the most to the stellar mass density at this redshift. Table 15.3 gives the corresponding stellar masses in the simulations as a function of redshift. The pseudo-continuum flux around the emission line is directly derived from the K-band magnitude; we did not apply any "colour" correction between the rest-frame wavelength of the emission line (e.g., 0.3727 μm for [OII]) and that of the K-band (e.g., 0.44 μm at z=4, see Fig. 15.7). Such an assumption is consistent with Spectral Energy Distributions of galaxies with a morphological type latter than Sa within a factor two in flux. As a reference case, we adopted a z=4, M_* galaxy (see Section 15.4.1).

Rest-frame emission line equivalent width: We assumed EW_0([OII])=30Å, which is an extrapolation of the mean value found at z=1 (Hammer et al. 1997). This parameter does not influence the flux distribution of the galaxy but is used to set its total integrated value (see Section 15.4.1).

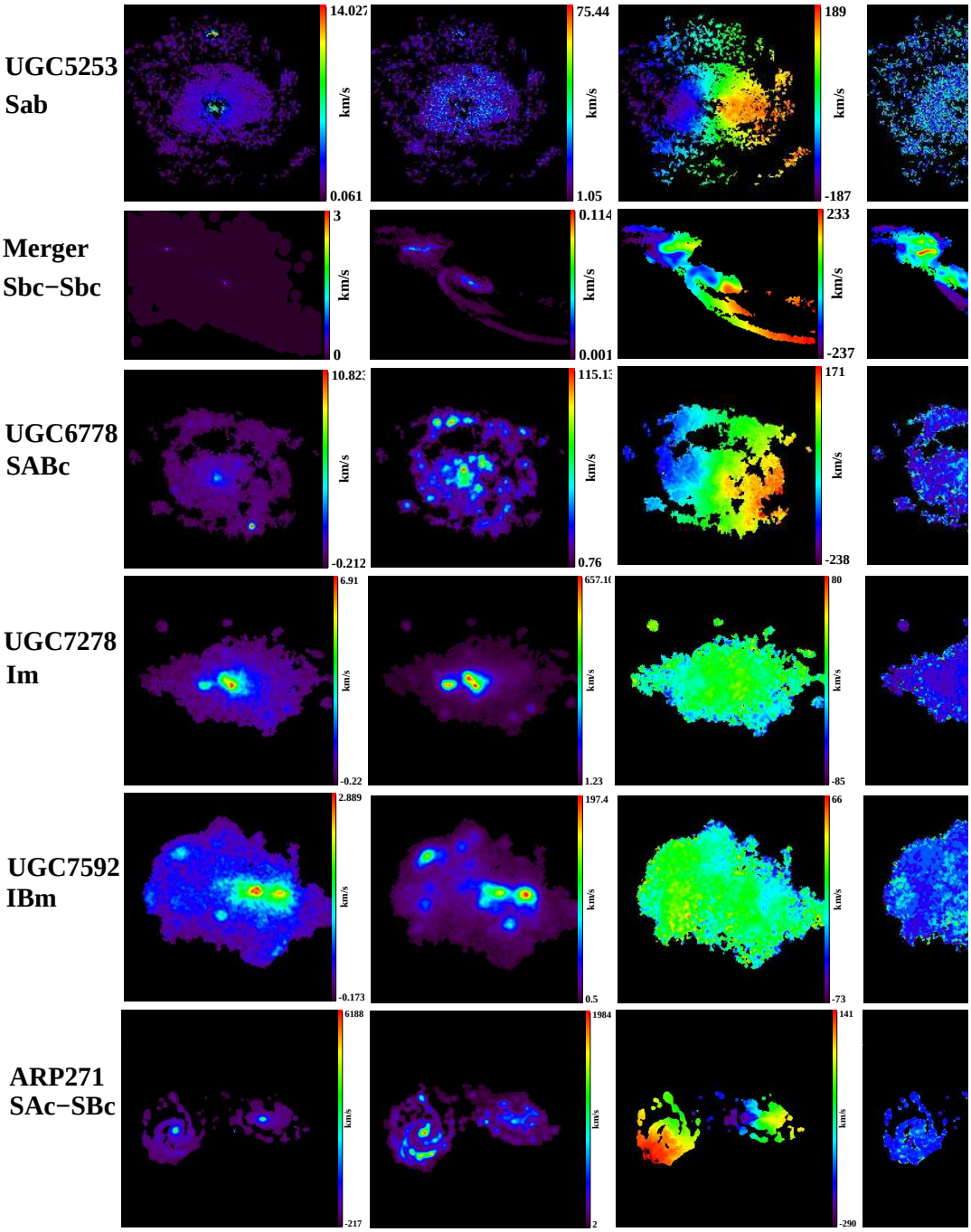

Figure 15.4: Morpho-kinematical templates used for the simulations. From left to right: continuum map, emission line map, gas velocity field and gas velocity dispersion map. For the major merger simulation, the first two maps are the stellar surface density and the gas surface density maps.

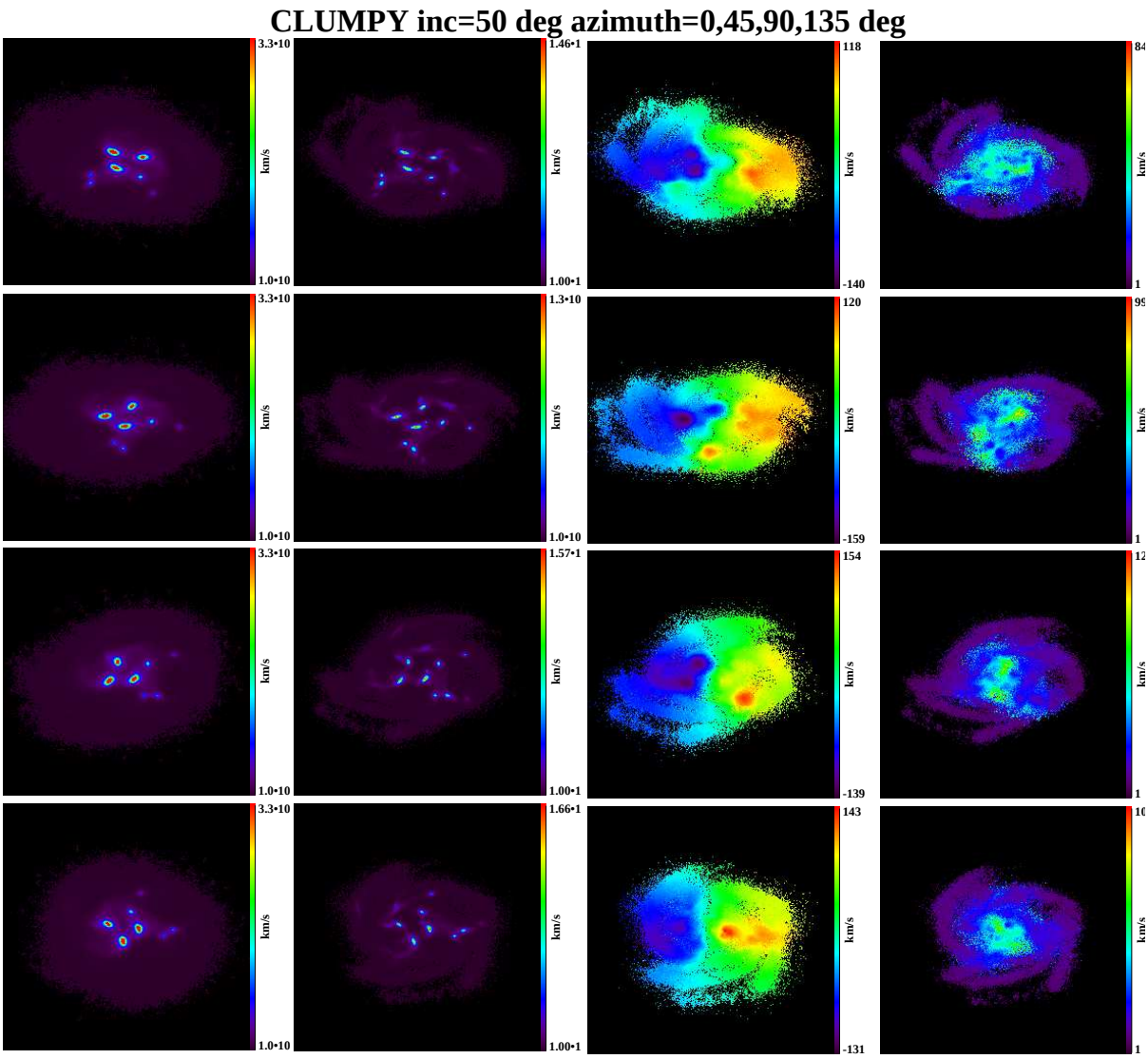

Figure 15.5: Morpho-kinematical templates from simulations of clumpy disks viewed from different azimuthal angles. From left to right: stellar surface density map, gas surface density map, gas velocity field and gas velocity dispersion map.

Table 15.1: Main properties of the morpho-kinematical templates used for the simulations. The first five templates are real observations from Garrido et al. (2002, 2004) and Fuentes-Carrera et al. (2004), while the two last are hydrodynamical simulations from Cox et al. (2006) and Bournaud et al. (2007).

Name	Morphological type	Inclination (deg)	M_B or $M_{\rm stellar}$	z	Comments
UGC5253	Sab	40	-20.8	0.00441	Rotating disk
UGC6778	SABc	30	-20.6	0.003226	Rotating disk
UGC7278	Im	44	-17.1	0.00097	No rotation
UGC7592	IBm	64	-17.8	0.00069	No rotation
ARP271	SAc-SBc	59/32	$-20.6/-21.2$	0.0087	Merging pair
Major merger	Sbc-Sbc	69	$5\times10^{10}\ M_\odot$	0.0	Simulation
Clumpy disks	—	50	$3\times10^{10}\ M_\odot$	1.0	Simulations

Table 15.2: Redshifts (and look-back times) considered for simulations, with the corresponding emission line targeted. The last column gives the broad-band filter within which the redshifted emission line falls.

z	Look-back time (Gyr)	Emission line	Observed broad-band
2.0	10	Hα λ6563Å	K
4.0	12	[OII]λ3727Å	H
5.6	12.6	[OII]λ3727Å	H

Table 15.3: Stellar mass, K-band magnitude, velocity gradient and size of the simulated galaxies as a function of redshift.

z	Quantity	0.1M_*	0.5 M_*	1.0 M_*	5.0 M_*	10.0 M_*
2.0	$\log(M_{\rm stellar}/M_\odot)$	10.2	10.9	11.2	11.9	12.2
	$K_{\rm AB}$	24.7	23.0	22.2	20.5	19.7
	ΔV (km/s)	160	210	260	350	430
	Size (arcsec)	0.68	1.19	1.52	2.67	3.40
4.0	$\log(M_{\rm stellar}/M_\odot)$	9.7	10.4	10.7	11.4	11.7
	$K_{\rm AB}$	26.8	25.1	24.3	22.6	21.8
	ΔV (km/s)	130	180	200	300	330
	Size (arcsec)	0.33	0.59	0.75	1.3	1.7
5.6	$\log(M_{\rm stellar}/M_\odot)$	8.9	9.6	9.9	10.6	10.9
	$K_{\rm AB}$	27.8	26.0	25.3	23.5	22.8
	ΔV (km/s)	90	110	140	200	240
	Size (arcsec)	0.28	0.50	0.63	1.11	1.41

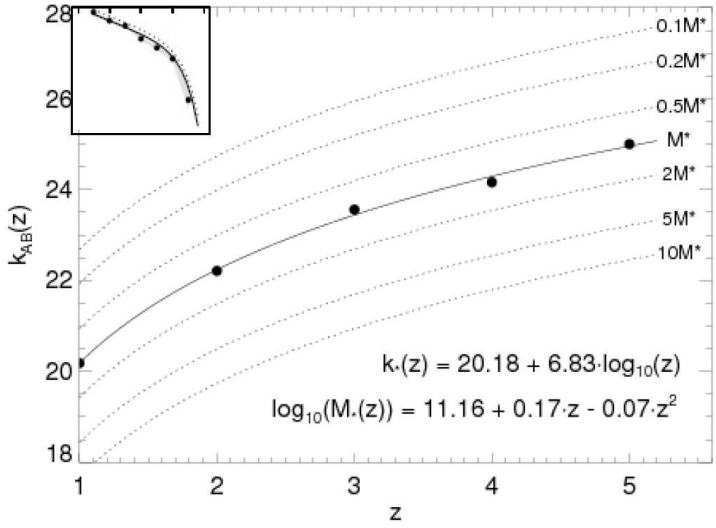

Figure 15.6: Empirical relations between the observed K-band magnitude and redshift used to rescale the morpho-kinematical templates to make them match the continuum flux of distant galaxies at the corresponding stellar-mass. Inset: Illustration of the fit of a Galaxy Stellar Mass Function by a Schechter function. The knee of the curve is called M_*; its variation with redshift is given by the formula on the bottom. The upper formula gives the evolution of the K-band magnitude of an M_* galaxy as a function of redshift.

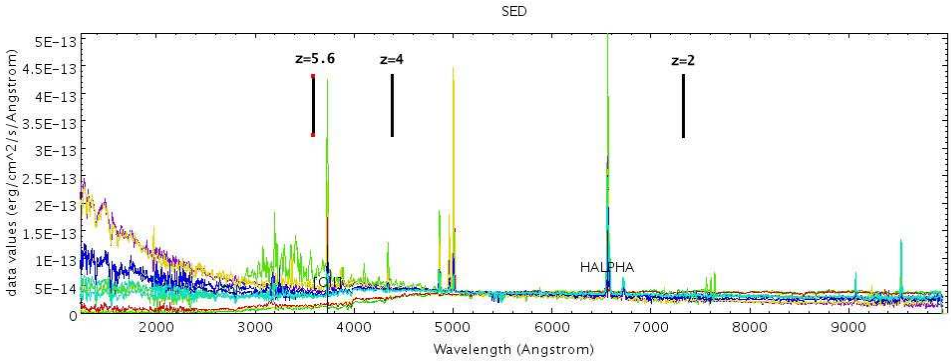

Figure 15.7: Averaged Spectral Energy Distribution for different morphological classes of local galaxies. From top to bottom (in the UV): starbursts with $E(B-V) \leq 0.1$, starbursts with $0.11 \leq E(B-V) \leq 0.21$, starbursts with $0.25 \leq E(B-V) \leq 0.35$, starbursts with $0.39 \leq E(B-V) \leq 0.50$, starbursts with $0.51 \leq E(B-V) \leq 0.60$, Sc galaxies, starbursts with $0.60 \leq E(B-V) \leq 0.40$, Sb galaxies, Sa galaxies and S0 galaxies (Kinney et al. 1996). The [OII] and Hα emission lines are labelled in the bottom of the spectra. The rest-frame position corresponding to the central wavelength of the K-band (2.2 μm) is indicated at the top, for different redshifts. The difference between the pseudo-continuum around the two emission lines and that corresponding to the rest-frame K-band is smaller than a factor two in flux for types latter than Sa.

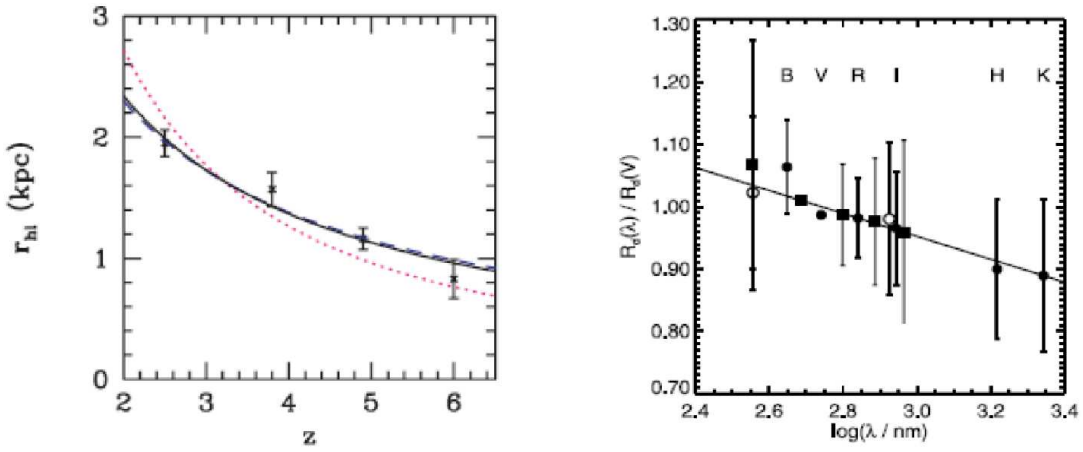

Figure 15.8: Empirical relations used to rescale the morpho-kinematical templates to make them match the size of distant galaxies at the correct observed wavelength. Left: relation between size and redshift from Bouwens et al. (2004). Right: relation between size and wavelength from Barden et al. (2005).

Size: We used empirical relations between redshift and half-light radius from the literature (see Fig. 15.8). To mitigate the impact of the different sample selection criteria used at high redshift, we average the different values found in the literature (Bouwens et al. 2004; Ferguson et al. 2004; Dahlen et al. 2007). The resulting "mean" half-light radius was then k-corrected using the empirical relation of Barden et al. (2005) (see Fig. 15.8) and re-scaled to the assumed stellar-mass using the local scaling between the K-band luminosity (used as a proxy for stellar mass) and size reported by Courteau et al. (2007), i.e., $R_{\rm half} \propto L_K^{0.35}$. The total size (diameter) was assumed to be four times the half-light radius. Note that in the case of a exponential thin disk, $R_{25} = 3.2R_d$ and $R_{\rm half} = 1.68R_d$, which approximately leads to $R_{25} \sim 2R_{\rm half}$. The diameters (in arcsec) used in the simulations are given in Table 15.3.

Velocity gradient amplitude: The velocity amplitude of the velocity field was rescaled using the local stellar-mass Tully-Fisher Relation, following Hammer et al. (2007) (see Table 15.3). Note that strictly speaking, this procedure is quantitatively correct only if applied to rotating disk morpho-kinematical templates.

15.4.3.2 Technical data

All simulations were performed before the technical data descibred in [RD1] had been assembled. Therefore, there are a few differences between [RD1] and the technical inputs actually used in the simulations. The main differences are listed below.

Site and background: A Paranal-like site is assumed with $T_{\rm site} = 280$ K. The atmospheric absorption for this site is modelled as in [RD1], although at a lower spectral resolution than the official current model. Sky emission (continuum and OH lines) was accounted for using a model from Mauna Kea, which also includes zodiacal emission, thermal emission from the atmosphere and an average amount of moonlight. Compared to the data for a Paranal-like site in [RD1], this Mauna Kea model is two times fainter in the H-band. In the K-band, the Mauna Kea sky background model falls in between the Paranal-like and the High & Dry models of [RD1]. The influence of the sky background will be discussed in Section 15.5.4.2.

Telescope: As described in [RD1] for a protected Ag+Al coating with an emissivity of 5%. However, we have neglected the primary mirror's central obscuration. Hence the collecting area of the telescope is assumed to be 8% larger than in [RD1].

Instrument: We assumed a reference pixel size of 50 mas and a reference spectral power of resolution of R=5000, as a compromise between our desire to minimise the impact of the OH sky lines and not wanting to over-resolve the line by a large factor. The instrument thermal background was modelled using two

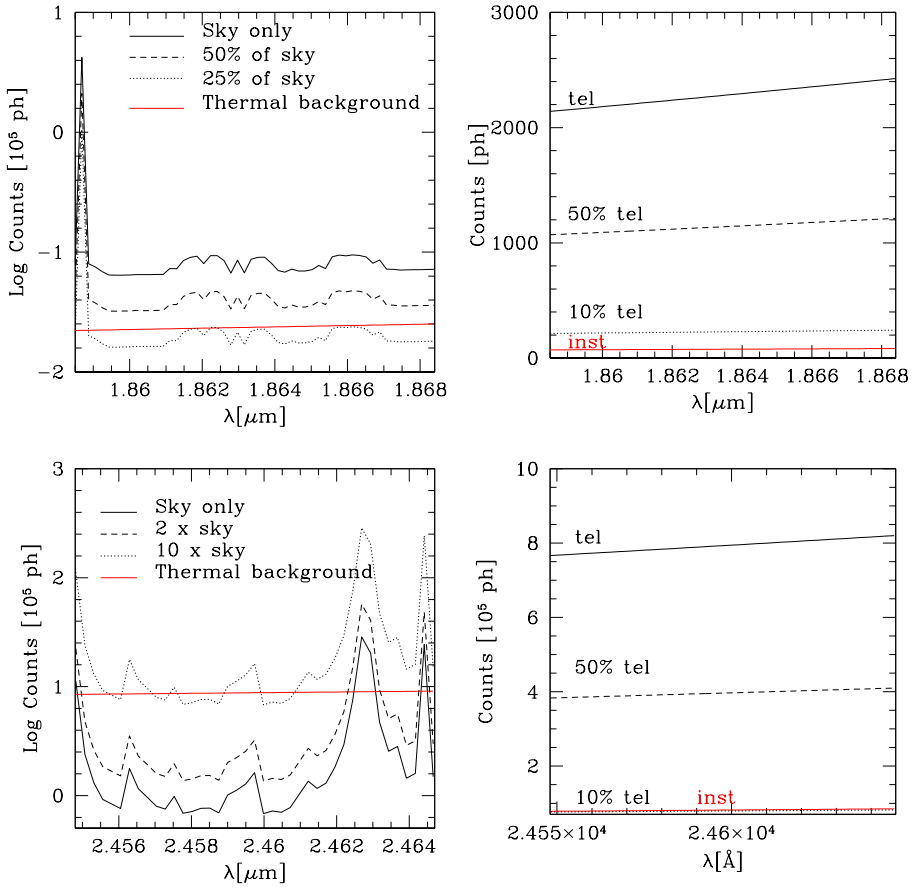

Figure 15.9: Contributions to the total background in the simulations, at z=4 (upper line) and z=5.6 (bottom line). Left column: Comparisons between the sky contribution (continuum and OH sky lines, in black) and the total thermal background (telescope and instrument, in red). Right column: Comparison between the telescope (black) and instrument (red) contributions to the total thermal background.

grey bodies. The first one models the effect of the Target Acquisition System and assumes a temperature of 240 K and an emissivity of 15%. The second one models the effect of the spectrograph and assumes a temperature of 150 K and an emissivity of 69%. These numbers were derived following a preliminary study of the thermal background of EAGLE (Laporte et al. 2008). The instrument background represents less than 10% of the telescope background in both the H and K bands (see Fig. 15.9), in agreement with the official requirements for EAGLE, which specify that "the instrument thermal background at the detector shall be less than 50% of that from the telescope" (goal=10%). The total background is dominated by the thermal contribution of the telescope in the K-band and by the sky background in the H-band. Therefore, we did not explore variations in the thermal contributions from the instrument, which is never the dominant source of background.

Detector and exposure time: We rely on the description of a cooled Rockwell HAWAII-2RG IR array working at ∼80 K by Finger et al. (2006), with a read-out noise of 2.3 e^-/pix and a dark current of 0.01 e^-/s/pix (corresponding values in [RD1]: 3 e^-/pix and a dark current of 0.0011 e^-/s/pix). Its thermal background, bias and saturation threshold were neglected. Because observations of distant galaxies in the NIR are generally not limited by the detector noise, we did not explore parameters that have a non-negligible influence only in such a regime, i.e., variations of the individual frame exposure time dit (we chose $dit = 3600$ s) and detector characteristics ($dark$ and RON). We assumed a reference case exposure time of $T_{\rm intg} = dit \times ndit = 24$ hours, i.e., $ndit = 24$.

Table 15.4: Sets of GLAO and MOAO PSFs generated for the simulations. The asterism size (dubbed "FoV", i.e., the position of guide stars) is varied in order to produce a range of EE sampling typical corrections. All values were measured in the H-band with a seeing of 0.8 arcsec.

GLAO FoV(')	EE in 100 mas	MOAO FoV (')	EE in 100 mas
1	15.00	0	63.59
2	12.51	0.25	56.13
5	8.15	0.5	45.62
10	6.04	1.0	33.73
15	5.31	2.0	27.24
		3.0	24.09
		4.0	23.06
		5.0	22.67

Global throughput: We assumed a global throughput $t_{\rm trans} = 20\%$, detector QE included. [RD1] lists a telescope transmission of 90–95% and a QE of 90% in our wavelength range of interest. The official requirements for EAGLE specify a throughput of at least 35%, including the detector QE. Therefore, according to the baseline, the global throughput should be 31%, which is larger than our assumption. This will be further discussed in Section 15.5.5. The integrated number of photons reaching the detector is a degenerate function of some parameters which have no impact on the spatial or spectral resolution, i.e., $T_{\rm intg}$, $t_{\rm trans}$, EW_0, and D. Therefore, we chose not to explore variations in global transmission, which can be directly derived by analogy with variations of associated degenerate parameters.

Adaptive Optics and PSFs: The coupling between the AO system and the 3D spectrograph is captured through the AO system PSF. Therefore, it is a crucial element that needs to be carefully simulated and cannot be approximated by, e.g., a simple Gaussian. Given the multiplex requirements for the present science case, only GLAO and MOAO have been considered. All PSFs include the effect of telescope aperture, but neglect the effects of the primary mirror's segmentation and of the spiders. In addition, the MOAO PSFs also neglect the the central obscuration. They were generated at the central wavelength of the corresponding filter (e.g., 1.65 μm for the H-band) and the difference between this wavelength and the actual wavelength where the emission line is observed (e.g., 1.86 μm for [OII] at z=4) was neglected.

GLAO PSFs were taken from [RD1]. Note that these PSFs correspond to a total integration time of 4 s and therefore include non-negligible speckle noise.

MOAO PSFs were simulated using an analytical code (calibrated against an end-to-end code) by B. Neichel (GEPI-Obs. de Paris/ONERA) and T. Fusco (ONERA), which has the advantage of producing PSFs free of speckle noise, i.e., more representative of long-exposure PSFs (Neichel et al. 2008). Briefly, we used assumptions as close as possible to the assumptions used for the GLAO PSF modelling: the pitch (inter-actuator distance) was assumed to be 0.5 m ($\sim$84$\times$84 actuators in the pupil plane), with a reference seeing of 0.8 arcsec and an outer scale of the turbulence $L_0 = 25$ m. The MOAO PSF modelling also used the same C_n^2 profile consisting of 10 turbulent layers as was used for the GLAO PSF modelling (see [RD1]). The wavefront was measured using three guide stars (assumed to be natural guide stars, i.e., specific issues related to laser guide stars like, e.g., the cone effect, were neglected), located on an equilateral triangle asterism.

In order to sample different performances for the MOAO and GLAO systems, different asterism sizes were considered (see Table 15.4). The PSFs were always estimated on-axis. It is beyond the scope of this report to study in detail the coupling of the MOAO with 3D spectroscopy; such a study can be found in Puech et al. (2008b). Of interest here is that it is the IFU pixel scale that determines the spatial resolution because the PSF FWHM is smaller than twice the IFU pixel scale. Improving the MOAO correction further does not provide any gain in spatial resolution but still improves the ensquared energy (EE) in a spatial resolution element, hence the SNR. This justifies the usual choice of characterising the EE in an aperture equal to twice the pixel size.

We chose reference PSFs for both AO systems as close as possible of the middle of the range of EE performances. Hence, the H-band MOAO reference PSF has EE=45.62% (in 100 mas), while the H-band GLAO reference PSF has EE=8.15% (in 100 mas). Unless stated otherwise (see Section 15.5.1.2), these two PSFs are those used in the simulations.

15.4.4 Outputs

For each simulation, the following FITS files were produced:

- A data-cube at high spatial resolution and high spatial sampling: this is the original data-cube produced from the set of four maps, once it has been rescaled in terms of size and flux (see Section 15.4.2.2).
- A data-cube at high spatial sampling but with degraded spatial resolution: identical to the previous data-cube but with all spectral channels convolved by a PSF. Hence, its spatial resolution is degraded accordingly to the simulated AO system performances.
- An IFU data-cube: this is the final product of the simulation, which corresponds to mock observations.
- An IFU data-cube without noise and PSF effects.
- A total background spectrum (sky continuum, OH sky lines and total thermal background).
- A thermal background spectrum.
- A SNR data-cube, which gives the expected SNR for each pixel of the IFU data-cube. This spectroscopic SNR is derived as follows:

$$\mathrm{SNR}(i_x, j_y, k_\lambda) = \frac{O(i_x, j_y, k_\lambda)\sqrt{ndit}}{\sqrt{O(i_x, j_y, k_\lambda) + S(i_x, j_y, k_\lambda) + \mathrm{RON}^2 + dark}}, \tag{15.1}$$

  where $O(i_x, j_y, k_\lambda)$ and $S(i_x, j_y, \lambda)$ are respectively the object and sky flux per dit (after accounting for atmospheric transmission) in the spatial position (i_x, j_y) of the data-cube (in pixels) and at the spectral position k_λ along the wavelength axis (in pixels). In the following, the "maximal SNR in the emission line in the pixel (i_x, j_y)" refers to $\mathrm{MAX}_{k_\lambda}[\mathrm{SNR}(i_x, j_y, k_\lambda)]$ and the "spatial-mean SNR" refers to the average of this quantity over the galaxy diameter.

The main parameters used for a given simulation were recorded in the headers of the corresponding FITS files.

In a second step, for each simulated data-cube, the analysis pipeline produced the following FITS files:

- A data-cube that contains all the fits to the emission lines.
- An emission line flux map.
- A velocity field.
- A velocity dispersion map.
- A map of the kinematic $\mathrm{SNR}_{\mathrm{kin}}$.
- A map of the maximal SNR, i.e., the SNR at the peak of the emission line within each spaxel of the simulated IFU data-cube. A spatial-mean value is recorded in the header (see above).

Some example data are available at http://www.eso.org/sci/facilities/eelt/science/drm/C10/.

15.5 Results of simulations

15.5.1 Simulation runs

Unless stated otherwise, parameters are by default set to their reference values as described in Section 15.4.3.

15.5.1.1 Influence of AO correction

We explore the influence of the AO correction by increasing the EE in a given aperture as described in Section 15.4.3.2 (see Table 15.4), for two types of AO systems, i.e., GLAO and MOAO. Simulations were done for all morpho-kinematical templates in the MOAO case and for the two UGC5253 and Major merger templates in the GLAO case.

15.5.1.2 Influence of technical and physical parameters

Simulations were performed for five distinct stellar masses, as described in Table 15.3, and in two different runs, as described in Tables 15.5 and 15.6.

15.5.2 Analysis

15.5.2.1 Influence of AO correction

The spatial-mean SNR obtained for the reference case (H-band) as a function of Ensquared Energy in a 100 mas aperture is plotted in Fig. 15.10. This figure shows that all simulations approximately fall along the same curve, independently of the AO type, although different codes (end-to-end vs. analytical) were used (see Section 15.4.3). This means that all PSFs can be safely compared. For comparison, to the GLAO and MAO PSFs described in Section 15.4.3, we have added MCAO PSFs (with 3 DMs and asterisms of 0.5 and 5 arcmin with corrections derived at 0, 0.5, 2 and 2.5 arcmin with relative to the centre of the FoV) and LTAO PSFs (FoV of 45 arcsec, on axis). The PSFs picked up as representative performances of the GLAO and MOAO modes have been encircled using a red ellipse (see Section 15.4.3). In the H-band, they can be considered as well representative of the AO performance within a range of $\pm \sim$2 in SNR for MOAO and $\pm \sim$3 for GLAO. A systematic study of the impact of EE on the kinematics of distant galaxies can be found in Puech et al. (2008b). Of interest here is that the reference MOAO PSF, with an EE of 45.62% in an aperture of 100 mas is above the minimal requirements derived in this study.

15.5.2.2 DRM STEP 1: 3D detection

We adopt a lower limit in SNR of $\mathrm{SNR}_{\mathrm{min}} = 5$ (spatial-mean SNR in the emission line, see Section 15.4.4) for the 3D detection of distant galaxies. The SNR obtained as a function of the observational parameters and mass is shown in Fig. 15.11. From this figure one finds that the minimal SNR scales as follows:

$$\mathrm{SNR}_{\mathrm{min}} = 5 \times \left(\frac{T_{\mathrm{intg}}}{24\,\mathrm{h}}\right)^{0.5} \left(\frac{D}{42\,\mathrm{m}}\right) \left(\frac{EW_0}{30\,\mathrm{\AA}}\right) \left(\frac{R}{5000}\right)^{-0.5} \left(\frac{\Delta_{\mathrm{pix}}}{50\,\mathrm{mas}}\right) \tag{15.2}$$

This scaling is the one expected in the background-limited regime. As discussed in Section 15.4.3, T_{intg} and the total instrument transmission (not explicitly explored by simulations) follow the same scaling.

Fig. 15.12 gives the stellar mass that can be reached as a function of redshift for MOAO corrections and a seeing of 0.8 arcsec for $\mathrm{SNR}_{\mathrm{min}} = 5$ (black line). One can see that the limit in stellar mass grows exponentially with redshift:

$$\frac{M_{\mathrm{lim}}}{M_*} \simeq 3.3 \times 10^{-6} \exp\left(\frac{z}{0.37}\right) - 0.1, \tag{15.3}$$

which is due to the exponentially increasing contribution of the thermal background from the telescope in the NIR (see right panel in Fig. 15.12). Noteworthy, the mass limit evolves quite smoothly with redshift

Table 15.5: Technical and physical parameters investigated in the first run of simulations. All other parameters are set to their reference values as described in Section 15.4.3.

z	$\Delta_{\rm pix}$ (mas)	Morpho-kin. Template	EW_0 (Å)	R	D (m)
4	25	UGC5253-Major merger	30	5000	42
	50	UGC5253-Major merger	30	5000	42
	50	UGC5253-Major merger	15	5000	42
	50	UGC5253-Major merger	30	10000	42
	50	UGC5253-Major merger	30	2500	42
	75	UGC5253-Major merger	30	5000	42
	50	UGC5253-Major merger	30	5000	30

Table 15.6: Technical and physical parameters investigated in the second run of simulations. All other parameters are set to their reference values as described in Section 15.4.3.

z	$T_{\rm intg}$ (hr)	$\Delta_{\rm pix}$ (mas)	AO correction	Morpho-kin. template	Seeing (arcsec)
2	8	50	MOAO	ALL	0.8
	24	50	GLAO	ALL	0.8
			MOAO	ALL	0.8-0.95
4	8	50	MOAO	ALL	0.8
	24	50	GLAO	ALL	0.8
			MOAO	ALL	0.8-0.95
5.6	8	50	MOAO	ALL	0.8
	24	50	GLAO	ALL	0.8
			MOAO	ALL	0.8-0.95

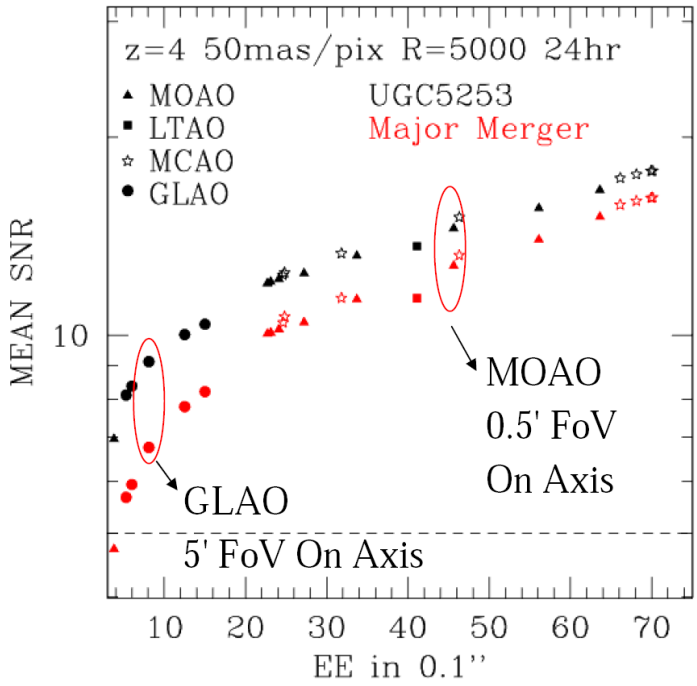

Figure 15.10: Spatial-mean SNR obtained for the reference case (H-band) as a function of Ensquared Energy in a 100 mas aperture for different type of AO systems. The black points correspond to a rotating disk template, while red points correspond to a major merger template. The left most symbols represent the seeing-limited case.

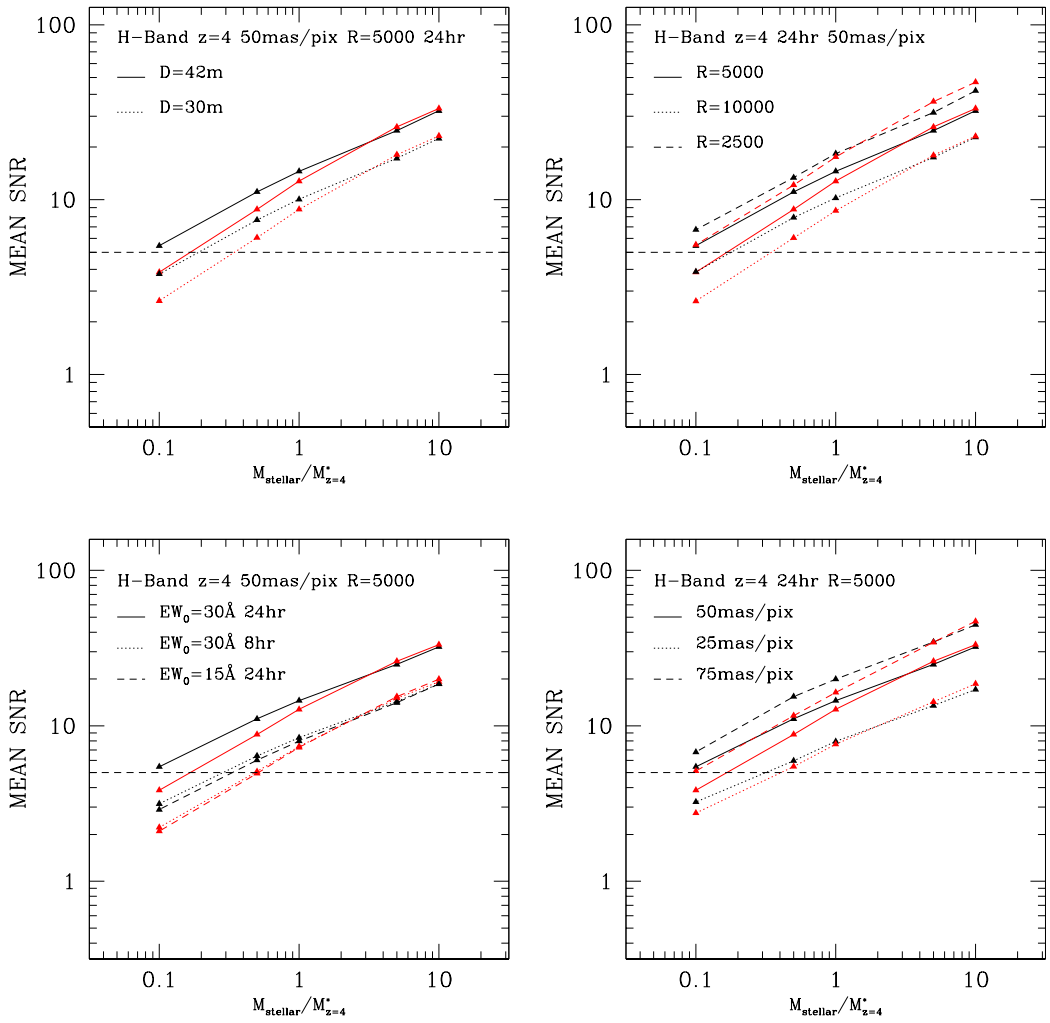

Figure 15.11: Spatial-mean SNR for different stellar masses at z=4 depending on observational parameters. Upper-left panel: influence of telescope diameter. Upper-right panel: influence of spectral resolution. Lower-left panel: influence of integration time and EW_0. Lower-right panel: influence of pixel size. In all panels, the stellar-mass ranges from 0.1 to 10 M_*, black lines represent simulations of a rotating disk, and red lines represent simulations of a major merger.

up to z$\sim$4-4.5: this means that this limit does not depend strongly on input parameters like, e.g., seeing (see the blue curve for a 0.95 arcsec), or AO correction (see the red curve for GLAO corrections). It is not surprising that MOAO and GLAO give similar results, as the SNR considered here is a spatial-mean over the galaxy diameter, which does not take into account differences in terms of spatial resolution (see next section). One can adopt $\mathrm{SNR}_{\mathrm{min}} = 10$ (instead of 5) without impacting strongly the resulting mass limit.

It is useful to summarise Fig. 15.12 by deriving the redshift up to which the Galaxy Stellar Mass Function (GSMF) can be probed down to M_*, as galaxies having such a stellar mass at a given z are those contributing the most to the stellar mass density at this redshift. From Fig. 15.12 one can see that the GSMF can be probed down to M_* up to z=4, independently of the SNR limit, AO correction and seeing conditions.

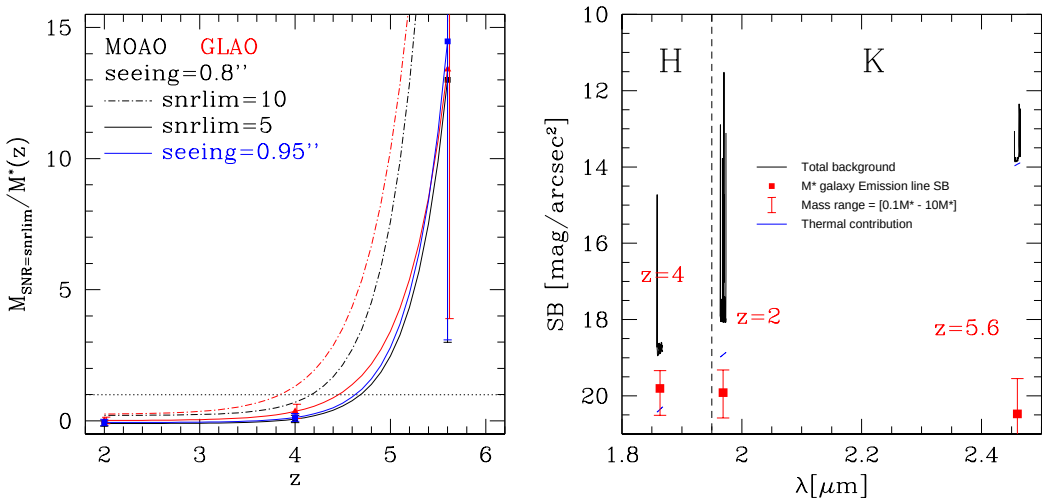

Figure 15.12: Left: Stellar mass limit that can be reach at a given SNR_{min} as a function of redshift, AO correction and seeing. Error-bars represent the range of mass limit obtained when considering all the morpho-kinematical templates. Right: Comparison between the median emission line surface brightness of the simulated galaxies (in red) with that from the thermal background (telescope and instrument, blue lines) and total background (thermal plus continuum and OH sky lines, black lines).

15.5.2.3 DRM STEP 2: Large scale motions

Because we are interested in recovering large-scale motions, it is sufficient to characterise the SNR over the galaxy size, using a simple spatial mean as defined in Section 15.4.4. Using MOAO, it has been suggested that a spatial-mean SNR of 5 is a minimum to recover large-scale motions and distinguish between a major merger and a rotating disk (Puech et al. 2008b). In Table 15.7, we give the SNR obtained for a range of stellar-mass using the reference case and MOAO or GLAO corrections. According to this table, it should be feasible to distinguish between both kinds of templates down to 0.5 M_* at z=4 using MOAO. A direct visual inspection of the simulations confirms this threshold (see Fig. 15.13).

Using GLAO, Table 15.7 suggests that the same distinction could be made down to M_* at z=4 (instead of 0.5 M_* using MOAO). Because GLAO corrections smooth out much more kinematic details than MOAO, it is therefore difficult to distinguish a major merger from a rotating disk down to the same limit in stellar masses/SNR: it is indeed difficult to visually distinguish non-circular motions in the 0.5 M_* major merger simulations at z=4 with GLAO, as shown in Fig. 15.13.

At z=5.6, the much higher background makes any distinction very difficult, and at best limited to the highest mass galaxies, as suggested by Table 15.7 and Fig. 15.14.

Table 15.7: Spatial-mean SNR obtained for simulations of the reference case as a function of mass and redshift, for MOAO and GLAO corrections. The first value corresponds to simulations with a rotating disk template (UGC5253), while the second value corresponds to simulations with a major merger template.

$M_{stellar}$ (in $M_*(z)$)	SNR w/ MOAO z=4	SNR w/ GLAO z=4	SNR w/ MOAO z=5.6	SNR w/ GLAO z=5.6
0.1	5.45-3.85	2.58-2.53	0.78-0.44	0.37-0.18
0.5	11.09-8.80	6.41-4.40	1.49-1.23	0.90-0.60
1.0	15.54-12.76	9.12-6.75	2.07-1.60	1.28-0.86
5.0	24.84-26.07	17.89-16.46	3.92-3.75	2.76-2.20
10.0	32.24-33.27	24.21-23.46	5.14-5.13	3.66-3.15

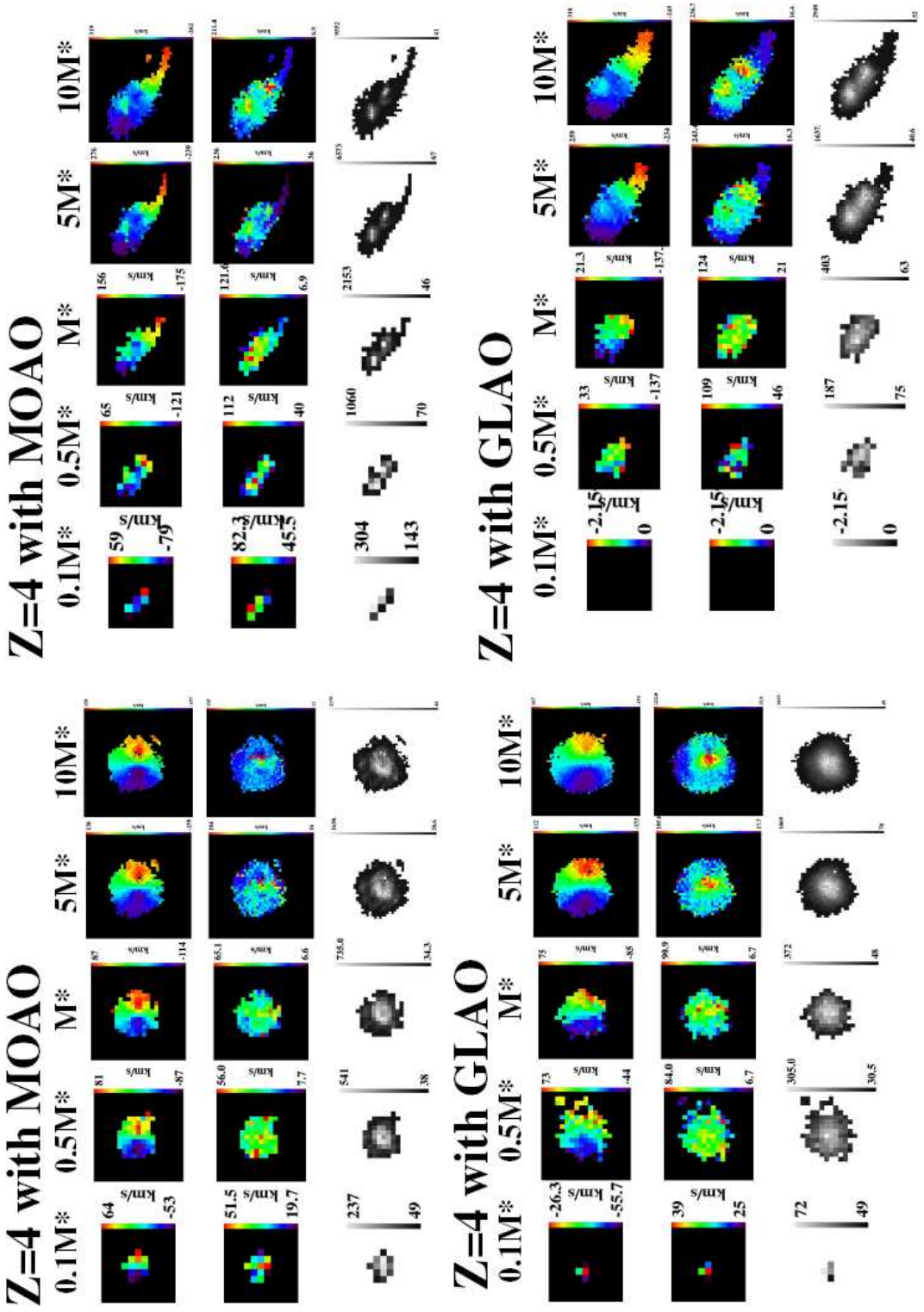

Figure 15.13: Results of simulations at z=4 with MOAO (upper panels) and GLAO (bottom panels). Each panel shows velocity fields (first line), velocity dispersion maps (second lines) and emission line maps (third lines) for different stellar masses (from 0.1 M_*, first column, to 10 M_*, last column). The two panels on the left show simulations using a rotating disk template (UGC5253), while the two panels on the right show simulations for a major merger.

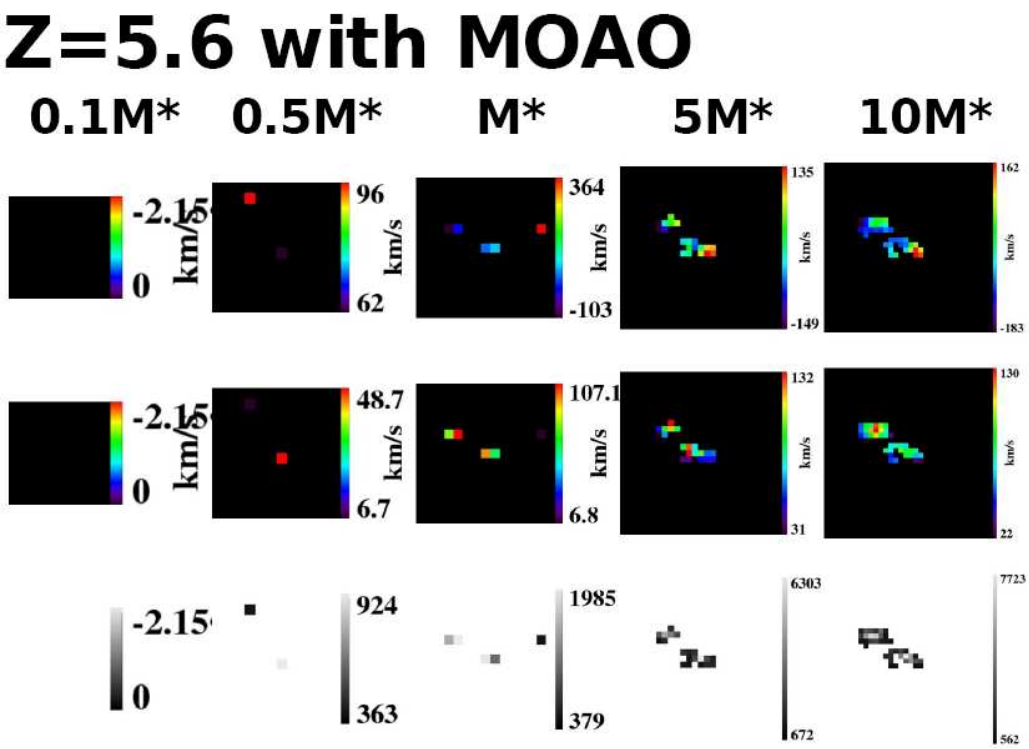

Figure 15.14: Results of major merger simulations at z=5.6 using MOAO. From top to bottom: Velocity fields (first line), velocity dispersion maps (second lines) and emission line maps (third lines) for different stellar masses (from 0.1 M_*, first column, to 10 M_*, last column).

In summary, provided a SNR of 5, it is possible, using MOAO, to recover large scale motions of galaxies and distinguish between different dynamical states at least up to z=4 and down to $M_{\rm stellar} = 0.5\ M_*$. GLAO provides very similar performances, although at lower spatial resolution, which limits the recovery of the dynamical state down to $\sim M_*$ at z=4. At higher redshift, the lost of SNR induced by the increasing contribution of the thermal background of the telescope, limits the recovery of large-scale motions to very massive systems.

15.5.2.4 DRM STEP 3: Rotation curves

In Fig. 15.15, we show rotation curves extracted from simulations at z=2 and z=4 using the UGC5253 rotating disk template. To focus on spatial features, we did not try to fit the inclination and held this parameter fixed during the fitting to the velocity field. We adopted a simple arctan rotation curve, which is fully described by two parameters (Courteau 1997).

At z=2, only MOAO can more or less recover the rising part of the RC. Due to its much lower spatial resolution (FWHM$_{\rm GLAO}$=161 mas and FWHM$_{\rm MOAO}$=11 mas), GLAO induces a beam smearing-like effect, as in HI observations of local galaxies. However, the spatial sampling of the IFU (50 mas/pix) limits the spatial resolution of the observations (to $\sim$100 mas) and does not allow to recover the true shape of the RC. Compared to the best RC that one can recover at this spatial scale (see the black dashed line), MOAO and even GLAO does not induce any further bias in terms of rotation velocity. However, to recover the true rotation velocity, one will need to use numerical simulations, as it is already the case in lower redshift kinematic studies of distant galaxies (see, e.g., Flores et al. 2006; Förster-Schreiber et al. 2006; Puech et al. 2008).

At z=4, the lack of spatial resolution will limit the study of rotation curves to super-M_* rotating disks, as seen in Fig. 15.15. At these redshifts, only MOAO will provide enough spatial resolution to limit biases in recovering the rotation velocity or the rising part of the RC.

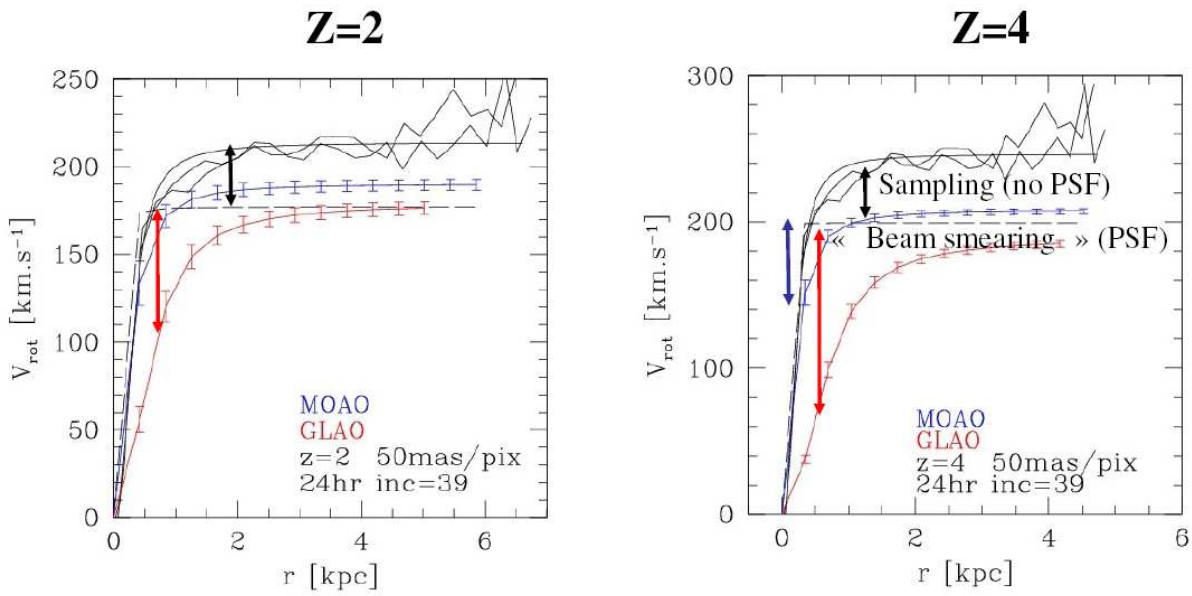

Figure 15.15: Rotation curve of a z=2, M_* galaxy (left panel), and a z=4, $5M_*$ galaxy (right panel). The black lines represent the original z=0 rotation curve rescaled to the size of simulated z=2 and z=4 galaxies; the rotation curve on each side of the galaxy is shown, along with an arctan fit. The rotation curve obtained using MOAO is shown in blue, while the one obtained using GLAO is shown in red. Error-bars represent only the uncertainty associated with the two parameters of the arctan function, and do not take into account uncertainties associated to other kinematic parameters like the dynamical centre or the Principle Axis. The black dashed line represent the rotation curve obtained in simulations without PSF: they correspond to the best rotation curve that one can obtain at the pixel scale of the simulations (50 mas/pix).

In summary, the recovery of the whole shape of the RC will be limited to z$\sim$2, using MOAO. The recovery of the true rotation velocity will require numerical simulations, which is already the case in lower-z studies. At higher redshifts, such measurements will be limited to super-M_* galaxies.

15.5.2.5 DRM STEP 4: Detailed kinematics

We present results of Jeans-unstable clumpy disks in Fig. 15.16. Clumps can clearly be distinguished in galaxies more massive than M_* at z=4, using MOAO (see emission line maps). Using GLAO does not allow to identify these clumps anymore: all clumps are smoothed together in the emission line map. One can still recover non-circular motions in the velocity fields, but without clear morphological signature, it is difficult to identify the underlying cause of this perturbation. At higher redshift, the limited SNR does not allow to recover such clumps anymore, even in very massive galaxies.

In summary, it is possible to recover clumps in rotating disks down to z=4, M_* galaxies using MOAO.

15.5.3 Compliance with figures of merit

The goal of the survey was to study $N_{\rm gal}$ $\sim$1000 galaxies at $2 \leq z \leq 6$ with $0.1 \leq M_{\rm stellar} \leq 5 \times 10^{11}$ $M_\odot$ in less than 100 nights. We adopt the following assumptions as a typical observational strategy for this science case:

- reference case with R=5000 and $\Delta_{\rm pix}$=50 mas (see Section 15.4.3);
- MOAO corrections;
- Mauna-Kea-like background, as described in Section 15.4.3;

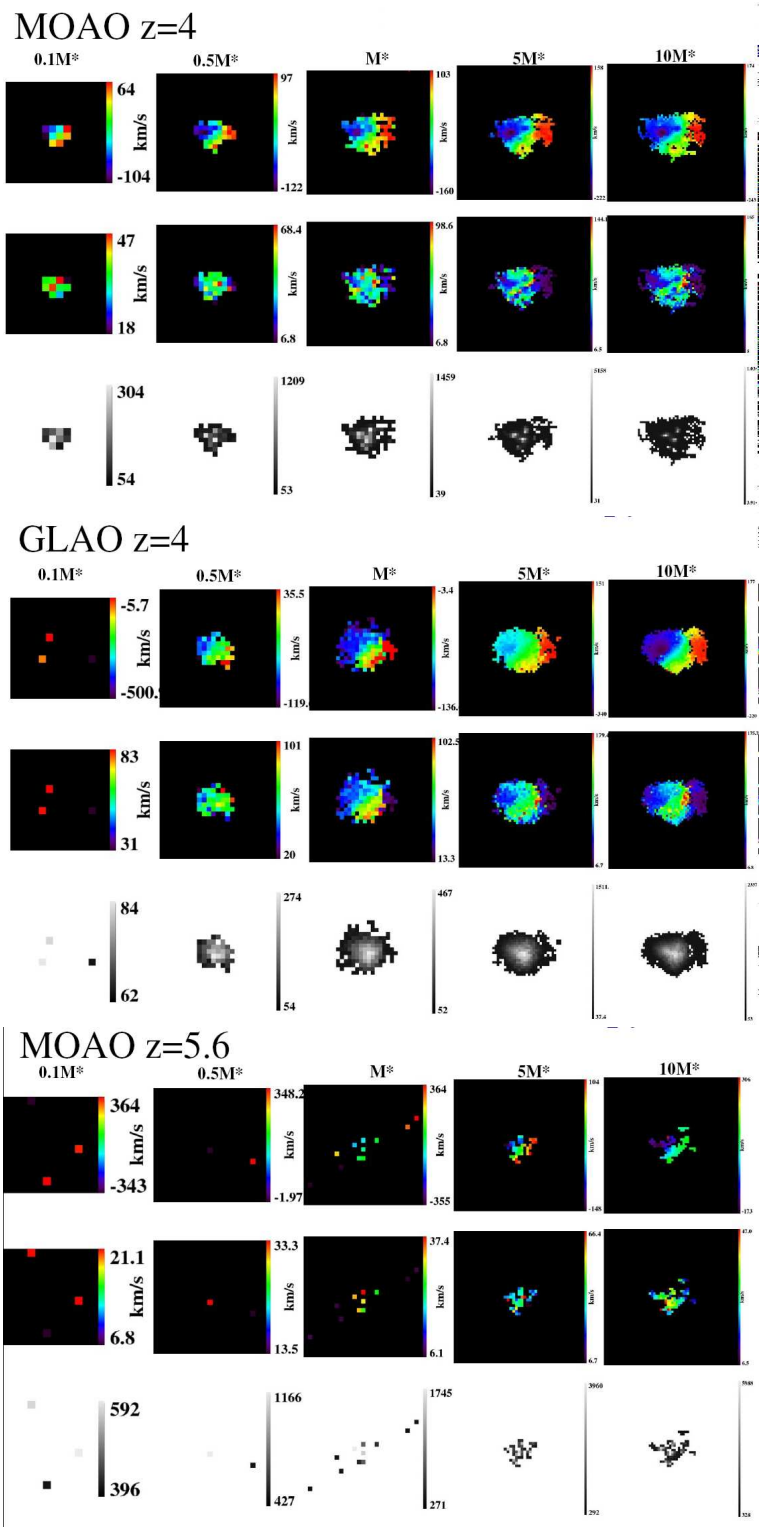

Figure 15.16: Simulations of clumpy disks at z=4 using MOAO (upper panel), GLAO (middle panel) and MOAO at z=5.6 (bottom panel). In each panel, the first line shows the velocity fields, the second line the velocity dispersion maps and the third line the emission line maps, for a range of stellar mass (from 0.1 M_*, first column, to 10 M_*, last column).

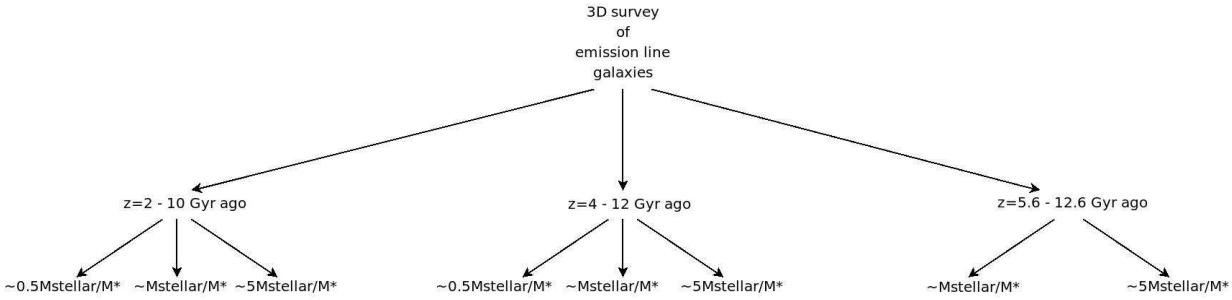

Figure 15.17: Structure of the survey.

- a limiting SNR of 10 (10-σ detection);
- overheads of 30% (i.e., overhead factor OH=1.3);
- three redshift bins (z=2, z=4 and z=5.6) and three mass bins per redshift bin ($\sim$0.5 M_*, $\sim$ M_* and $\sim$5 M_*), except for the z=5.6 bin, which has only two mass bins ($\sim$ M_* and $\sim$5 M_*), since we are interested in galaxies having $0.1 \leq M_{\rm stellar} \leq 5 \times 10^{11}$ $M_\odot$ and that M_*(z=5.6)=0.8$\times$10^{10} $M_\odot$. The structure of the survey is illustrated in Fig. 15.17;
- the number of galaxies per elementary bin is assumed to be $N_{\rm gal}$=100, in order to allow one to derive statistics over the morpho-kinematic types within each bin. This translates into a total number of galaxies in the survey of $N_{\rm gal} \times N_{\rm bins} = 800$;
- within each bin of the survey, the $N_{\rm gal}$=100 galaxies can be observed in $N_{\rm setup}$ setups, with $N_{\rm setup}$=$N_{\rm gal}/M$ where M is the multiplex capability of the instrument. We assume $M \geq 100$, which means that all galaxies in a given bin can be observed in one instrument setup ($N_{\rm setup} = 1$);
- the number of effective observed hours per night is assumed to be 8 hr.

The corresponding integration time, in nights, needed in each elementary bin is given in Table 15.8. The great collecting power of the E-ELT is clearly reflected in the time needed to survey $\sim$667 galaxies in the redshift range 2–4, which requires only $\sim$7 nights. Regarding the z=5.6 bin, the large thermal background of the telescope (see Figs. 15.9 and 15.12) translates into very large integration times.

The total time survey, in nights, can be expressed as follows:

$$T_{\rm survey}(n) \simeq 90 \left(\frac{100}{M}\right) \left(\frac{\rm SNR_{lim}}{10}\right)^2 \left(\frac{N_{\rm gal}}{800}\right) \left(\frac{OH}{1.3}\right) \quad (15.4)$$

In summary, with a total integration time of 90 nights and within the assumptions listed above, it will be possible to:

- Observe 800 galaxies with SNR=10 in three redshift bins (2–4–5.6) and three mass bins (0.5 M_*–M_*–5 M_*), except for the z = 5.6 bin, which has two mass bins (M_* and 5 M_*); all bins have 100 galaxies, which will allow one to do statistics over morpho-kinematic types;

Table 15.8: Integration time needed to complete the survey, in nights, assuming a multiplex of M=125.

T$_{\rm intg}$	0.5 M_*	M_*	5 M_*	Total
z=2	1.2	0.8	0.3	2.3
z=4	2.3	1.4	0.6	4.3
z=5.6	—	66.9	16.39	83.2
Total	3.5	69.1	17.2	89.8

- Recover the large-scale motions of galaxies at least up to z=4 and down to $M_{\rm stellar} = 0.5\ M_*$, i.e., of most galaxies in the survey;
- Recover the detailed kinematics of the most massive galaxies (the detection of clumps in rotating disks will be possible down to z=4, M_* galaxies).

15.5.4 Sensitivity to input parameters

15.5.4.1 Impact of telescope

Diameter: MOAO provides only partial seeing corrections, therefore the resulting PSF is dominated by residual atmospheric perturbations. This implies that the telescope diameter does not directly influence the spatial resolution of observations, but only the achieved SNR. The scaling relation of Section 15.5.2.2 shows that there is no breaking point in telescope diameter: with a smaller, 30 m telescope, one would need two times longer exposures to reach the same SNR.

Thermal background: The telescope is by far the dominant source of background in the K-band. Because the SNR is in a background-limited regime at z=5.6 (see Fig. 15.12), the telescope thus limits the detectability of sources at very high redshift. Note that in the simulations, we have assumed the most optimistic assumption for the emissivity of the 5-mirror E-ELT design, with $\epsilon_{\rm tel}$ =5%.

15.5.4.2 Impact of site

Sky background: The sky is the dominant source of background only for $z \leq 4$ observations. Note that in the simulations, we have used a sky model from Mauna Kea, which is ∼2 times fainter in the H-band that the official DRM Paranal model (see Section 15.4.3.2). As SNR $\propto 1/\sqrt{\rm background}$ in such a regime, a site having two (four) times higher background than the one used in the simulations will reduce the achieved SNR from, e.g., 10 to 7 (5). The impact on high-redshift observations is therefore relatively limited (see Fig. 15.12).

Seeing: The SNR loss between observations with a seeing of 0.8 arcsec and observations with a seeing of 0.95 arcsec is 5-15%. The strongest impact of worse seeing conditions will be to limit the capability to recover rotation curves and detailed kinematics of distant galaxies.

15.5.4.3 Impact of Instrument

GLAO vs. MOAO: GLAO limits the achievable SNR to lower values, compared with MOAO. As a consequence, the recovery of the large scale motions in distant galaxies using GLAO is limited to higher masses. Any further objective of the DRM will be strongly limited in using GLAO instead of MOAO (e.g., rotation curves, detailed kinematics).

Pixel scale: The choice of the IFU pixel scale drives the spatial resolution of MOAO-fed 3D spectroscopy observations, as in the range of EE provided by MOAO, the PSFs have a FWHM largely smaller than twice the pixel size. Therefore, the choice of the optimal IFU pixel scale is related to the optimal "scale-coupling" between the IFU pixel scale and the spatial scale of the physical feature that one wants to recover using this IFU (Puech et al. 2008b). It can be quantified by using the ratio between the size of this feature (here, the galaxy diameter, as one wants to recover the large-scale rotation) and the size of the IFU resolution element. 3D observations of z∼0.6 galaxies with FLAMES/GIRAFFE have demonstrated that a scale-coupling of about 3 is already enough to recover large-scale motions (Flores et al. 2006). It corresponds to the minimum value necessary to ensure that each side of the galaxy is at least spatially sampled by the IFU at the Nyquist rate. In the z=4 simulations, this minimal scale-coupling leads to a pixel scale of ∼55, 125 and 280 mas for 0.1, 1 and 10 M_* galaxies, respectively. Hence, a minimal pixel scale of 50 mas is required if one wants to be able to recover, at least in principle, large-scale motions in z=4 galaxies, and for a large range of stellar mass.

Table 15.9: Accuracy on the velocity (in km/s) and relative accuracy (in %) on the velocity dispersion measurement as a function of spectral resolution (Monte-Carlo simulations).

SNR	3-4	4-5	5-7	7-10	$\geq$10
R=2500	70 km/s - 73%	46 km/s - 59%	31 km/s - 42%	25 km/s - 21%	9 km/s - 9%
R=5000	59 km/s - 50%	38 km/s - 42%	31 km/s - 29%	21 km/s - 20%	8 km/s - 8%
R=10000	54 km/s - 50%	37 km/s - 36%	57 km/s - 26%	20 km/s - 18%	8 km/s - 7%

In Fig. 15.13, we show the velocity fields, velocity dispersion maps and emission line maps extracted from the simulations using a 50 mas pixel scale. One can see that the lowest mass case barely provides enough spatial information to clearly distinguish between the rotating disk and the major merger. Even if the scale-coupling is in principle large enough to properly recover large-scale motions, the pixel scale limits the achieved SNR to relatively low values (see Fig. 15.11). If one only wants to recover large-scale motions in the most massive objects, then an IFU pixel scale of 75 mas can be used, providing more SNR at constant integration time (see Fig. 15.11).

Spectral resolution: Appropriate targets in the NIR are usually selected such as they have emission lines that fall in regions free of strong OH lines. This requires a minimal spectral resolution of $R \sim 3000$ to resolve OH sky lines with enough accuracy. On the other hand, the higher the spectral resolution, the better the accuracy on the recovered kinematics (see Table 15.9): at least R=5000 is required if one wants to recover the velocity dispersion with no more that 50% of relative uncertainty. This value appears to be a good compromise between the desire to minimise the impact of the OH sky lines and not wanting to over-resolve the line by a large factor: Fig. 15.11 clearly demonstrates the interest of having the smallest spectral resolution possible, which optimises the achieved SNR. We estimated that about one third of the H-band is free from strong sky background variations (i.e., larger than 10% of the sky continuum) in continuous windows of at least 200 km/s.[5]

Multiplex capability: In Section 15.5.3 we have assumed an optimal multiplex capability M=125. However, it is likely that the future multi-object integral field spectrograph on the E-ELT, EAGLE, will not provide us with a so large multiplex capability. At constant $T_{\rm survey}$, the relation between the multiplex capability M and the total number of galaxies in the survey is $N_{\rm gal} = 8M$. Therefore, if one assumes $M = 25$, only 200 galaxies can be observed in $T_{\rm survey} = 90$ nights, each bin having 25 galaxies. Such a preliminary survey would certainly provide a very interesting and useful first glimpse of the galaxy mass assembly process as a function of time.

15.5.5 Limitations

The main limitations of this work are:

Variations with wavelength: Most of the inputs are currently considered to be invariant with wavelength (e.g., throughput, emissivity), because the design of the telescope and the instrument were not fully known at the time this word was carried out. Simulations shall be updated once these characteristic curves will be known.

Range of morpho-kinematical templates: The real morpho-kinematical range that one would need to sample to be exhaustive is much larger than the one used for the simulations. However, the exact physical conditions in high-redshift galaxies is far from being well known at the moment. Therefore, we have intentionally limited the simulations to a few well-motivated cases.

Accuracy on the scientific inputs: Most of quantities are not known better than to a factor of two. However, the simulations are internally consistent in the sense that neglected corrections on physical quantities are not larger that this factor. For instance, the empirical relations between stellar mass and

[5]This obviously depends on the spectral resolution and sampling of the sky background. These figures were determined using the Mauna Kea sky model, which has a spectral resolution of 0.4 Å and is sampled at the Nyquist rate. This gives a spectral resolution of $R \sim 22\,500$ at $\lambda = 9000$ Å.

K-band magnitude are uncertain within a factor two (likely more at z≥4), but the flux correction between the rest-frame wavelength corresponding to the K-band and that of the emission line considered is also less than a factor of two (see Section 15.4.3); also, the Tully-Fisher relation appears to evolve smoothly with redshift (a factor of two in mass between z=0 and z∼0.6, see Puech et al. 2008).

Representativeness of the simulations: For simplicity, we have adopted several assumptions that might impact the results of the simulations in terms of SNR (see Section 15.4.3). First, EAGLE specifications suggest that the global throughput might be ∼55% higher than assumed, which would translate into an increase of the SNR by ∼25%, since SNR $\propto \sqrt{t_{\rm trans}}$. Second, we have neglected the central obscuration of the E-ELT, which leads to an overestimate of the SNR by ∼8%, since SNR $\propto D$. Finally, we have assumed that the [OII] doublet is a single emission line instead of a doublet. The average [OII] line ratio is 1.4 (Flores et al. 2006), which leads to an overestimation of the maximal SNR in the emission line by 41% at constant flux (based on the crude assumption of two resolved Gaussians), and in turn leads to an overestimate of the spatial-mean SNR by the same factor. Therefore, one can estimate that all our assumptions lead to an overestimate of the spatial-mean SNR by ∼24%. We have assumed EW_0=30 Å, and since SNR $\propto EW_0$, this means that our simulations would actually correspond to EW_0 ∼37 Å. Therefore, one should consider that the reference EW_0 is 37 Å instead of 30 Å, which was chosen as an extrapolation of the *mean* EW_0([OII]) observed at z∼1 (Hammer et al. 1997). This new value is actually closer to the *median* EW_0([OII]) observed at z∼1 (Hammer et al. 1997): such a value is as relevant as the first one and therefore the adopted simplifications do not impact strongly the results.

15.6 Concluding remarks

We have conducted simulations of the "Physics and mass assembly of galaxies out to z∼6" science case exploring a wide range of observational and physical parameters. We have defined figures of merit for this science case despite the inherent complexity of the science goals, derived empirical scaling relations between the signal-to-noise ratio and the main telescope and instrument parameters, as well as a relation between the limit in stellar mass that can be reached for a given signal-to-noise ratio and redshift. We specifically investigated the impact of AO performance on the science goals. We did not identify any break points with respect to any of the parameters (e.g., the telescope diameter), with the exception of the telescope thermal background, which strongly limits the performance in the highest (z>5) redshift bin. We find that the full range of science goals can be achieved with a ∼100 night program on the E-ELT. We stress that several assumptions and guided guesses had to be made on both the observational conditions and physical characteristics of distant galaxies under study. This introduces an inherent uncertainty, which can be mitigated with future simulations as the telescope design will be consolidated and more details about the physics of high-z galaxies will become available.

References

Amram P., Adami C., Balkowski C., et al., 2002, Ap&SS, 281, 393
Barden M., Rix H.-W., Somerville R.S., et al., 2005, ApJ, 635, 959
Beauvais C., Bothun G., 1999, ApJS, 125, 99
Blais-OUellette S., Amram P., Carignan C., 2001, AJ, 121, 1952
Bournaud F., Elmegreen B.G., Elmegreen D.M., 2007, ApJ, 670, 237
Bournaud F., Daddi E., Elmegreen B.G., 2008, ApJ, 486, 741
Bouwens R.J., Illingworth G.D., Blakeslee J.P., et al., 2004, ApJ, 611, 1
Courteau S., 1997, AJ, 114, 2402
Courteau S., Dutton A.A., van den Bosch F.C., et al., 2007, ApJ, 671, 203
Cox T.J., Jonsson P., Primack J., et al., 2006, MNRAS, 373, 1013
Dahlen T., Mobasher B., Dickinson M., et al., 2007, ApJ, 654, 172
Epinat B., Amram P., Marcelin M., et al., 2008, MNRAS, 388, 500
Ferguson H.C., Dickinson M., Giavalisco M., et al., 2004, ApJ, 600, 107
Finger G., et al., 2006, Nuclear Instruments & Methods in Physics Research A, 565, 241

Flores H., Hammer F., Puech M., et al., 2006, A&A, 455, 107
Förster-Schreiber N., Genzel R., Lehnert M.D., et al., 2006, ApJ, 645, 1062
Fried D.L., 1981, JOSA, 72, 52
Fuentes-Carrera I., Rosado M., Amram P., et al., 2004, A&A, 415, 451
Garrido O., Marcelin M., Amram P., et al., 2002, A&A, 387, 821
Garrido O., Marcelin M., Amram P., 2004, A&A, 349, 225
Genzel R., Tacconi L.J., Eisenhauer F., et al., 2006, Nature, 442, 786
Grazian A., Fontana A., de Santis C., et al., 2006, A&A, 449, 951
Hammer F., Flores H., Lilly S.J., et al., 1997, ApJ, 481, 49
Hammer F., Puech M., Chemin L., et al., 2007, ApJ, 662, 322
Kinney A.L., Calzetti D., Bohlin R.C., et al., 1996, ApJ, 467, 38
Laporte P., Chemla F., Puech M. et al., 2008, Proc. SPIE, 7014, 701463
Neichel B., Fusco T., Conan J.-M., et al., 2008, Proc. SPIE, 7015, 701573
Puech M., Hammer F., Flores H., et al., 2006, A&A, 455, 119
Puech M., Flores H., Hammer F., et al., 2006, A&A, 455, 131
Puech M., Flores H., Hammer F., et al., 2008, A&A, 484, 173
Puech M., Flores H., Lehnert M., et al., 2008, MNRAS, 390, 1089
Roddier F., 1981, Progress in Optics, 19, 281
Sarzi M., Falcon-Barroso J., Davies R.L., et al., 2006, MNRAS, 366, 1151

16 C10-2: High-resolution imaging of high-redshift galaxies

Authors: M. Puech, P. Rosati, S. Toft

16.1 The science case

We propose to take very high resolution imaging of high-redshift galaxies to understand their physical conditions. These observations will be taken at the diffraction limit of the telescope (6, 8, 10 mas in J, H, K for a 42 m telescope). The resolution will be approximately 60 pc in physical length and will be of comparable quality to 1 arcsec imaging of Virgo galaxies. They will give typically more than 100 resolution elements over the galaxies and produce detailed information about the morphology, dynamical state, and variations in physical parameters across the galaxy. These observations will image galaxies beyond (redward of) the Balmer break up to z=4.5, enabling the separation of old and young stars. High contrast observations of QSOs will produce host morphologies, colours, and physical properties. At lower redshifts, the E-ELT will offer exquisite diffraction-limited images of galaxies, which will provide unrivalled details about their morphological structure.

16.2 Goals of the DRM simulations

1. Recovery of morphological properties of high-z galaxies of different masses and types, using E-ELT diffraction limited (DL) images. This will be addressed with incremental complexity as follows: (1) create images of z=4 and z=8 L_* galaxies based on the z=0 “Tadpole” galaxy (UGC 10214/ARP 188), using MCAO corrections; (2) extend the template library to include observed ACS images of local LBGs, which are likely more similar to high star-forming galaxies; (3) include empirical scaling relations between luminosities, masses, sizes, surface brightnesses and morphological k-corrections; (4) constrain the parameter space (mass, z, AO mode, OH suppression efficiency, etc.) in which morphological details (surface brightness profiles, structure parameters, knotty substructures, etc.) can be recovered.
2. Comparison to expected performance of JWST. The simulated images (using GLAO or MCAO corrections) will be compared to simulations of imaging observations with JWST to locate in which areas imaging with the E-ELT can compete with (or is superior to) JWST (e.g., at very high angular resolution of high-z galaxies with significant substructure).

16.3 Metrics / figures of merit

- Recovery of morphological features with DL images: performances will be quantified using structural parameters (e.g., concentration, asymmetry, clumpiness, etc.). We shall require that DL images provide us with sufficient signal-to-noise (S/N) to guarantee useful morphological information. In practice, structural parameters will be compared to those obtained by analysing the morphological templates at the highest spatial resolution (i.e., not degraded by the E-ELT PSF); a threshold on the mean S/N per pixel will be determined empirically, above which structural parameters can be derived in a meaningful way. The sub-parameter space providing useful DL images will be determined.
- Comparison with JWST: it is likely that DL images will not provide enough S/N in the case of very distant galaxies. The JWST/NIRCAM will be the main competitor of the E-ELT in this area, providing images at a 32 mas scale (i.e., sampling the diffraction limit of the 6.5 m primary mirror of JWST). Therefore, a comparison of the spatial resolution and S/N achieved with JWST and the E-ELT at a pixel scale of 30 mas will be conducted. At any redshift, the majority of the stellar mass is locked into M_* galaxies, which makes these objects of particular interest. Results will be therefore quantified using the limit in redshift within which it will be possible to probe the galaxy stellar mass function (GSFM) down to M_*.

16.4 DRM simulations

16.4.1 Methodology

We use the same pipeline as in Section 15. This pipeline produces simulations of data-cubes with two spatial dimensions (α, δ) and one spectral dimension (λ). Images are a simple particular case of 3D data, where the spectral dimension reduces to one large spectral slice, whose width corresponds to the bandwidth of the instrument filter.

Galaxies in the distant Universe are substantially different from nearby galaxies: they are much smaller, brighter in the UV and appear irregular in high-resolution imaging (Elmegreen et al. 2005). Therefore, the simulations will focus on E-ELT observations of UV bright starbursts, i.e., Lyman Break Galaxies (LBGs). The simulations will use high-resolution templates of such galaxies, rescaled in terms of size and flux depending on redshift (see below), and convolved with the proper AO PSF. Realistic sky and thermal backgrounds will also be taken into account.

As a reference, we will assume observations in the H-band, where the influence of the thermal background is minimised in comparison with the K-band. This makes simulations as independent as possible of the telescope design (e.g., number of mirrors), environmental conditions (site selection) and instrument characteristics (e.g., number of warm mirrors), which were not all fully known at the time this work was carried out.

The parameter space will then be explored in order to assess separately the influence of:

- the S/N needed to recover the morphological features of high-z galaxies as a function of integration time, telescope diameter, pixel scale, mass and redshift;
- the influence of the AO correction (MCAO vs. GLAO) at a pixel scale of 30 mas;
- the influence of the background (JWST vs. OH suppression vs. total background).

16.4.2 Pipeline

A complete description of the pipeline is given in Section 15 and Puech et al. (2010). The main steps are summarised here:

- A high-resolution image is produced by interpolating the input template at a sampling rate of $\lambda_{\rm obs}/2D$, where $\lambda_{\rm obs}$ is the central wavelength of the photometric filter and D is the telescope diameter. This sampling rate corresponds to the telescope cut-off frequency.
- This high-resolution image is convolved with a PSF which includes the effects of the telescope aperture and of the AO system. This step produces an image with a degraded spatial resolution, but still sampled at the telescope cutoff frequency.
- The image is spatially re-binned to a given instrument pixel scale. Sky and thermal backgrounds, as well as detector dark current are added.
- Photon and detector noises are generated. In the NIR, the detector noise is a quadratic combination of the dark current photon noise and read-out noise.
- The simulation pipeline generates $ndit$ images with individual exposure times dit. For each individual exposure, a sky frame is simulated separately (i.e., with a different noise realisation). The median value of this frame is used as an estimate of the total background, which is subtracted from the associated "science" frame, reproducing the usual procedure in the NIR. All sky-subtracted images are finally combined by estimating the median of each pixel to simulate several realistic exposures.

Following the above, DL images are analysed using the structural parameters CAS (concentration, asymmetry and smoothness/clumpiness) defined by Conselice et al. (2009). Note that these structural parameters will not be used to classify galaxies into morphological types (e.g., Conselice et al. 2009; Lotz et al.

2008 and references therein), but only as a quantitative metric in order to assess how the morphological information is impacted by degraded spatial resolutions and finite S/N ratios.

16.4.3 Inputs

All magnitudes are given in the AB system.

16.4.3.1 Scientific data

Scientific data are needed in order to (i) provide the simulation pipeline with morphological templates to be used as inputs, and (ii) to rescale these templates in terms of size and flux to realistic values corresponding to distant galaxies.

Redshift: Given the objective of the survey, we have defined a grid of redshift "bins": z=1, 2, 4, 6 and 8. In a concordance cosmology, these correspond to look-back times of 7.7, 10.3, 11.5, 12.1, 12.7 and 13 Gyr.

Morphological templates: The galaxy images that will serve as morphological templates should be sampled as close as possible to 0.03 kpc/pix in order to resolve high-z galaxies at the E-ELT diffraction limit (e.g., at z=4, 0.03 kpc corresponds to 4 mas/pix, i.e., half the diffraction limit in the H-band). Therefore, we use ACS images of local galaxies as morphological templates. These are shown in Figs. 16.1 and 16.2), and Table 16.1 lists their main properties.

Table 16.2 lists the rest-frame central wavelengths corresponding to observations at z=4 and z=8 using the J, H and K broadband filters. Of particular interest is the F330W filter, which directly corresponds to the observed H-band filter for z=4 observations. In principle, in simulations at given redshift, one should

Figure 16.1: Morphological template constructed from ACS observations of the tadpole galaxy in the F475W filter (log scale). Note that the tail was cut off and only the main part of the galaxy was used for this template.

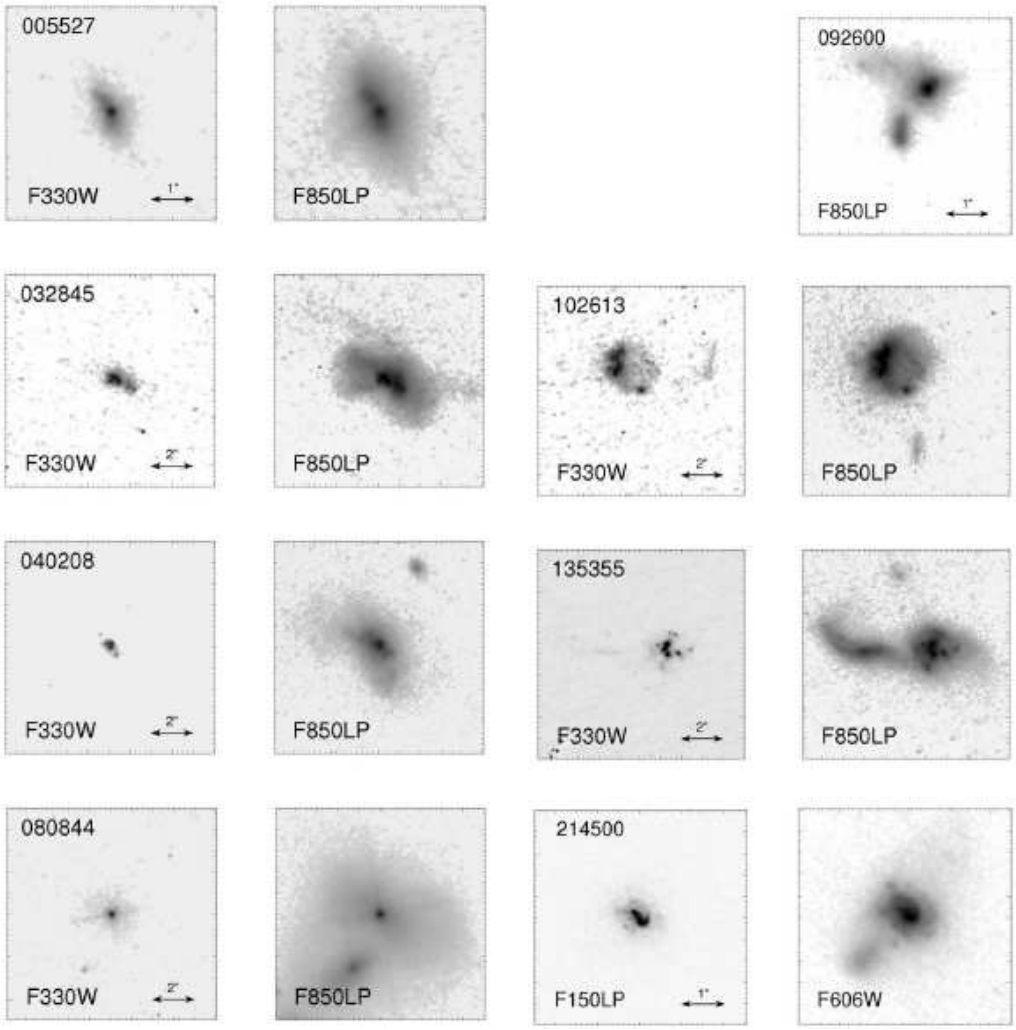

Figure 16.2: Morphological templates created from ACS observations of local LBGs (Overzier et al. 2008).

use the filter that is the closest to the observed band. However, this would complicate the comparison between the morphological information recovered at different redshifts or in different bands. Therefore, we choose to use the F330W images (or the F475W image for the tadpole template) regardless of redshift or observed band.

Stellar mass and flux: We used empirical relations between the K-band magnitude and redshift from observed luminosity functions of high-z galaxies (MUSIC survey, Grazian et al. 2006; see Puech et al. 2010 for more details). In the simulations, we did not apply any colour correction between this K-band estimate and the observed filter, when different of the K-band. In a sample of 33 LBGs at $2.0 \leq z \leq 3.5$, Papovich et al. (2001) found a mean $H - K_S$ colour of 0.42 mag, which means that assuming $H = K$ in the simulations implies that the H-band flux is underestimated by less than 50% in flux on average. Additional evolution in surface brightness is taken into account empirically (e.g., more compact galaxies at high-z). Table 16.3 gives the corresponding stellar masses in the simulations as a function of redshift.

Size: We used empirical relations between size and redshift: (1) $R_{\rm half}$ vs. z in the UV from, e.g., Bouwens et al. (2004) or Ferguson et al. (2004); (2) $R_{\rm half}$ vs. λ "k-correction" from Barden et al. (2005); (3) local L_K vs. $R_{\rm half}$ to scale the size of the templates from Courteau et al. (2007). See Puech et al. (2010) for more details. The total size (diameter) was assumed to be four times the half-light radius. The diameters (in arcsec) used in the simulations are given in Table 16.3.

Table 16.1: Properties of morphological templates observed with HST/ACS and used for the simulations. The third column lists the filters for which images are available. The LBG templates were kindly provided by R. Overzier (Overzier et al. 2008).

ID	z	ACS filters	Spatial sampling [arcsec/pix]
Tadpole	0.031	F475W,F606W,F814W	0.03 [F475W]
LBG0055	0.167	F330W,F850LP	0.05 [F330W]
LBG0328	0.142	F330W,F850LP	0.05 [F330W]
LBG0402	0.139	F330W,F850LP	0.05 [F330W]
LBG1026	0.160	F330W,F850LP	0.05 [F330W]
LBG1353	0.198	F330W,F850LP	0.05 [F330W]

Table 16.2: Rest-frame central wavelengths (in μm) corresponding to the J, H and K broadband filters for observations at the redshifts probed by the simulations.

	J	H	K
z=1	0.63	0.83	1.10
z=2	0.42	0.55	0.73
z=4	0.25	0.33	0.44
z=6	0.18	0.24	0.31
z=8	0.14	0.18	0.24

Table 16.3: Stellar masses, K-band magnitudes and sizes of the simulated galaxies as a function of redshift.

z	Quantity	0.1 M_*	0.5 M_*	1 M_*	5 M_*	10 M_*
1.0	$\log(M_{\rm stellar}/M_\odot)$	10.7	11.0	11.3	12.0	12.3
	$K_{\rm AB}$	22.7	20.9	20.2	18.4	17.7
	Size (arcsec)	1.10	1.93	2.46	4.32	5.51
2.0	$\log(M_{\rm stellar}/M_\odot)$	10.2	10.9	11.2	11.9	12.2
	$K_{\rm AB}$	24.7	23.0	22.2	20.5	19.7
	Size (arcsec)	0.68	1.19	1.52	2.67	3.40
3.0	$\log(M_{\rm stellar}/M_\odot)$	10.0	10.7	11.0	11.7	12.0
	$K_{\rm AB}$	25.9	24.2	23.4	21.7	20.9
	Size (arcsec)	0.50	0.87	1.11	1.96	2.50
4.0	$\log(M_{\rm stellar}/M_\odot)$	9.7	10.4	10.7	11.4	11.7
	$K_{\rm AB}$	26.8	25.1	24.3	22.6	21.8
	Size (arcsec)	0.33	0.59	0.75	1.3	1.7
6.0	$\log(M_{\rm stellar}/M_\odot)$	8.7	9.4	9.7	10.4	10.7
	$K_{\rm AB}$	28.0	26.3	25.5	23.8	23.0
	Size (arcsec)	0.29	0.51	0.65	1.14	1.45
8.0	$\log(M_{\rm stellar}/M_\odot)$	7.2	7.7	8.0	8.7	9.0
	$K_{\rm AB}$	28.9	27.1	26.4	24.6	23.9
	Size (arcsec)	0.29	0.51	0.66	1.15	1.47

Table 16.4: Assumptions for the surface brightness of the sky emission (in mag$_{\rm AB}$ arcsec^{-2}). The continuum plus OH backgrounds in the J and K bands are extrapolations of the H-band values, assuming a relative increase of 2.1 mag in comparison to the pure continuum brightness.

Background	J	H	K	Comments
Continuum only	18.9	17.9	17.6	Full OH suppression
Continuum + OH	16.8	15.8	15.5	Extrapolation of H-band

16.4.3.2 Technical data

Unless stated otherwise, we used the technical data from [RD1]. Important differences or further assumptions are listed below.

Site and background: We follow [RD1]. A Paranal-like site is assumed but with $T_{\rm site} = 280$ K (instead of 285 K as in [RD1]). The atmospheric transmission is assumed to be the mean transmission over the filter spectral range. Sky emission values are listed in Table 16.4. Values in the K-band are extrapolated from the H-band, since the real background is not well known in this region. The thermal background of the telescope was modelled as a grey body (as in [RD1]). The effective thermal background is assumed to be the mean value of this grey body over the filter spectral range, which leads to a total background of 14.95 mag$_{\rm AB}$ arcsec^{-2} at 2.12 μm. For comparison, the K-band total background in Paranal (sky plus VLT thermal background) is found to be 14.9 mag$_{\rm AB}$ arcsec^{-2} (Cuby et al. 2000); this background corresponds to the Cassegrain focus, which has an emissivity of 6%, which is very similar to the value of 5% assumed here (see below).

Telescope: As described in [RD1] for a protected Ag/Al coating with an emissivity of 5%. However, we have neglected the primary mirror's central obscuration. Hence the collecting area of the telescope is assumed to be 8% larger than in [RD1]. We also consider primary mirror sizes of 30 and 60 m, but only for the purpose of exploring variations in the telescope's photon collecting power, not its resolution.

Instrument: We assume an imaging camera operating in the NIR. The pixel size is one of the parameters that will be varied. We set the thermal background from the AO module and the instrument to 15% and 10% of that of the telescope, respectively (which is in line with the specifications for the MICADO concept study). Because the thermal emission from the instrument is never the dominant source of noise, we choose not to explore any variations in these parameters. Finally, photometric filters were assumed to have a square shape with a width equal to the FWHM of the filter.

Detector and exposure time: We rely on the description of a cooled Rockwell HAWAII-2RG IR array working at ∼80 K, as described by Finger et al. (2006), with a read-out noise of 2.3 e$^-$/pix and a dark current of 0.01 e$^-$/s/pix (corresponding values in [RD1]: 3 e$^-$/pix and a dark current of 0.0011 e$^-$/s/pix). Its thermal background, bias and saturation threshold were neglected. Because observations of distant galaxies in the NIR are generally not limited by detector noise, we did not explore parameters that have a negligible influence only in this regime, i.e., variations of the individual frame exposure time dit (we chose dit=600 s) and detector characteristics ($dark$ and RON). We adopt as a reference a total integration time of 10 hr (= $dit \times ndit$)

Global throughput: According to the official specifications for the MCAO module and MICADO, these devices shall have a global throughput of at least 80% and 50% (detector QE included), respectively. The telescope transmission in the NIR is given as 90–95% in [RD1] for the protected Ag/Al coating. Therefore, we assumed a total throughput of 37% and 45% (independent of wavelength) for MCAO and GLAO observations respectively. The integrated number of photons reaching the detector is a degenerate function of several parameters, i.e., the total integration time $T_{\rm intg}$, the telescope diameter D and the global throughput $t_{\rm transm}$. It is therefore pointless to explore variations due to all these parameters, and in the following only variations due to D will be explicitly explored. One can easily translate the results obtained for D to these other parameters.

Adaptive Optics and PSFs: All PSFs include the effect of telescope aperture, but neglect the effects of the primary mirror's segmentation and of the spiders. In addition, the MCAO PSFs also neglect the central obscuration. They were generated at the central wavelength of the corresponding filter (e.g., 1.65 μm for the H-band), and their variations as a function of wavelength were neglected.

Note that all simulations, even those for a 30 or 60 m telescope, use the PSFs for a 42 m telescope.

GLAO PSFs were taken from [RD1]. Note that these PSFs correspond to a total integration time of 4 s and therefore include non-negligible speckle noise.

MCAO PSFs were simulated using an analytical code (calibrated against an end-to-end code) by B. Neichel (GEPI-Obs. de Paris/ONERA) and T. Fusco (ONERA), which has the advantage of producing PSFs free of speckle noise, i.e., more representative of long-exposure PSFs (Neichel et al. 2008). Briefly, we used assumptions as close as possible to the assumptions used for the GLAO PSF modelling: the pitch (inter-actuator distance) was assumed to be 0.5 m ($\sim$84$\times$84 actuators in the pupil plane), with a reference seeing of 0.8 arcsec and an outer scale of the turbulence $L_0 = 25$ m. The MCAO PSF modelling also used the same C_n^2 profile consisting of 10 turbulent layers as was used for the GLAO PSF modelling (see [RD1]). The wavefront was measured using six guide stars (assumed to be natural guide stars, i.e., specific issues related to laser guide stars like, e.g., the cone effect, were neglected), located on the same asterism shape as the one used for the LTAO PSFs (see [RD1]). The MCAO module was assumed to have 3 deformable mirrors, conjugated to 0, 5 and 12 km above the ground.

Given the multiplex requirements of the DRM proposal, we will consider the Large Field mode of the MCAO module only (2 arcmin diameter). The specifications of the MCAO module require in this case a Strehl ratio (SR) of (24$\pm$13)% in the H-band and (43$\pm$13)% in the K-band. We produced a model PSF with an asterism of 3 arcmin and a line-of-sight located 2 arcmin away from the asterism centre; this PSF has a SR of 28% in the H-band and 41% in K, which makes it well suited for MCAO simulations. For a fair comparison, we will consider a GLAO PSF evaluated on-axis with an asterism of 2 arcmin in diameter.

JWST/NIRCAM model: We followed the description of JWST given on the JWST web pages, with M1=6.5 m, and M2=3.163 m. A description of NIRCAM is given on the NIRCAM web pages, including the definition of filters. We used the finest spatial scale with 31.7 mas/pix. The total background is assumed to be 23.1 mag$_{\rm AB}$ arcsec^{-2} in the H-band, following Windhorst et al. (2002). This background can be considered to be constant as a function of wavelength in the NIR (Windhorst et al. 2002) within $\sim$10%. We assumed a dark current of 0.01 e$^-$/s/pix and a read-out noise of 6 e$^-$/pix, as measured by Garnett et al. (2004). The global throughput is assumed to be 50% (i.e., this assumes 4 reflections in the telescope with a reflectivity of 97%, a NIRCAM throughput of 75%, a detector entrance window transmission of 95% and a detector QE of 80%). PSFs were generated using JWPSF v3 (Cox & Hodge 2006).

16.4.4 Outputs

For each simulation the following FITS files were produced:

- An image at high spatial resolution and high spatial sampling: this is the original image produced from the morphological template, once it has been re-sampled at the telescope cut-off frequency and rescaled in terms of size and flux (see Section 16.4.2).
- An image at high spatial sampling but with degraded spatial resolution: identical to the previous image but convolved with a PSF. Hence, its spatial resolution is degraded according to the simulated AO system performances.
- A sky-subtracted image: this is the final product of the simulation, which corresponds to mock observations.
- A simulated image without noise and PSF effects.
- A SNR image, which gives the expected SNR for each pixel. The SNR is derived as follows:

$$\mathrm{SNR}(i_x, j_y) = \frac{O(i_x, j_y)\sqrt{ndit}}{\sqrt{O(i_x, j_y) + S(i_x, j_y) + \mathrm{RON}^2 + dark}}, \quad (16.1)$$

  where $O(i_x, j_y)$ and $S(i_x, j_y)$ are respectively the object and sky flux per dit (after accounting for atmospheric transmission) in the spatial position (i_x, j_y) of the image (in pixels). In the following, we will use the flux-weighted mean S/N, as determined within the intrinsic galaxy diameter, i.e., irrespective of what parts of the galaxy are detected. For simplicity, we will refer to this quantity simply as the "mean S/N" or even "S/N".

The main parameters used for a given simulation were recorded in the headers of the corresponding FITS files. The CAS structural parameters measured on simulated images were recorded in ASCII files.

16.5 Results of simulations

16.5.1 Simulation runs

Unless stated otherwise, parameters are by default set to their reference values as described in Section 16.4.3, and for all morphological templates.

16.5.1.1 Influence of S/N

We explore the influence of the S/N on our ability to recover morphological features of high-z galaxies as a function of telescope diameter D, pixel scale $\Delta_{\rm pix}$, mass and redshift. Note that all morphological templates were simulated in every case. Simulations are listed in Table 16.5.

In the case of simulations with D=30 m, we did not re-simulate new PSFs but assumed that the detector PSF (i.e., the spatial re-binning at 30 mas/pix) will completely dominate the resulting total PSF (i.e., telescope PSF convolved by the detector PSF).

16.5.1.2 Influence of AO correction

We used a pixel scale of 30 mas to explore the influence of GLAO vs. MCAO corrections at z=2 and z=4 (see Table 16.6). Note that all morphological templates were simulated in every case.

16.5.1.3 Influence of background

We conducted the simulations listed in Table 16.7 to study the influence of the background at constant integration time. All simulations assume a 42 m E-ELT, except JWST simulations. Note that all morphological templates were simulated in every case.

16.5.2 Analysis

16.5.2.1 DL images

Examples of the results of simulations at z=1 and z=6 are shown in Fig. 16.3. It is difficult to build a simple criterion that would allow us to assess and quantify the quality of the recovered morphology. At first order, such a criterion should be a function of the number of spatial resolution elements per galaxy intrinsic diameter, and of the surface brightness detection limit. As a simple first-order criterion, it is possible to use a threshold on the spatial-mean S/N per pixel over the intrinsic galaxy diameter, as defined in Section 16.4.4. This quantity indeed captures the two required dependencies on surface brightness detection and number of spatial elements of resolution. A visual inspection of Fig. 16.3 supports this idea and suggests that a minimal S/N of 5 is required to recover the detailed morphology of distant galaxies using DL images (see how the tail at the bottom-right of the main galaxy component tends to disappear below S/N$\sim$5).

To quantify this more precisely, we derived the CAS structural parameters for all simulations at the DL, regardless of redshift or any other parameter: to be meaningful, these non-parametric measures must be derived above a given S/N threshold (see, e.g., Lisker 2008[6]). Because we are not interested in the absolute values of these parameters (which can be used to classify galaxies into rough morphological classes, see, e.g., Conselice et al. 2009; Lotz et al. 2008), we systematically compared their values to those obtained on the highest resolution templates (i.e., using the original HST images without any degradation due to spatial resolution or noise). Fig. 16.4 compares the derived CAS values as a function of the mean S/N, in all simulations at the DL. This figure confirms that a minimal S/N of 5 is required for the CAS parameters to be consistent.

[6]Although the discussion focuses on the Gini parameter which is not used in the present study, their conclusion remains of general interest for our study.

Table 16.5: Simulations conducted to study the influence of the S/N on the recovery of morphological features of high-z galaxies.

z	$\Delta_{\rm pix}$ (mas)	$T_{\rm intg}$ (hr)	D (m)	Mass ($M_*(z)$)
1	30	10	42	0.1–10
	4	10	42	0.1–10
2	30	10	42	0.1–10
	4	10	42	0.1–10
	30	10	30	0.1–10
3	30	10	42	0.1–10
	4	10	42	0.1–10
4	30	10	42	0.1–10
	4	10	42	0.1–10
	30	10	30	0.1–10
6	30	10	42	0.1–10
	4	10	42	0.1–10
	30	10	30	0.1–10
8	30	10	42	0.1–10
	4	10	42	0.1–10
	30	10	30	0.1–10

Table 16.6: Simulations conducted to study the influence of the AO correction.

z	$\Delta_{\rm pix}$ (mas)	$T_{\rm intg}$ (hr)	AO type	Mass ($M_*(z)$)
2	30	10	GLAO	0.1–10
3	30	10	GLAO	0.1–10
4	30	10	GLAO	0.1–10
6	30	10	GLAO	0.1–10
8	30	10	GLAO	0.1–10

Table 16.7: Simulations conducted to study the influence of the background.

z	$\Delta_{\rm pix}$ (mas)	$T_{\rm intg}$ (hr)	Background conditions	Mass ($M_*(z)$)
1	30	10	JWST (zodiacal light limited) with F150W	0.1–10
2	30	10	JWST (zodiacal light limited) with F150W	0.1–10
3	30	10	JWST (zodiacal light limited) with F150W	0.1–10
4	30	10	JWST (zodiacal light limited) with F150W	0.1–10
6	30	10	JWST (zodiacal light limited) with F150W	0.1–10
8	30	10	JWST (zodiacal light limited) with F150W	0.1–10
2	30	10	Full OH suppression in H-band	0.1–10
3	30	10	Full OH suppression in H-band	0.1–10
4	30	10	Full OH suppression in H-band	0.1–10
6	30	10	Full OH suppression in H-band	0.1–10
8	30	10	Full OH suppression in H-band	0.1–10
2	4	10	Full OH suppression in H-band	0.1–10
3	4	10	Full OH suppression in H-band	0.1–10
4	4	10	Full OH suppression in H-band	0.1–10
6	4	10	Full OH suppression in H-band	0.1–10
8	4	10	Full OH suppression in H-band	0.1–10

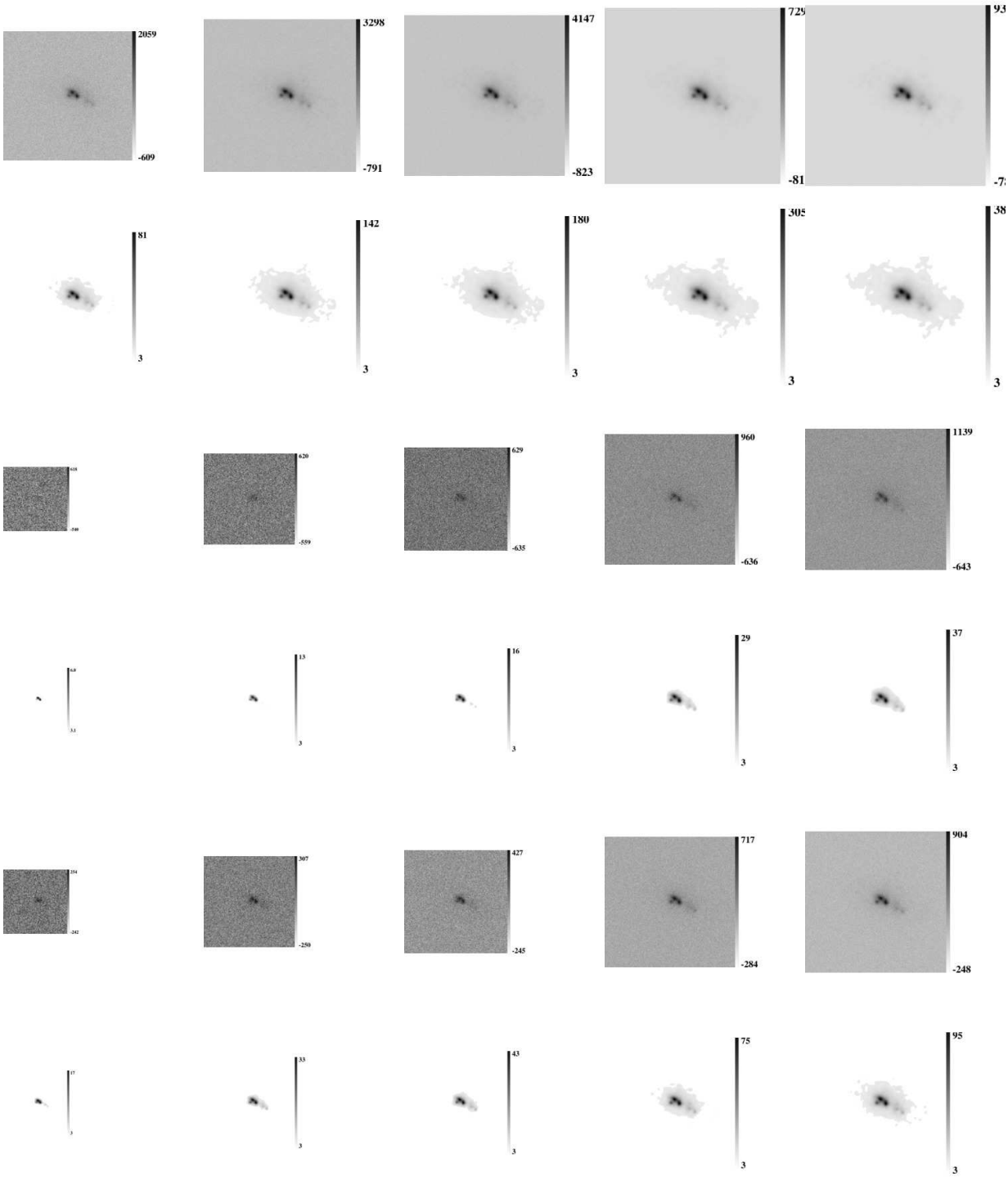

Figure 16.3: Results of simulations using the LBG0328 template with a Paranal-like sky background at z=1 (top two rows), z=6 (middle two rows) and at z=6 but assuming full OH suppression (bottom two rows). From left to right: simulations for stellar masses of 0.1, 0.5, 1, 5 and 10 M_*. First, third and fifth rows: images simulating what would be observed with the E-ELT using MCAO at z=1 and z=6, respectively, after sky subtraction. Second, fourth and sixth rows: expected S/N per pixel. A threshold at S/N=3 has been applied. The mean S/N per pixel over the intrinsic galaxy diameter is, from left to right, 21, 37, 48, 89, 116 at z=1 (second row) and 2, 3, 4, 7, 9 at z=6 (fourth row), and 6, 9, 12, 19 and 25 with OH suppression at z=6 (sixth row).

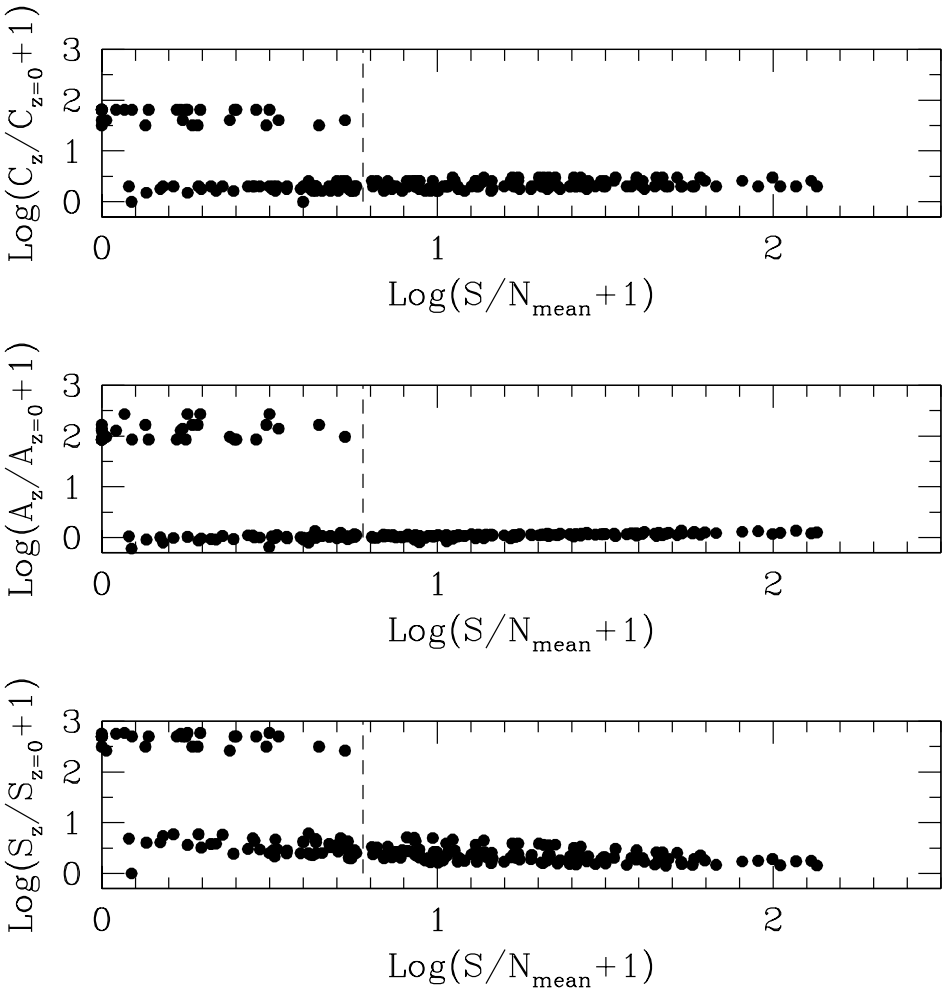

Figure 16.4: Structural parameters derived in simulated DL images (normalised to the parameters in the original high resolution images) as a function of S/N. The vertical dashed lines represent the threshold above which structural parameters can be safely measured, which corresponds to S/N=5. Below this threshold, the high level of noise can severely affect the estimation.

Fig. 16.5 shows the stellar mass that can be reached as a function of redshift using MCAO DL images with a minimal S/N of 5. In this plot, error-bars represent the range of stellar mass derived considering all the morphological templates. We adopted the middle of this range as a simple estimator of the stellar mass one can reach at a given redshift under the specific observational conditions considered in this plot.

The limit in stellar mass grows exponentially between z=1 and z=8 as:

$$\frac{M_{\rm lim}}{M_*} \simeq -0.7 + 0.4 \exp\left(\frac{z}{1.7}\right). \quad (16.2)$$

This increase is due to a fast decreasing surface brightness in the simulated galaxies as a function of redshift (see right panel of Fig. 16.5). In conclusion, these simulations suggest that the galaxy stellar mass function can be probed using DL images down to M_* up to z$\sim$2.5. With full OH suppression, this limit shifts to z$\sim$5.

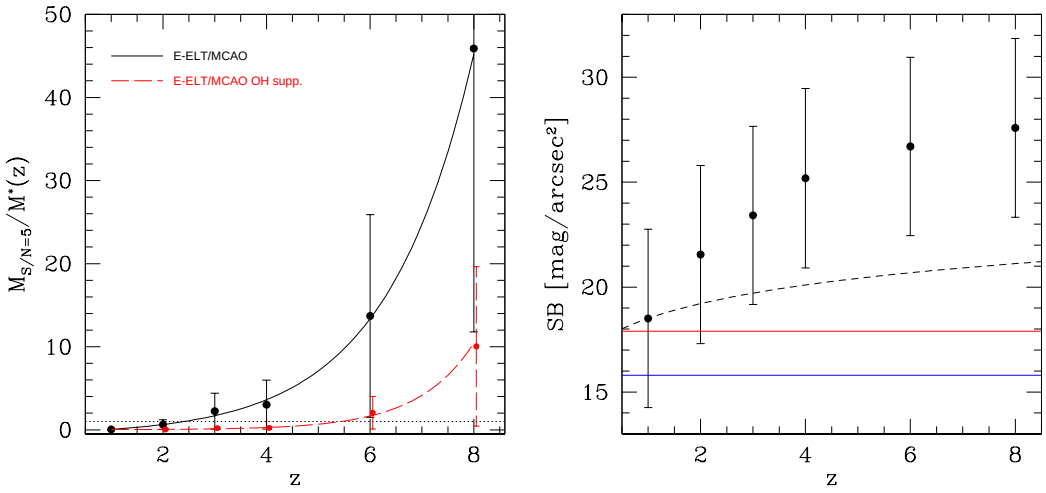

Figure 16.5: Left: Stellar mass limit that can be reached at a fixed S/N=5 as a function of redshift. Error-bars represent the range of mass limits obtained when considering all the morphological templates. The horizontal dashed line denotes M_* at all redshifts. The black line assumes a sky background as typically measured at Paranal, while the red line assumes OH suppression. Right: Comparison between the surface brightness of the simulated galaxies and sky background. The error-bars represent the range of values corresponding to masses between 0.1 and 10 M_* at a given redshift, while the dots represent the middle of this range. The blue line represents the H-band Paranal-like sky surface brightness, while the red line represents the H-band background with full OH suppression. The dashed line represents the surface brightness evolution due to cosmological dimming ($\propto (1+z)^4$). The decrease in surface brightness of the simulated galaxies is dominated by the evolution of their physical size and luminosity.

16.5.2.2 Exploration of the observational parameter space

Images of distant galaxies will be obtained in a background-limited regime (see right panel of Fig. 16.5). In such a regime, one expects the limiting S/N to scale as follows[7]:

$$S/N_{\rm lim} = 5 \times \left(\frac{T_{\rm intg}}{10\,{\rm h}}\right)^{0.5} \left(\frac{D}{42\,{\rm m}}\right) \left(\frac{\Delta_{\rm pix}}{4\,{\rm mas}}\right) \tag{16.3}$$

Using the simulations with different pixel sizes and telescope diameters we derive the following relations:

$$S/N_{\Delta_{\rm pix}=30\ {\rm mas}} = (7.2 \pm 0.2) \times S/N_{\Delta_{\rm pix}=4\ {\rm mas}}, \tag{16.4}$$

$$S/N_{D=42\ {\rm m}} = (1.4 \pm 0.2) \times S/N_{D=30\ {\rm m}}, \tag{16.5}$$

which are in excellent agreement with the expected scalings.

16.5.2.3 Comparison with JWST

As shown in Fig. 16.5, above z∼2.5, the E-ELT will be able to provide DL images of relatively massive galaxies only. To increase the S/N per pixel, it will be necessary to use a larger pixel scale. In this area, the main competitor of the E-ELT will be NIRCAM on-board JWST. As described in Section 16.4.3, JWST will be a 6.5 m telescope, which will provide NIR DL images at a 32 mas scale. This is why we use a 30 mas pixel scale to compare the performances of the E-ELT and JWST.

[7] Recall that the integrated number of photons reaching the detector is a degenerate function of several parameters, i.e., the total integration time $T_{\rm intg}$, the telescope diameter D, and the global throughput $t_{\rm transm}$, of which only variations in D were explicitly simulated. One can easily translate the results obtained for D to these other parameters.

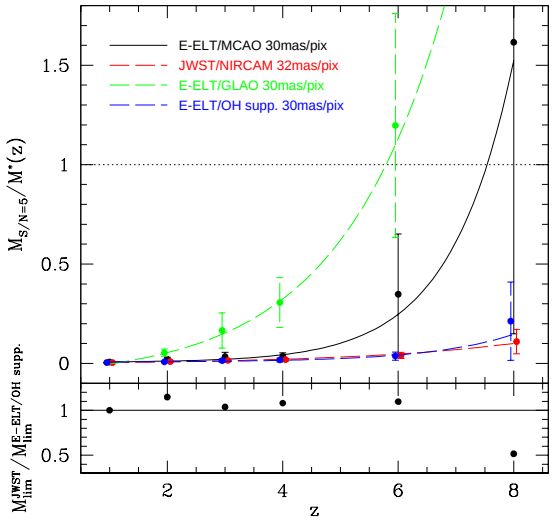

Figure 16.6: Top: Stellar mass limit that can be reached at a fixed S/N=5 as a function of redshift with 30 mas/pix. Error-bars represent the range of mass limits obtained when considering all the morphological templates. The horizontal dashed line denotes M_* at all redshifts. The black line assumes a sky background as typically measured at Paranal, the blue line assumes OH suppression and the green line GLAO corrections instead of MCAO. The red line shows the results for JWST/NIRCAM (zodiacal light limited observations). Bottom: Ratio of the mass limit obtained using JWST/NIRCAM to the mass limit obtained using E-ELT/MCAO with OH suppression, as a function of redshift.

The mass that can be reached at a fixed S/N=5 as a function of redshift is shown in Fig. 16.6. At a pixel scale of $\sim$30 mas, the E-ELT will allow us to probe the GSMF down to (at least) M_* galaxies up to z$\sim$7.5. Full OH suppression would remove this limitation and provide performances (in terms of S/N) comparable or even slightly superior to JWST. Only the highest redshifts will benefit from slightly better performances of JWST compared to the E-ELT with full OH suppression. In conclusion, the niche in imaging for JWST will clearly lie in observing very distant and faint galaxies (z$\geq$7).

Some illustrative results of simulations at z=4 and z=8 for LBG0328 are shown in Figs. 16.7 and 16.8, respectively. At z$\geq$7, the much lower background seen by JWST compared to the E-ELT provides a real advantage in S/N (see how the 0.1 M_* case at z=8 is detected by JWST whereas it is barely seen in E-ELT images). However, it can clearly be seen that the MCAO PSF provides better spatial resolution than the JWST PSF. A central cut of the MCAO and NIRCAM PSF is shown in Fig. 16.9. This figure confirms that, as expected, the MCAO PSF is much sharper than the NIRCAM one.

16.5.3 Compliance with figures of merit

In agreement with the figures of merit described in Section 16.3, we summarise the niche of the E-ELT in terms of imaging as follows:

- A minimal mean S/N per pixel of 5 is required to recover the structural parameters of distant galaxies at the E-ELT DL.
- Within the limit of S/N=5, the E-ELT will provide meaningful DL images of (at least) M_* galaxies up to z$\sim$2.5. Using full OH suppression shifts this limit to z$\sim$5.
- At a pixel scale of $\sim$30 mas, the E-ELT will allow measurements of the galaxy stellar mass function up to z$\sim$7.5. Compared to JWST/NIRCAM, the E-ELT will benefit from MCAO to provide images at a better spatial resolution. However, JWST/NIRCAM will provide better S/N above z$\sim$4. This cross-over point shifts to z$\sim$7 if full OH suppression is used.

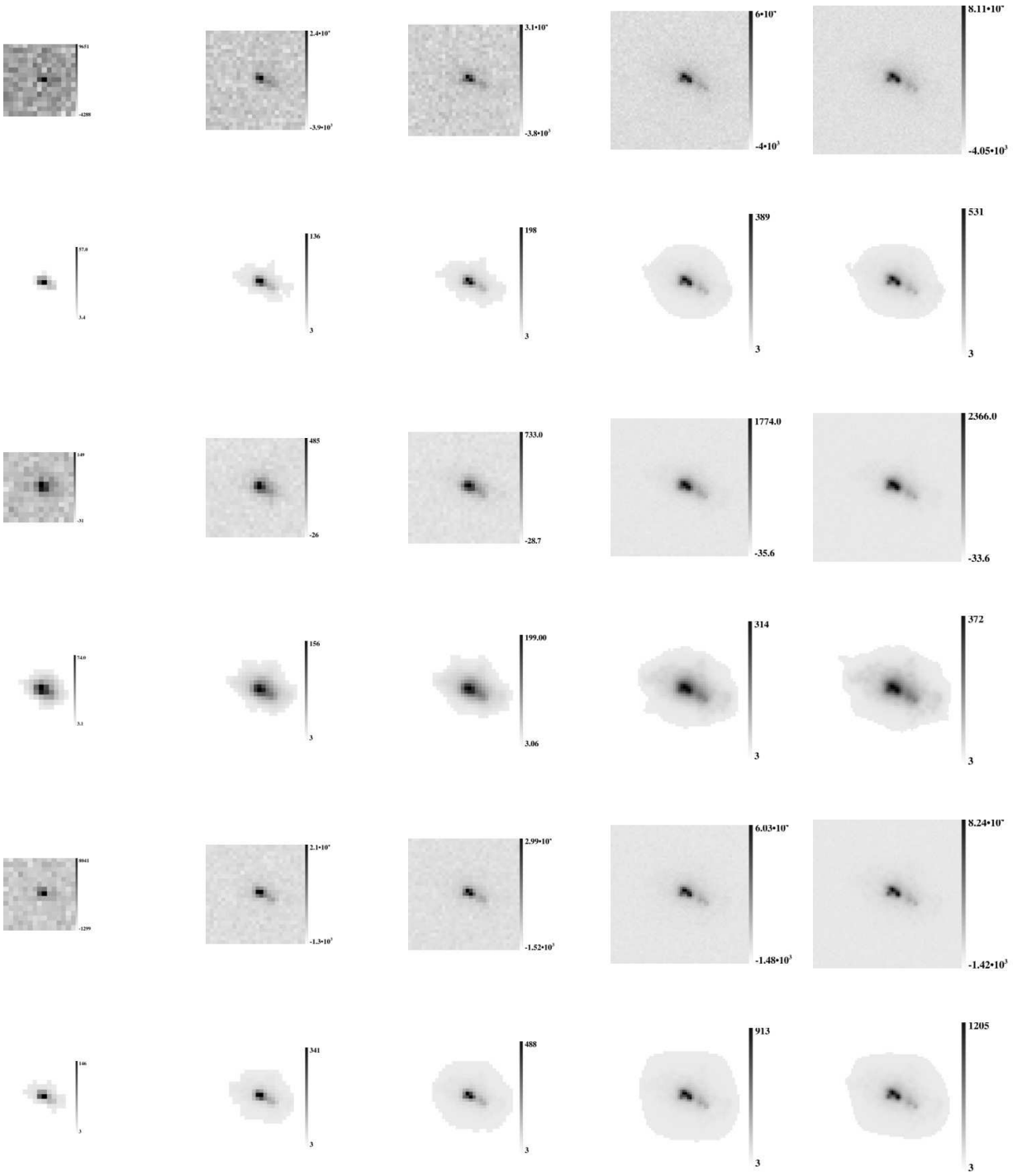

Figure 16.7: Results of simulations at z=4 using the LBG0328 template with the E-ELT and assuming a Paranal-like sky background (top two rows), using JWST (middle two rows) and using the E-ELT plus full OH suppression (bottom two rows). From left to right: simulations for stellar masses of 0.1, 0.5, 1, 5 and 10 M_*. First, third and fifth rows: images simulating what would be observed with the E-ELT using MCAO, after sky subtraction. Second, fourth and sixth rows: expected S/N per pixel. A threshold at S/N=3 has been applied. The mean S/N per pixel over the intrinsic galaxy diameter is, from left to right, 25, 52, 58, 114 and 149 (second row), 37, 73, 93, 148 and 175 (fourth row), and 56, 127, 157, 276 and 357 (sixth row).

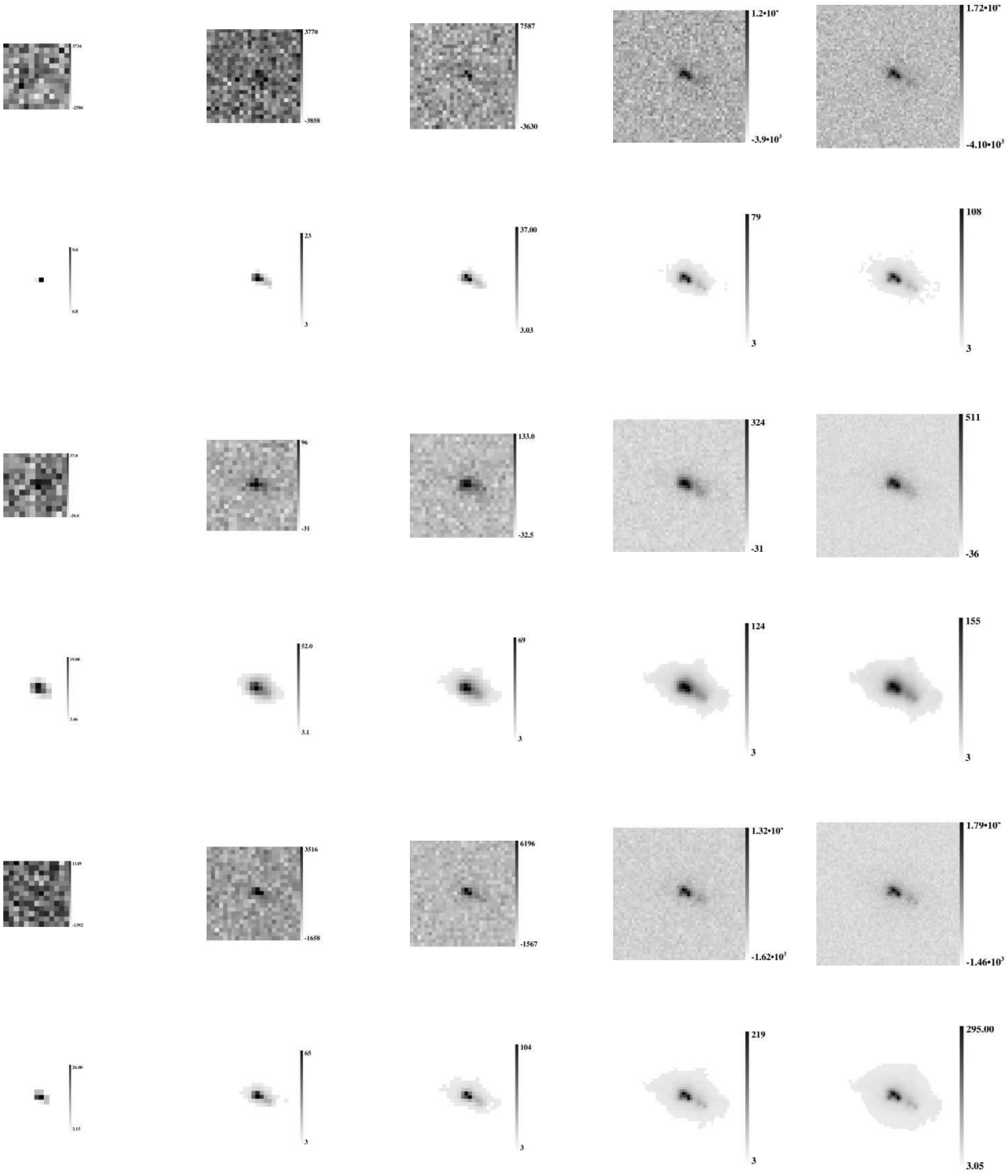

Figure 16.8: Results of simulations at z=8 using the LBG0328 template with the E-ELT and assuming a Paranal-like sky background (top two rows), using JWST (middle two rows) and using the E-ELT plus full OH suppression (bottom two rows). From left to right: simulations for stellar masses of 0.1, 0.5, 1, 5 and 10 M_*. First, third and fifth rows: images simulating what would be observed with the E-ELT using MCAO, after sky subtraction. Second, fourth and sixth rows: expected S/N per pixel. A threshold at S/N=3 has been applied. The mean S/N per pixel over the intrinsic galaxy diameter is, from left to right, 2, 8, 11, 22 and 30 (second row), 10, 21, 27, 52 and 66 (fourth row), and 0 , 29, 33, 65 and 80 (sixth row).

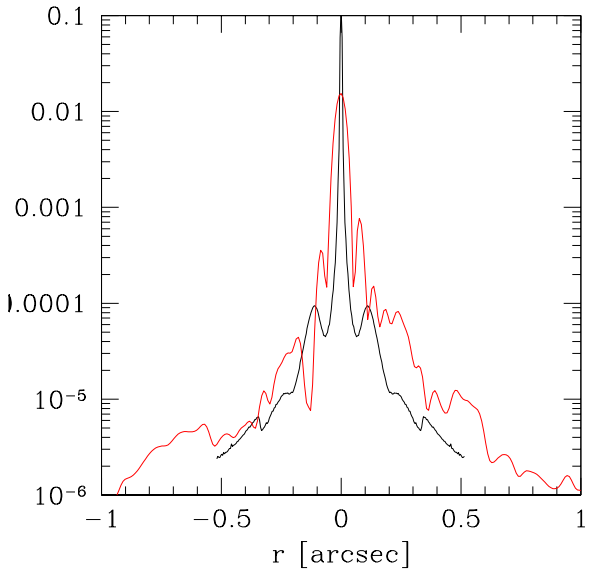

Figure 16.9: Comparison between the H-band MCAO PSF (black) and the F150W JWST/NIRCAM PSF (red) used for the simulations (log stretch). For both PSFs a central cut of the PSF is shown.

16.5.4 Sensitivity to input parameters

16.5.4.1 Impact of telescope

Diameter: The scaling relation of Section 16.5.2.2 shows that with a smaller, 30 m telescope one would need two times longer exposures to reach the same SNR. In terms of spatial resolution, MCAO provides us with high-Strehl PSFs, therefore one expects the spatial resolution of images to depend directly on the telescope diameter and to scale as $\propto 1/D$. With a smaller, 30 m telescope the spatial resolution element would be 28% larger, with 11.2 mas instead of 8 mas in the H-band. This translates to $\sim$55 (95) pc instead of 39 (68) pc at z=8 (z=2). This is far below the size of clumps observed in distant galaxies ($\sim$kpc), and well above the size of the most numerous H II regions (Elmegreen et al. 2006). There is therefore no clear break point in telescope diameter in terms of S/N or spatial resolution.

Thermal background: The telescope is by far the dominant source of background in the K-band. Above z=4 (i.e., where the Balmer break enters the K-band), it will therefore become difficult to obtain both H and K-band imaging in a limited amount of exposure time, and then separate emission between old and young stars (see Section 16.1). Note that in the simulations, we have already made the most optimistic assumption of 5% for the emissivity of the 5-mirror E-ELT.

16.5.4.2 Impact of site

Sky background: The sky is the dominant source of noise in almost all simulated observations (except the brightest galaxies at z=1). In this regime, one has $S/N \propto 1/\sqrt{\mathrm{background}}$: a site having a two times lower background than the one used in the simulations will increase the achieved SNR from, e.g., 5 to 7. Note that in the simulations we have used a sky model for a Paranal-like site, which is $\sim$2 times brighter in the H-band than that for Mauna Kea.

Seeing: We did not explore variations in terms of seeing, since we required a minimal Strehl ratio on the PSF (see Section 16.4.3). The strongest impact of worse seeing conditions will be to degrade the spatial resolution of the images. Such variations were investigated indirectly, through the comparison between MCAO and GLAO simulations (see below).

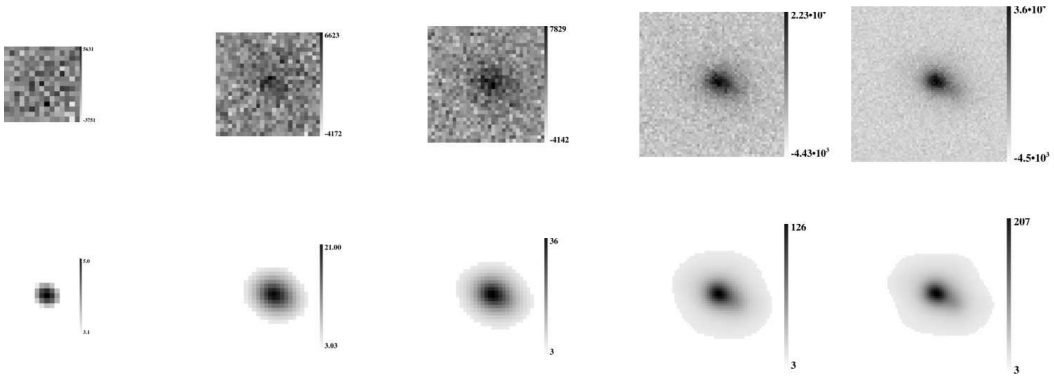

Figure 16.10: Simulations of LBG0328 at z=4 using GLAO. The top row shows the simulated images after sky subtraction, while the bottom row shows the expected S/N per pixel. A threshold at S/N=3 has been applied. The mean S/N per pixel over the intrinsic galaxy diameter is, from left to right, 1, 11, 17, 53, 82.

16.5.4.3 Impact of instrument

GLAO vs. MCAO: Results of simulations with GLAO are shows in Fig. 16.10. As expected, GLAO (or worse seeing conditions) results in lower S/N and spatial resolution (see Fig. 16.7 for comparison).

Pixel scale: The choice of the pixel scale directly drives the spatial resolution of images. At the highest redshifts (z$\geq$2.5), the E-ELT will not achieve high enough S/N per pixel to allow DL images. A natural larger pixel scale is $\sim$30 mas, i.e., close to what JWST/NIRCAM will provide. With this new pixel scale, the GSMF will be probed up to z$\sim$7.5. However, spatial resolution will be degraded by a factor of 7.5, with $\sim$290 pc per element of spatial resolution. This will not impact the ability of the E-ELT to resolve clumps in distant galaxies, as illustrated in Figs. 16.7 and 16.8.

OH suppression: OH suppression consists of reducing the sky background (see above) by adding filters along the optical path in the instrument (see, e.g., Ellis & Bland-Hawthorn 2008). We explored the extreme case of full OH suppression, which clearly demonstrates the interest of developing such a technique for the E-ELT. In the meantime, JWST will remain an important competitor for the E-ELT, especially for the most distant and faint sources.

16.5.5 Limitations

The main limitations of this work are the following:

Variations with wavelength: Most of the inputs are currently considered as invariant with wavelength (e.g., sky background, throughput, emissivity), because the design of the telescope and the instrument are not fully known at the moment. Simulations shall be updated once these characteristics are known.

Accuracy of the scientific inputs: Most of the quantities used here are not known to a factor better than two. However, simulations are internally consistent in the sense that neglected corrections on physical quantities are not larger than this factor. For instance, the empirical relations between stellar mass and K-band magnitude are uncertain within a factor two (likely more at z$\geq$4), but the flux correction between the rest-frame wavelength corresponding to the K-band and that of the simulation is less than 50% (see Section 16.4.3). For a more detailed discussion of this issue see Puech et al. (2010).

Representative of the simulations: For simplicity, we have adopted several assumptions that might impact the results of the simulations in terms of S/N (see Section 16.4.3). First, we have neglected the central obscuration of the E-ELT, which leads to an overestimate of the S/N by $\sim$8%, since $S/N \propto D$.

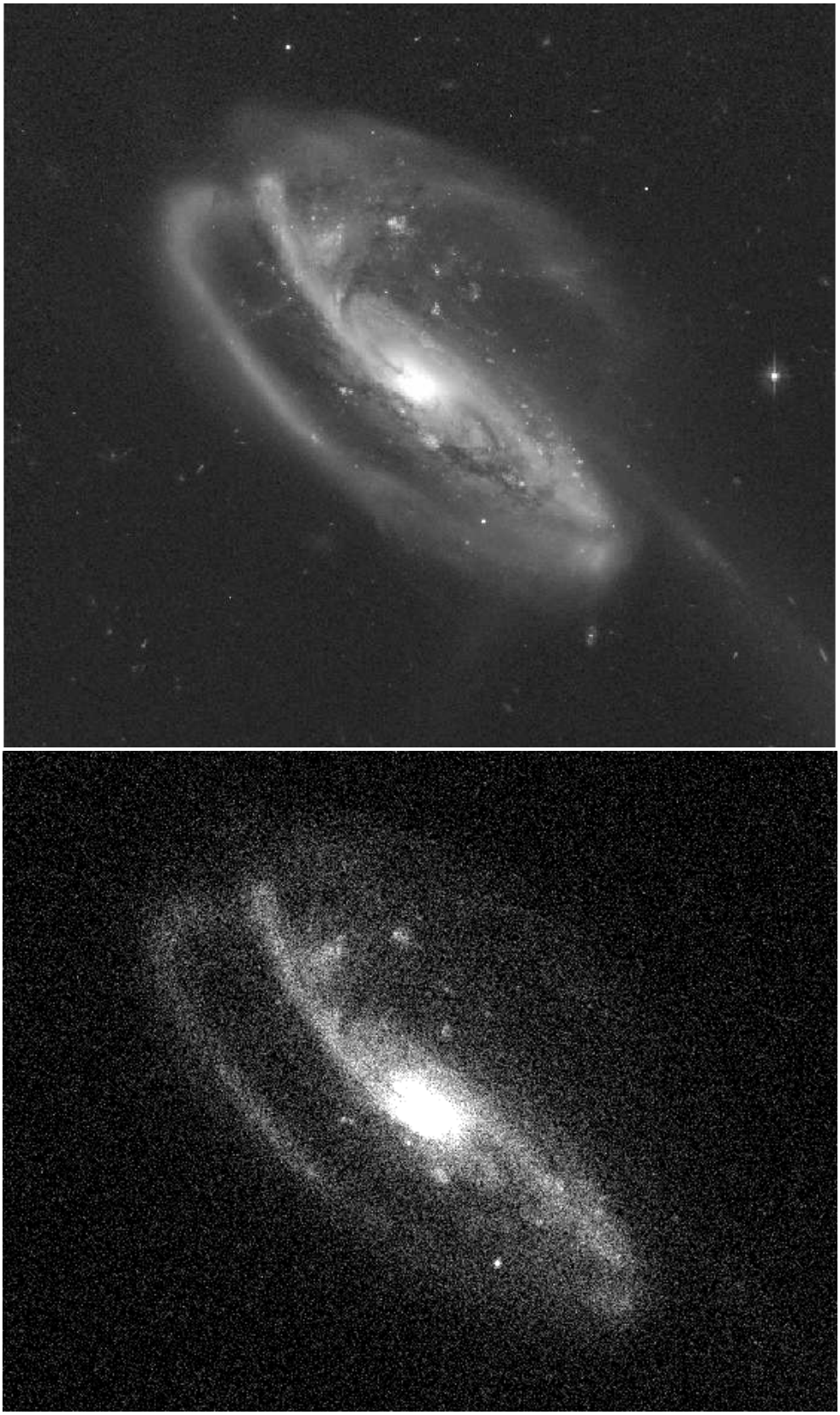

Figure 16.11: Result of simulation using the tadpole template. Bottom: This simulation corresponds to an M_* galaxy at z=1 imaged in the H-band at the E-ELT DL (i.e., 4 mas/pix). Top: The original template as shown in Fig. 16.1 is reproduced here for comparison.

As $S/N \propto \sqrt{T_{\rm intg}}$, this also impacts integration times by 17%, meaning that the real integration time of reference should be 11.7 hr instead of 10 hr.

16.6 Concluding remarks

We have conducted simulations of the "High-resolution imaging of high-redshift galaxies" science case exploring a wide range of observational and physical parameters. We have derived empirical scaling relations between the signal-to-noise ratio and the main telescope and instrument parameters, as well as a relation between the limit in stellar mass that can be reached for a given S/N as a function of redshift. We specifically investigated the impact of AO performance and OH suppression on the science goals. We did not identify any break points with respect to any of the parameters (e.g., the telescope diameter), with the exception of the total background, which strongly limits the observations of the farthest and faintest galaxies. Another critical parameter is the pixel scale, which directly drives the spatial resolution and achieved S/N of the observations. Unless full OH suppression is used, JWST/NIRCAM will remain an important competitor above z$\sim$4. However, the E-ELT will provide unrivalled diffraction-limited images of distant galaxies at z$\leq$2.5 (see Fig. 16.11). These look-back times correspond to the critical ages where the local Hubble sequence is built up, which will provide a well-defined niche for the E-ELT.

References

Barden M., Rix H.-W., Somerville R.S., et al., 2005, ApJ, 635, 959
Bouwens R.J., Illingworth G.D., Blakeslee J.P., et al., 2004, ApJ, 611, 1
Conselice C. J., Yang C., Bluck A. F. L., 2009, MNRAS, 394, 1956
Courteau S., Dutton A.A., van den Bosch F.C., et al., 2007, ApJ, 671, 203
Cox C., Hodge P., 2006, Proc. SPIE Vol. 6265
Cuby J.-G., Lidman C., Moutou C., 2000, The ESO Messenger, 101, 2
Ellis S.C., Bland-Hawthorn J., 2008, MNRAS, 386, 47
Elmegreen D.M., Elmegreen B.G., Rubin D.S., Schaffer M.A., 2005, ApJ, 631, 85
Elmegreen B.G., Elmegreen D.M., Chandar R., Whitmore B., Regan M., 2006, ApJ, 644, 879
Ferguson H.C., Dickinson M., Giavalisco M., et al., 2004, ApJ, 600, 107
Finger G., et al., 2006, Nuclear Instruments & Methods in Physics Research A, 565, 241
Garnett J.D., Farris M., Wong S., et al., 2004, Proc. SPIE Vol. 5499, 35
Grazian A., Fontana A., de Santis C., et al., 2006, A&A, 449, 951
Lisker T., 2008, ApJS, 179, 319
Lotz J. M., et al., 2008, ApJ, 672, 177
Neichel B., Fusco T., Conan J.-M., et al., 2008, Proc. SPIE Vol. 7015, 701573
Overzier R.A., Heckman T.M., Kauffmann G., et al., 2008, ApJ, 677, 37
Papovich C., Dickinson M., Ferguson H.C., 2001, ApJ, 559, 620
Puech M., Rosati P., Toft S., Cimatti A., Neichel B., Fusco T., 2010, MNRAS, 402, 903
Windhorst R.A., Conselice C., Petro L., 2002, draft of ApJ paper, not published

17 C4: First light – the highest redshift galaxies

Authors: B.P. Venemans, M. Lehnert, M. Franx

17.1 The science case

When and where did the first galaxies form? From very general considerations, the critical mass scale at which the "first galaxies" likely formed is in $\sim$5$\times$10^7 M$_\odot$ halos assembled at $z\sim$10. At this scale, the virial temperature of the halo exceeds 10^4 K and atomic cooling becomes the dominant cooling mechanism. Halos above this mass can retain photoionised gas and therefore likely maintained a self-regulated multi-phase ISM sustained by the energy due to star-formation (e.g. Mac Low & Ferrara 1999; Madau et al. 2001; Ricotti et al. 2002). These are the "first galaxies".

Many other questions surround the formation of the first galaxies. What was the shape of the initial mass function (IMF), if such a designation even makes sense for the first galaxies? The IGM is likely to have been heavily enriched (Becker et al. 2009), the radiation fields may have been intense and the turbulence high during this epoch. The balance between metallicity, turbulence generated by collapsing gas and mechanical energy injection from the stars, and the strength of the radiation is not at all academic as it determines whether the star-formation is regulated by cooling through molecular or atomic lines. These two cooling mechanisms could lead to substantially different IMFs (Bromm et al. 2003; Greif et al. 2007; Wise & Abel 2008). Moreover, the accretion of cold gas from filaments in combination with a softened equation of state drives strong shocks and turbulence (e.g. Kereš et al. 2005) which might lead to vigorous fragmentation and the formation of the first star clusters. Such a situation would be an important early step towards conditions that are similar to star-formation in the local Universe and hence a "normal" IMF.

However, to know whether or not any of our modelling represents reality, and to solve this "tension" in the models of the first galaxies, observers must find the first galaxies and then confirm that they are indeed in the early Universe. To overcome the problem of discovering galaxies at high redshift, regardless of their properties, we should use selection methods which rely on extrinsic rather than intrinsic galaxy properties and which, if possible, can be applied over the entire history of the Universe. The dropout technique, which selects galaxies according to the strength of their Lyman break (so-called Lyman break galaxies, LBGs) is such a selection method. LBGs have been the benchmark against which other samples of high-redshift galaxies have been compared since their discovery by Steidel et al. (1996), and they allow us to study galaxies selected in a consistent manner over the range $z = 0$–7 and now perhaps even up to $z = 8.5$ (Bouwens et al. 2010). So, while LBGs can in principle select every type of galaxy, we inherently add a bias which is that they must be actively star-forming with little reddening to "clean" the samples of "interlopers". At higher and higher redshifts it becomes difficult to remove interlopers as data of sufficient depth at longer wavelengths (i.e., beyond 2.5 μm) are difficult to obtain. However, despite this, it is the simplicity of selection of the LBG technique, its robustness, and wide redshift span probed that makes the LBG selection technique unrivalled for (comparative) galaxy evolution studies.

To understand the properties of the "first galaxies", we propose to use integrated JHK spectroscopy on the E-ELT (in GLAO mode) to identify the first UV emitting sources in the Universe responsible for reionisation at $z > 6$. These integrated spectra will be useful for obtaining redshifts, removing interlopers, i.e., galaxies not in the early Universe but at somewhat lower redshifts or stars, and for determining the correlation lengths of galaxies in the early Universe. In addition, more detailed studies of individual galaxies at the bright end of the high-redshift galaxy luminosity function can be undertaken using a high multiplex near-infrared spectrometer with deployable integral field units fed by a high-correction AO system such as LTAO or MOAO. Because of the ability of such a spectrograph to obtain high spatial resolution and to concentrate the light into individual spaxels, these data will be useful in characterising the nature of the UV continuum, investigating the properties of the rest-frame UV lines to determine the covering fraction, column densities, and kinematics of the UV absorption line gas over a wide range of ionisation states (e.g., C IV, S II, etc.), unveiling the nature of structure seen in many distant galaxies, and possibly finding Lyman-α scattering halos in some sources due to neutral H within the epoch of reionisation.

The galaxy candidates will be identified via the Lyman break technique with upcoming deep space-based and ground-based near-IR surveys (with WFC3 on HST, Spitzer, VISTA and the JWST). The source density at $z \sim 8$ is about 1 arcmin^{-2} at $H_{\rm AB} \sim 28.5$, for which an E-ELT spectrograph equipped with a GLAO or LTAO/MOAO system will be necessary to obtain redshifts and to understand their physical properties. These observations will probe an era of a few hundred million years right at the end of the "dark ages", will identify the sources of reionisation and yield their luminosity function. Such a program will give the first insights into the crucial early stages of galaxy formation and what caused reionisation.

17.2 Goals of the DRM simulations

- What integration time is needed to measure redshifts of high-redshift galaxy candidates identified by the ULTRA-VISTA survey? These galaxies will have magnitudes up to $H_{\rm AB} = 26$ mag.
- Determine the integration time needed to measure redshifts of very faint, high-redshift galaxies identified through deep HST and JWST imaging. These galaxies have $H_{\rm AB} = 28.5$ mag.
- Measure physical parameters of a sub-sample of these high-redshift galaxies with $H_{\rm AB} = 27$ mag that have been preselected to have redshifts which are favourable for these measurements (i.e. emission/absorption lines not in regions where contamination from terrestrial line emission and absorption is significant).

We did not, for the purpose of this DRM case, consider very low resolution spectroscopy ($R \sim 100$ to a few $\times 1000$). We made this decision to both simplify the analysis and because of competition with JWST. NIRSpec on JWST will have low resolution modes of $R \sim 100$, $R \sim 1000$, and $R \sim 3000$. Because of the lower background and the lack of night-sky features compared to observing from the ground, and a broad wavelength response in the low resolution mode (about 0.8 to 5 μm), JWST will have a great sensitivity advantage over the E-ELT at resolutions less than 1000. We feel this low resolution science for distant galaxies is better suited for programs on the JWST than with E-ELT.

17.3 Metrics / figures of merit

For the first two goals outlined above, we want to estimate the integration time needed to accurately measure the redshift of $z \sim 8$ galaxy candidates. The goal is reached if, with a certain integration time, 90% of the galaxies have a redshift measured to an accuracy better than 0.01, which corresponds to 300 km s^{-1} for a galaxy at $z \sim 9$. Obviously, how long it takes to reach these goals depends on a number of factors including the quality of the AO correction, the efficiency of the instrument, etc.

The third goal above is reached if we are able to measure the absorption column density within a factor of 3 in at least 50% of the galaxies. This is the most difficult challenge for any spectrograph on the E-ELT as it requires high enough signal-to-noise (S/N) in the UV continuum to accurately measure equivalent widths of the lines as well as determine if the lines are resolved or not.

17.4 DRM simulations

17.4.1 Methodology

To investigate the possibility of measuring accurate redshifts and physical parameters of high-redshift galaxies, we created mock galaxies at $z > 7$ with different types of morphology and spectral energy distributions (SEDs). Since virtually nothing is currently known about very high-redshift galaxies other than that they exist, we based these "scientific" parameters on observations of galaxies at lower redshifts (17.4.3.1). The results of the simulations also depend on many "technical" parameters describing e.g. the telescope, site, type of instrument, AO correction, etc.

The goals of this DRM case could in principle be met with various types of spectrographs. To keep the number of simulations to a manageable level, we have limited ourselves to two basic spectrograph types

(see Section 17.4.3.2). One is an integral field spectrograph and the other is a fibre-fed spectrograph. The difference between these is the kind of AO technique that is used to feed the spectrograph and the size of the entrance aperture or spaxel size. For the IFU spectrometer, we will assume that the spaxel scale is 37.5 milliarcseconds (mas) while adding one set of simulations with a finer scale, 10 mas. This will allow us to investigate the impact of read-out noise and dark current on the sensitivity of the observations. For the fibre-fed spectrograph, we will assume a 900 mas entrance aperture. All simulations will be at $R = 4000$ (with a set at higher resolution of 10 000 to test the impact of increasing resolution on data quality and scientific return). These numbers correspond approximately to the range proposed in various instrument concept studies for the E-ELT and other ELTs.

To compare the influence of the various scientific and technical parameters we have defined a reference case to which we compare the different simulations. The reference case has the following scientific and technical parameters (see Sections 17.4.3.1 and 17.4.3.2 for details):

- The galaxies are located randomly in redshift between $7.4 < z < 9.2$ and have $H_{\rm AB} = 26.0 \pm 0.2$ (in AB magnitudes, which are used throughout this report). For the morphology three different types are used, ranging from centrally concentrated to a merging, extended, clumpy system. This range will allow us to test if we can recover the morphologies of distant galaxies with the performance expected for various AO systems and spectrograph types on the E-ELT.
- The telescope has a 42 m primary mirror, coated with bare aluminium and is located in a site like Paranal with a typical seeing of 0.6″. For the adaptive optics the corrections provided by a Multi-Conjugate Adaptive Optics (MCAO) system are used. The galaxies are simulated to have an airmass of 1.5.
- The instrument contains an integral field unit (IFU) with a spaxel scale of 37.5 mas spaxel^{-1} and a field of view (FoV) of 1.65″. The resolution of the spectrograph is 4000, critically sampled at the centre of the J-band. The wavelength range of the spectra is $1.04640 < \lambda < 1.35345$ μm over 2048 pixels of the detector.

These are simply the parameters used for the baseline simulations. The full suite of simulations will include different methods of AO corrections, a fibre-fed spectrograph with a single entrance aperture of constant size, and a few simulations with different spaxel scales for the IFU, and spectral resolutions for both types of spectrographs.

17.4.2 Pipeline

In this section we describe the pipeline used to create mock observations, reduce the resulting 3D data cubes, and analyse the extracted spectra. The reduction of the mock observations depended on the type of instrument modelled in the simulation (see Section 17.4.3.2). For both cases, the IFU spectrograph and the fibre-fed spectrograph, the reduction of the mock observations is described in detail below.

In short, we first generated 3D data cubes containing the spatial and spectral information of mock galaxies. Noise of both the sky and the object itself was randomly added assuming an underlying Poissonian distribution and the resulting data cubes were subsequently analysed without prior knowledge of the shape, position or extent of the simulated galaxy. A 1D spectrum was extracted from the data cube, which was then analysed to recover the redshift and some physical parameters of the galaxy.

17.4.2.1 Simulating observations

To simulate observations of high-redshift galaxies, we first created artificial $z > 7$ galaxies of one of the morphological types, with a random redshift (see Section 17.4.3.1) and position angle on the sky. These mock galaxies were subsequently convolved with a PSF appropriate for the AO correction assumed and re-mapped onto a coarser (instrument dependent) spaxel scale. As baseline a spaxel scale of 37.5 mas spaxel^{-1} was used.

The SED of the galaxy (Fig. 17.1) was generated at a given resolution R, redshift and magnitude. The spectra are roughly Nyquist sampled, with a dispersion of $\lambda_c/2R$, where λ_c is the central wavelength of the spectrum. In all our simulations the central wavelength was set to 1.22 μm. The spectrum was converted from flux density to photon density and multiplied by the atmospheric extinction, the telescope throughput and the total instrument throughput (including detector). The baseline spectra had 2048 spectral pixels and a resolution of 4000, giving a wavelength range of $1.046 < \lambda < 1.353$ μm, and a spectral sampling of 1.5 Å.

The spatial light distribution and the SED were combined into a 3D data cube. Photons from the background (sky), modified by the telescope and instrument throughput, were added to each pixel. The number of photons in each pixel in the data cube was converted to counts by multiplying by the total collecting area of the telescope, a given exposure time, and the detector quantum efficiency. Noise was randomly added to each pixel. The sources for the noise were assumed to have a Poissonian distribution and included the object, the sky background and the dark current, and detector read-out noise.

These mock 3D data cubes should mimic real observations, i.e., the sky background is not subtracted in advance, but is part of the data reduction. We did not take into account any flat field errors, we assumed perfect flat fielding. Individual exposure times were limited to 3600 s, longer exposures were broken up in several shorter exposures, which increased the total read-out noise.

17.4.2.2 Reduction of IFU data

The reduction of the mock 3D data cube in case of an IFU instrument was done in several steps. First of all, pixels strongly affected by the sky emission lines and/or atmospheric extinction were marked and not used in the remainder of the reduction. The marked pixels had a background level of at least twice the median background, and/or an atmospheric transmission of less than 50%. To reduce the influence of the wings of the sky emission lines, pixels neighbouring marked pixels were also masked. At a resolution of 4000, this procedure roughly removes half of the pixels in the J-band.

The data were reduced assuming no prior knowledge of the redshift, morphology, SED of the galaxy or its location within the IFU FoV. To identify the rough location of the object on the IFU, a 2D image was created by summing the pixels in the spectral direction, and smoothing the created image by averaging the pixels in 3×3 blocks. The highest peaks in the image were located and the centre of light was computed. All pixels at least 0.5″ away from the peak were used to determine the sky background in each spectral pixel and subsequently subtracted from the data cube.

High-redshift objects are expected to have no signal below the redshifted Lyman-α line due to the Lyman-α forest and perhaps the Gunn-Peterson effect (depending on the neutral fraction of H in the early Universe). For galaxies at roughly $z > 7.5$ the Lyman-α line shifts into the spectral range of our simulated J-band. To more accurately determine the location of the object on the IFU, an image was created by summing only the highest wavelength spectral pixels. All spatial pixels with a S/N above 0 in this image were selected and combined to create a 1D spectrum. Since we are only interested in determining the redshift, the 3D information, if any, is lost in this analysis. However, the morphology of the galaxy still plays a role in determining the final S/N of the spectrum in our reduction method which realistically mimics what would happen at the telescope.

17.4.2.3 Reduction of fibre data

In the case when the simulated instrument was a fibre-fed spectrograph, the steps to convert the 3D mock data cube to a 1D spectrum were simple. We assumed a perfect sky subtraction (i.e., no extra sky noise was added with the sky subtraction) and added up all the flux within the fibre. The resulting 1D spectrum contains object and sky noise and noise from the dark current and the reading of the detector. This is perhaps rather favourable for the fibre-fed spectrograph as in reality, another fibre or set of fibres would have to be used to construct a sky spectrum. Differences in fibre transmissivities, flat-fielding errors, etc., would reduce the S/N of the extracted spectrum.

17.4.2.4 Determination of the redshift and physical parameters

To determine a redshift from the 1D spectrum, the data were binned into initially only four bins of equal size. As a first guess of the redshift the location of a break in the binned spectrum was determined. This procedure was repeated with bin sizes that were smaller than the previous one by a factor of two, to more accurately measure the location of the break in the spectrum. This initial guess of the redshift and the width of the bin was then used to fit absorption lines to the spectrum.

The column density of the absorption line was measured by fitting Voigt profiles to each of the absorption lines separately. The Voigt profiles had three free parameters: the central wavelength of the absorption line, the velocity dispersion of the absorbing gas and the column density. The Voigt profile fitting was done both with the velocity dispersion as a free parameter for each absorption line and with the velocity dispersion kept constant.

17.4.2.5 Imperfections of the pipeline reduction

The reduction of the IFU data cubes has several imperfections. It assumes perfect flat-fielding, while the sky subtraction and extraction of the 1D spectrum is done with no prior knowledge of the sky background and galaxy extent and morphology. Especially the combining of the individual spectra of the data cube into one 1D spectrum will significantly improve if one knows the morphology and location of the galaxy on the IFU. Extracting the spectrum using the full knowledge of the light distribution of the galaxy using the input image increased the S/N of the 1D spectrum. Assuming perfect sky subtraction and optimal extraction of the spectrum will reduce the exposure time needed to reach the goals of the simulations by 25% in some cases, and removes any dependance of the results on spaxel size (as long as noise in the simulated observations are dominated by the sky noise) and removes most of the dependence on galaxy morphology.

The automated fitting of the redshift to the 1D spectrum is not doing a perfect job. This is especially true at the low-redshift end of our simulated z-range, where the Lyman-α break is shifted blueward of the J-band. Also, the eye can more easily reject redshift solutions that would leave significant emission blueward of the Lyman-α break. We tested the redshift estimates by manually fitting absorption lines to the data, and found that this manually approach reduced the number of outliers (defined as $|\Delta z| > 0.01$, see Section 17.5.2) by half. We therefore decided that the goal of the simulation was reached when 80% of the galaxies have an accuracy of $|\Delta z| < 0.01$, assuming we would have a 90% success rate when measuring the redshifts manually.

17.4.3 Inputs

17.4.3.1 Scientific data

Observations of galaxies at $z > 7$ are virtually non-existent. Galaxy candidates potentially lying at these high redshifts have been identified in very deep HST observations. These observations provide us with number densities and (rough) luminosity functions for $z > 7$ galaxies. However, spectroscopic confirmation does not exist and would have to rely on a single emission line – Lyman-α. In contrast, many thousands of galaxies have been confirmed at redshifts around $z \sim 3$. We will assume that the SED in the rest-frame far-UV is similar for $z \sim 3$ and $z \sim 8$ galaxies, hence we will use the numerous observations of $z \sim 3$ galaxies to model the SED of $z > 7$ galaxies.

In contrast, for the morphologies, the high-resolution images of $z \sim 3$ galaxies made with the HST have a resolution of $\sim$50 mas, while the E-ELT should be capable of providing a spatial resolution which is at least a factor of 10 better. We will therefore not use images of $z \sim 3$ galaxies, but will model the morphology of $z > 7$ galaxies based on local star-forming galaxies which are thought to be good analogs for LBGs in the distant Universe.

Below, we will describe the modelling of the $z > 7$ galaxies in more detail.

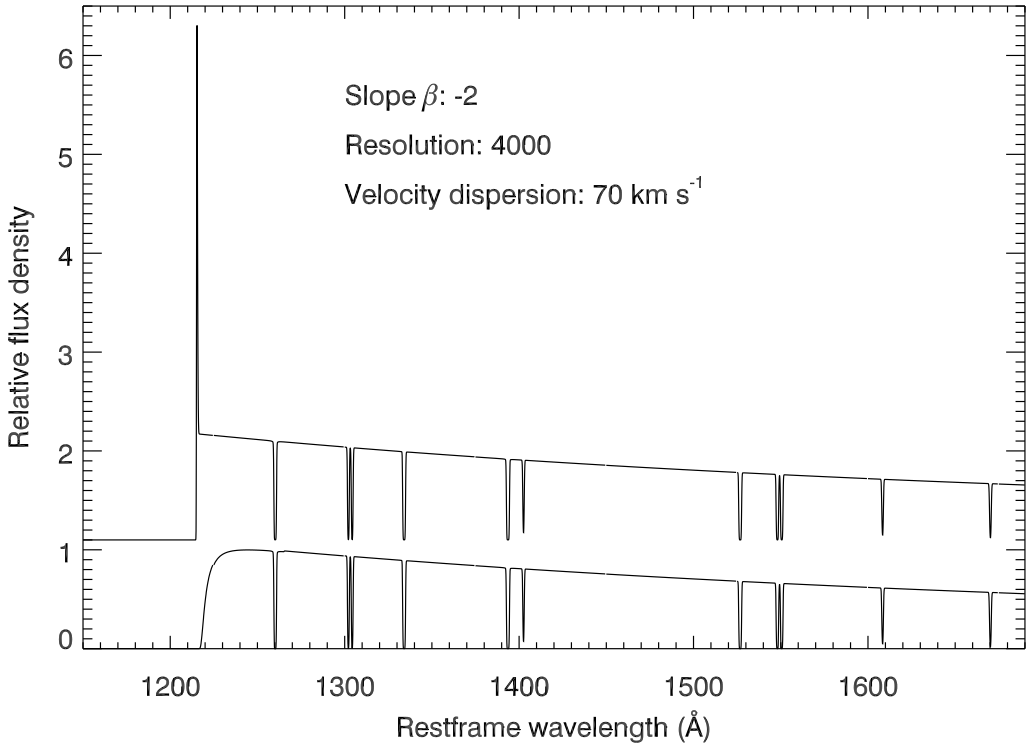

Figure 17.1: Example of the model SEDs used in the simulations. The spectrum at the top has Lyman-α in emission, while in the bottom spectrum the Lyman-α is absorbed. The spectrum at the top is offset for easier comparison. Both spectra have a resolution of 4000.

Spectral Energy Distribution: The spectra of high-redshift LBGs were created using the following assumptions. The continuum has a slope of $\beta = -2$ ($f_\lambda \propto \lambda^\beta$). We assumed that no flux was transmitted blueward of Lyman-α at 1215.67 Å. For the first two science cases, the absorption line strengths were taken to be similar to the composite $z \sim 3$ LBG spectrum as presented by Shapley et al. (2003). In Table 17.1 the wavelength, absorption equivalent width and corresponding column densities are listed.

The intrinsic velocity dispersion of the lines was set to 70 km s^{-1}, which is similar to the value measured in a highly lensed LBG at $z \sim 3$, CB58 (Pettini et al. 2002). Lyman-α was included either in absorption, with a column density of log $N_{\rm HI}$ = 20.65 (Pettini et al. 2002) or in emission, with a line FWHM of 200 km s^{-1} and a rest-frame equivalent width of 50 Å. All flux shortward of the peak was removed to mimic absorption by neutral Hydrogen. In Fig. 17.1 the two spectra are shown at a resolution of 4000.

Table 17.1: Wavelengths and equivalent widths (EWs) of the absorption lines as measured in $z \sim 3$ LBGs by Shapley et al. (2003). Our model $z \sim 8$ spectra with these line EWs are shown in Fig. 17.1. For the conversion between equivalent width and column density we assumed a velocity dispersion of 70 km s^{-1}.

λ [Å]	EW [Å]	Column density [log cm^{-2}]	λ [Å]	EW [Å]	Column density [log cm^{-2}]
1260.42	1.63	16.42	1526.71	1.72	16.65
1302.17	1.10	16.10	1548.20	1.515	15.89
1304.37	1.10	15.84	1550.78	1.515	16.18
1334.53	1.72	17.50	1608.45	0.91	15.21
1393.76	1.83	16.93	1670.79	1.04	13.83
1402.77	0.81	14.64			

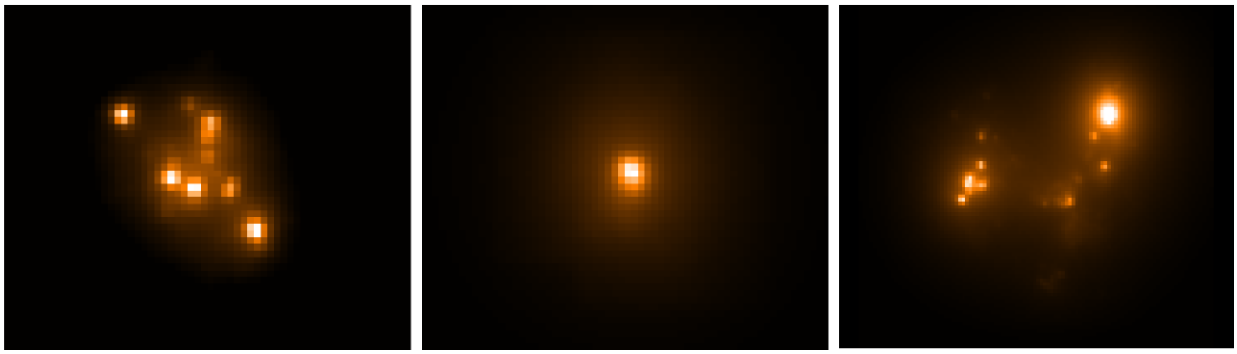

Figure 17.2: The three morphological types of high-redshift galaxies that are used in the simulations. From left to right: a clumpy galaxy, a Dominant Central Object, and Haro 11 shifted to high redshift, representing a merging system. The sizes of the images are roughly 1.0″, which corresponds to 4.8 kpc at $z = 8$.

Morphology: Not much is known about the morphology of galaxies at $z \gg 2$, as even in HST images such galaxies are barely resolved. The best information regarding the shape and sizes of high-redshift galaxies comes from highly lensed objects and low redshift starburst galaxies, the so-called Lyman break analogs (Heckman et al. 2005). Based on this information we decided to simulate high-redshift galaxies with three different types of morphology: (i) a clumpy galaxy with most of the light in compact starburst regions (based on observations by Swinbank et al. 2009); (ii) an object with a single, dominant luminous point source, containing 50% of the light, at the centre of the galaxy, referred to as a "Dominant Central Object" (DCO, Overzier et al. 2009); (iii) an object similar to Haro 11, a very nearby star-forming galaxy. This more extended morphology could represent a merging system. These three morphological types are shown in Fig. 17.2.

Redshift distribution: Very recently candidate $z \sim 8$ galaxies were found in the Hubble Ultra Deep Field using new HST/WFC3 data (Bouwens et al. 2010; Bunker et al. 2010; McLure et al. 2010). Using simulations, Bouwens et al. (2010) computed the expected redshift distribution of these $z \sim 8$ candidates (see Fig. 17.3). Using these figures we assume that the galaxies are distributed over the range $7.4 < z < 9.2$. For simplicity we further assume a flat redshift distribution. For the third science case, all model galaxies were located at $z = 8$. This redshift was chosen to make sure that the absorption lines in the LBG spectra were located in regions clear of strong background emission lines.

17.4.3.2 Technical data

The technical data used for the simulations were mainly taken from [RD1].

Site and background: We follow [RD1]. We consider the background emission and transmission models of both the Paranal-like and High & Dry sites. However, because the simulations are performed in the J-band (∼1.0–1.4 μm), we can ignore the thermal emission from the atmosphere and telescope as demonstrated in Fig. 17.4. We assume an airmass of 1.5.

Telescope: As described in [RD1]. We consider both the bare aluminium and the protected silver/aluminium coatings.

Instrument: We model two different instrument types. The first is an IFU with a default spaxel size of 37.5 mas and a FoV of 1.65″, the other is a fibre-fed spectrograph with an entrance aperture of 0.9″ diameter, sampled by three pixels in the cross-dispersion direction. Both are assumed to operate in the J-band at a spectral resolution of $R = 4000$. A spectral resolution element is sampled by two pixels. We ignore the thermal background of the instrument, which can be done safely as it is required that the instrument background is much lower than that of the telescope (see Fig. 17.4). The instrument throughput is set to 35% including the QE of the detector. This value is similar to that of SINFONI but is much higher than FLAMES/GIRAFFE on the VLT. However, for the purposes of this work, we have

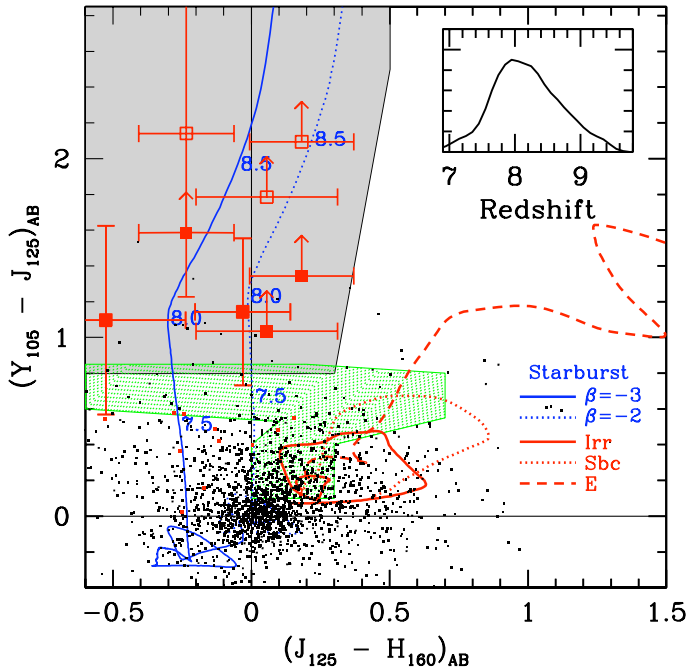

Figure 17.3: Fig. 1 from Bouwens et al. (2010), showing the colour selection of $z \sim 8$ galaxies using deep HST images. The inset shows the redshift distribution of the colour selected candidates.

assumed that both the IFU and the fibre-fed spectrographs have the same total efficiencies. The read-out noise of the detector is 3 e^- pixel^{-1} and the dark current 4 e^- h^{-1} pixel^{-1} (see [RD1]).

AO mode and PSFs: For the shape of the PSF we considered not only Laser Tomography Adaptive Optics (LTAO), Ground Layer Adaptive Optics (GLAO) and natural seeing (noAO), for which PSFs are available in [RD1], but also Multi-Conjugate Adaptive Optics (MCAO). Simulated PSF profiles for MCAO corrections were taken from the web site of the MAORY consortium. The MAORY PSF version used was 12 October 2009. The PSFs taken from the MAORY consortium are labelled "Maory" in Table 17.3. These calculated PSF profiles do not include errors due to laser guide star cone effects, natural guide star wavefront sensors or implementation errors, and have therefore more optimistic Strehl ratios and encircled energies than a more realistic PSF simulation that included these error sources (cf. Fig. 17.5). Also, for the MAORY PSFs a seeing of 0.6″ was assumed, while the simulated PSFs of [RD1] assume a seeing of 0.8″. More recent simulations from the MAORY consortium (version 01 November 2009) include MCAO PSFs in 0.8″ seeing. We will also use these PSFs to investigate the effect of the seeing. As can be seen in Fig. 17.5, the 0.6″ seeing MAORY PSF has high encircled energies, while the 0.8″ seeing MCAO PSF (with correction factors included) more closely follows the LTAO PSF. Note that the optimistic MCAO PSF has encircled energies very similar to Multi-Object Adaptive Optics (MOAO), as simulated by Puech et al. (2008). We will use both MAORY and LTAO PSFs to evaluate the influence of the PSF on the results.

17.4.4 Outputs

The outputs of the simulations consist of a fits image of the input galaxy convolved with a PSF, a 3D data cube (in the case where the modelled instrument was an IFU), a reduced 1D spectrum and a redshift determination. In Fig. 17.6 an example is shown of a spectrum extracted from a simulated observation of a $z \sim 8.36$ galaxy. The absorption lines used to measure the redshift are indicated. For the analysis of the simulations we will only use the redshift determination to see how accurately we can recover the input redshift. To assess how precisely we can measure the physical parameters of the galaxies we will use the extracted 1D spectra and fit the individual absorption features (see Fig. 17.7 for an example).

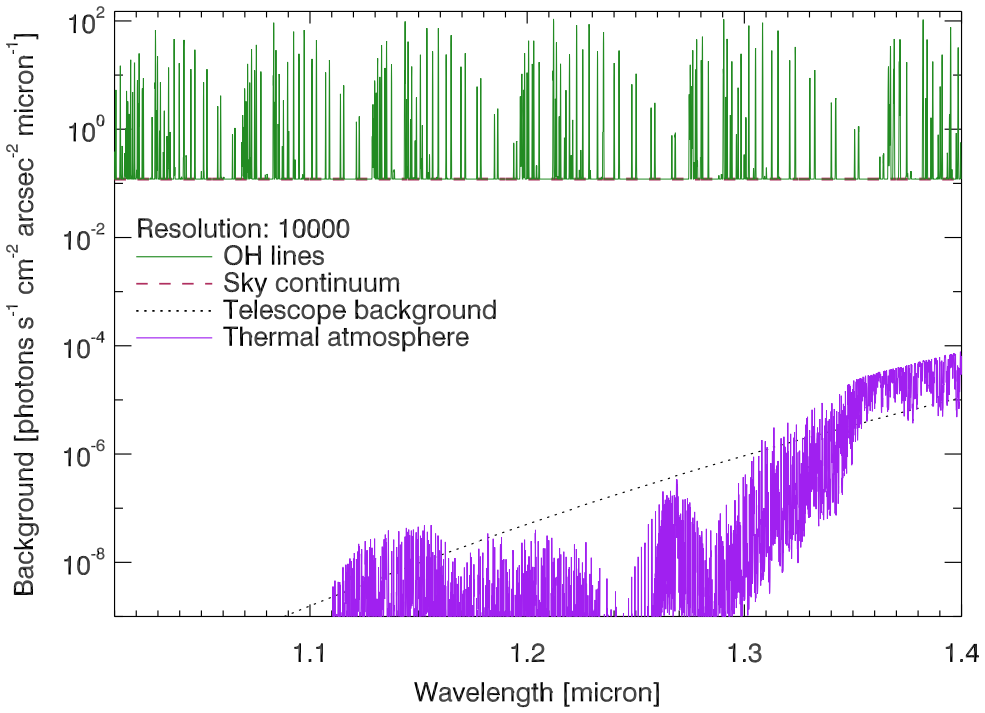

Figure 17.4: The total background in the J-band as seen by an instrument behind the E-ELT at a resolution of 10 000. In the J-band the thermal emission from both the telescope and the atmosphere are much lower than the OH emission lines and the sky (pseudo-)continuum and their contribution to the total background can be ignored. Note that the instrument background is required to be only a fraction of that of the telescope and can thus also be ignored.

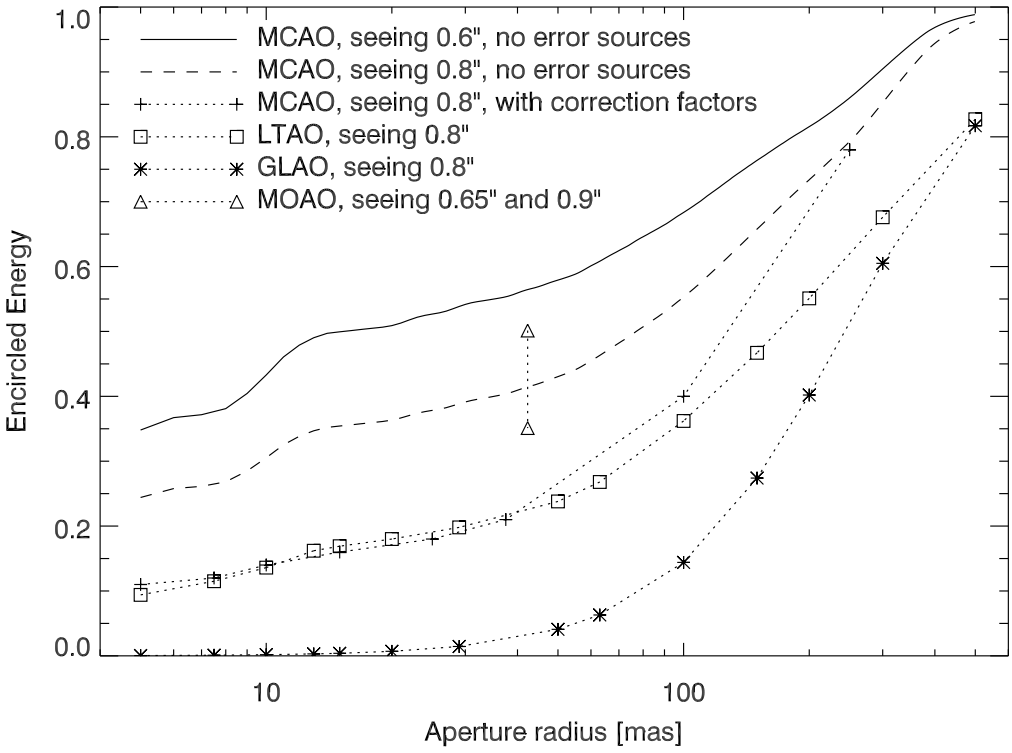

Figure 17.5: Encircled energy as a function of aperture radius for various types of AO correction. As reference we used MCAO with a seeing of 0.6″ and no error sources.

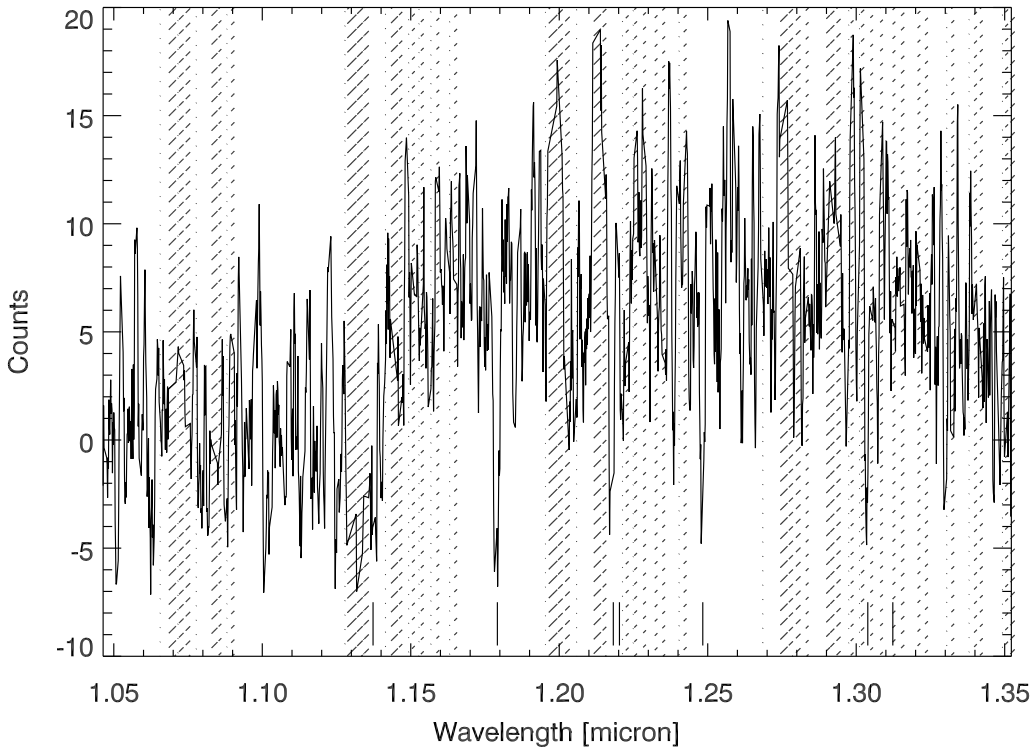

Figure 17.6: Example of an extracted 1D spectrum of a simulated galaxy from run s1_1800 (see Table 17.2). The absorption lines used to measure the redshift of the galaxy are indicated with tick marks below the spectrum. Regions with strong telluric emission and/or absorption are masked.

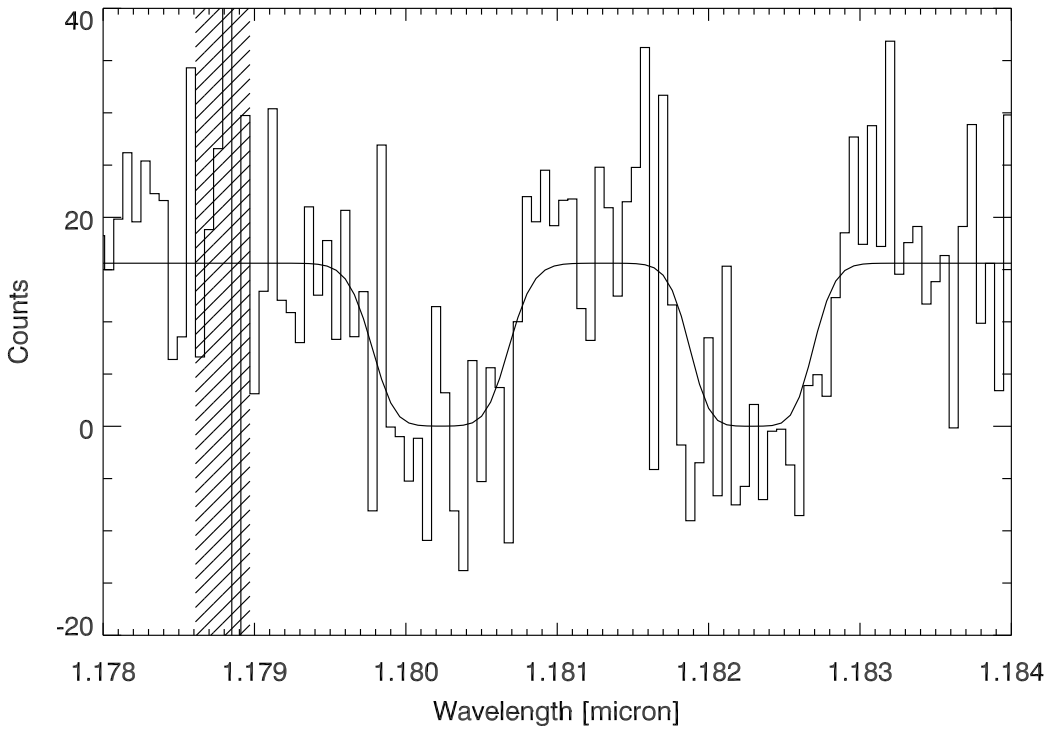

Figure 17.7: Example of the fit of two absorption line systems in a galaxy at $z = 8$, taken from run s4_50h_r10000 (see Table 17.2). The goal of these simulations was to recover the physical parameters of the model galaxies. The fit recovers the input column density of both lines within a factor of 2 and the line centre within 20 km s^{-1}.

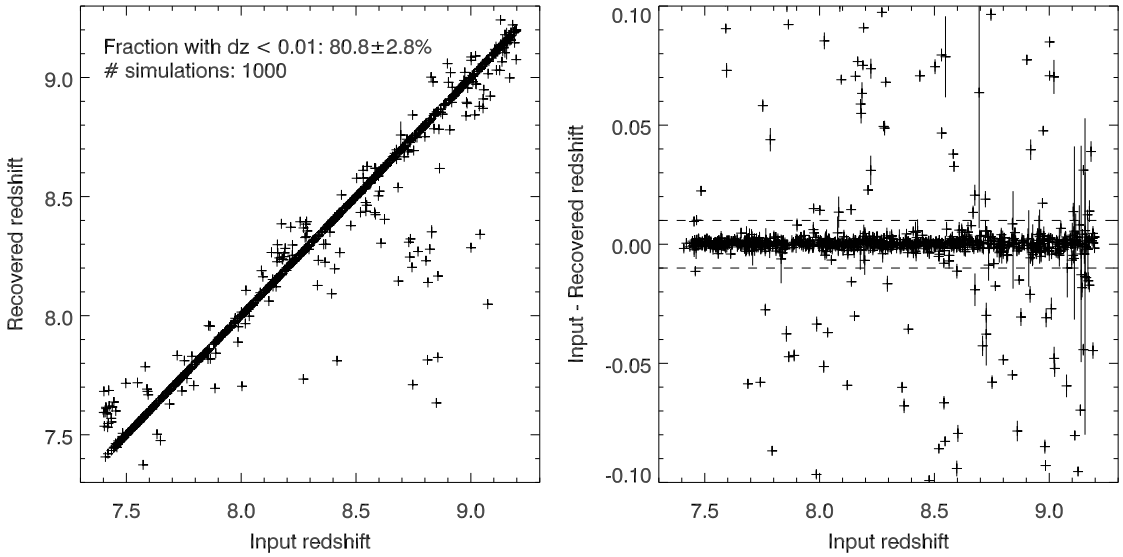

Figure 17.8: Results obtained for simulations done with all parameters set to the reference case. The figure on the left shows the recovered redshift as function of input redshift. The second plot shows a zoom of the redshift errors. The total fraction of galaxies for which the recovered redshift was within 0.01 of the input redshift is 80.8%. If the redshifts would have been measured manually from the 1D spectra, the fraction would be higher, and approximately 90% would have a redshift within 0.01 of the input redshift.

17.5 Results of simulations

17.5.1 Simulation runs

Several sets of simulations were performed to investigate the influence of the various model assumptions and input parameters on the results of the simulations. By default the parameters were set to the value of the reference case (see Section 17.4.1). We explored the impact of various input parameters on the results. We divided the model parameters into three broad groups and studied the effect on the outcome of the simulations for each group separately. The three groups of parameters are:

1. **Intrinsic parameters:** We simulated galaxies with different physical properties, to reach the specific goals of these DRM simulations as outlined in Section 17.2. The different parameter values of these simulations are listed in Table 17.2.
2. **Technical parameters:** To look into the impact of the instrument and AO mode chosen in our reference case, we ran simulations with a range of different technical parameters. The simulations, listed in Table 17.3, use two different instrument types and several distinct PSFs.
3. **Site parameters:** A third set of simulations was conducted to study the influence of the site, telescope coating and atmospheric conditions (Table 17.4).

17.5.2 Analysis

Fig. 17.8 shows our results for simulation run s1_1800, our reference case. On the left we show the input redshift (z_i) versus the recovered redshift (z_r) and on the right the redshift error ($\Delta z = z_i - z_r$) is shown as a function of input redshift. The fraction of galaxies with $|\Delta z| < 0.01$ is 80.8%. A group of points is visible above general trend at $z_i \sim 7.4$–7.5. For galaxies at these redshifts, the Lyman-α break is shifted to a wavelength of < 1.034 μm, which is not covered by the simulated spectra. Hence the pipeline cannot make a first guess of the redshift based on the location of the break, and is therefore less successful in subsequently locating the absorption lines. A higher success rate in finding the right redshift is achieved

Table 17.2: First set of simulations, conducted to investigate the physical properties of the input galaxies. The last column lists the number of simulated galaxies in the run. For the telescope, instrument, site and atmosphere the reference values were assumed. To summarise: the telescope had a bare aluminium coating and was located on a Paranal-like site, the instrument simulated was an IFU with 2048 spectral pixels, a spaxel scale of 37.5 mas pixel^{-1} and a FoV of 1.65″. For the PSF the Maory analytical PSFs in 0.6″ seeing were taken for the s1 and s3 runs, and an LTAO PSF was assumed for the s4 simulations.

Run	type	$H_{\rm AB}$ [mag]	redshift	Resolution	$t_{\rm exp}$ [s]	#
s1_1800	All	26.0±0.2	7.4–9.2	4000	1800	1000
s1_900	All	26.0±0.2	7.4–9.2	4000	900	100
s1_2400	All	26.0±0.2	7.4–9.2	4000	2400	200
s1_3600	All	26.0±0.2	7.4–9.2	4000	3600	100
s1_7200	All	26.0±0.2	7.4–9.2	4000	7200	100
s1_1350_dco	DCO	26.0±0.2	7.4–9.2	4000	1350	342
s1_1200_dco	DCO	26.0±0.2	7.4–9.2	4000	1200	342
s1_2400_haro11	Haro11	26.0±0.2	7.4–9.2	4000	2400	306
s3_25h	All	28.5±0.2	7.4–9.2	4000	90 000	78
s3_40h	All	28.5±0.2	7.4–9.2	4000	144 000	100
s3_45h	All	28.5±0.2	7.4–9.2	4000	162 000	100
s3_50h	All	28.5±0.2	7.4–9.2	4000	180 000	300
s4_10h_r4000	DCO	27.0	8.0	4000	36 000	500
s4_10h_r10000	DCO	27.0	8.0	10 000	36 000	500
s4_25h_r4000	DCO	27.0	8.0	4000	90 000	500
s4_25h_r10000	DCO	27.0	8.0	10 000	90 000	500
s4_50h_r4000	DCO	27.0	8.0	4000	180 000	500
s4_50h_r10000	DCO	27.0	8.0	10 000	180 000	500

Table 17.3: Simulations done to study the dependence on the exposure time necessary to reach the first goal of this DRM case on several technical parameters. In all runs the telescope was located on a Paranal-like site and the mirrors had a bare aluminium coating. The resolution was 4000, except for run s1_1800_r6000, where the resolution was set to 6000.

Run	PSF	Seeing [arcsec]	Instrument	Spaxel scale [mas]	FoV [arcsec]	$t_{\rm exp}$ [s]	#
s1_1800	Maory	0.6	IFU	37.5	1.65	1800	1000
s1_1800_m0p8	Maory	0.8	IFU	37.5	1.65	1800	250
s1_3600_m0p8	Maory	0.8	IFU	37.5	1.65	3600	250
s1_1800_ltao	LTAO	0.8	IFU	37.5	1.65	1800	233
s1_3600_ltao	LTAO	0.8	IFU	37.5	1.65	3600	233
s1_5400_ltao	LTAO	0.8	IFU	37.5	1.65	5400	233
s1_1800_glao	GLAO	0.8	IFU	37.5	1.65	1800	200
s1_1800_pix75mas	Maory	0.6	IFU	75.0	1.65	1800	500
s1_1800_pix10mas	Maory	0.6	IFU	10.0	1.0	1800	100
s1_1800_r6000	Maory	0.6	IFU	37.5	1.65	1800	250
s2_0p5h	GLAO	0.8	Fibre	300	0.9	1800	500
s2_1h	GLAO	0.8	Fibre	300	0.9	3600	500
s2_1p5h	GLAO	0.8	Fibre	300	0.9	5400	500
s2_2h	GLAO	0.8	Fibre	300	0.9	7200	500
s2_2p5h	GLAO	0.8	Fibre	300	0.9	9000	500
s2_3h	GLAO	0.8	Fibre	300	0.9	10 800	500
s2_4h	GLAO	0.8	Fibre	300	0.9	14 400	500
s2_3h_fibre0p7	GLAO	0.8	Fibre	300	0.7	10 800	500
s2_3h_fibre1p1	GLAO	0.8	Fibre	300	1.1	10 800	500
s2_3h_noao	noAO	0.8	Fibre	300	0.9	10 800	500

Table 17.4: Simulations carried out to investigate various telescope, site and atmospheric parameters. In all simulations reference values are used for input parameters not listed in the table: the galaxies had a magnitude of $H_{\rm AB} = 26.0 \pm 0.2$, the instrument simulated was an IFU with 2048 spectral pixels, a spatial pixel scale of 37.5 mas pixel^{-1} and a FoV of 1.65″, the resolution was 4000, the PSF was the Maory analytical PSF with 0.6″ seeing, and the exposure time was 1800 s.

Run	Site[a]	OH lines[b]	Continuum[c]	Coating[d]	Airmass	#
s1_1800	P	1.0	1.0	Al	1.5	1000
s1_1800_hd	HD	1.0	1.0	Al	1.5	400
s1_1800_agal	P	1.0	1.0	Ag/Al	1.5	400
s1_1800_hd_agal	HD	1.0	1.0	Ag/Al	1.5	400
s1_1800_am1p0	P	1.0	1.0	Al	1.0	500
s1_1800_contx0p5	P	1.0	0.5	Al	1.5	500
s1_1800_contx0p2	P	1.0	0.2	Al	1.5	250
s1_1800_ohx2	P	2.0	1.0	Al	1.5	500
s1_1800_ohx0p5	P	0.5	1.0	Al	1.5	500

[a]Telescope site: P = Paranal-like; HD = High & Dry
[b]Relative strength of the OH lines
[c]Relative level of the atmospheric continuum
[d]Mirror coating: Al = bare aluminium; Ag/Al = protected silver/aluminium

when these spectra are analysed by eye, see the discussion in Section 17.4.2.5. At the high redshift end, the fraction of outliers increases due to the small number of absorption lines in the J-band spectrum.

In Fig. 17.9 the results of the simulation runs with varying "intrinsic" parameters as listed in Table 17.2 are presented. Galaxies with a more compact morphology require a shorter exposure time (by $\sim$25%) compared to a clumpy morphology to reach the same redshift accuracy. On the other hand, very extended galaxies need $>$35% more integration time. Part of this difference is caused by the pipeline and the exposure times become more similar when doing an optimal extraction as described in Section 17.4.2.5.

In Fig. 17.10 we show the outcome of the fitting of absorption lines in the spectra of $H_{\rm AB} = 27$ mag galaxies at $z = 8$. The fraction for which the column density was recovered to within a factor of 3 strongly differs between the various absorption lines. If the fitting has the velocity dispersion as an extra free parameter, the fraction is lower.

17.5.3 Compliance with figures of merit

The first two goals of the simulations were to determine the exposure times to accurately ($\Delta z < 0.01$) measure redshifts for $z > 7$ galaxies with $H_{\rm AB} = 26$ and $H_{\rm AB} = 28.5$. The estimated exposure times are roughly 0.5 hr with an IFU and 3 hr with a fibre-fed spectrograph (see Section 17.5.4 and Fig. 17.11) for a galaxy with $H_{\rm AB} = 26$ mag, and $\sim$100 times longer for galaxies with $H_{\rm AB} = 28.5$ mag. These exposure times are not uncommon in Large Programmes at the VLT, so such spectroscopic studies of $z > 7$ galaxies appear to be feasible with the E-ELT.

To derive physical parameters as detailed in Section 17.2 we need an exposure time of (at least) 25 hr. With this integration time, it is possible to accurately measure the column density of half of the absorption lines if the velocity dispersion of the absorbing gas is known. In 50 hr the goal is reached for all absorption lines. If the velocity dispersion is not known, in 50 hr the science goal can be reached for half of the absorption lines. These exposure times are large but realistic, and could even be reduced if one stacks several individual galaxy spectra similar to the work done at $z \sim 3$ (Shapley et al. 2003).

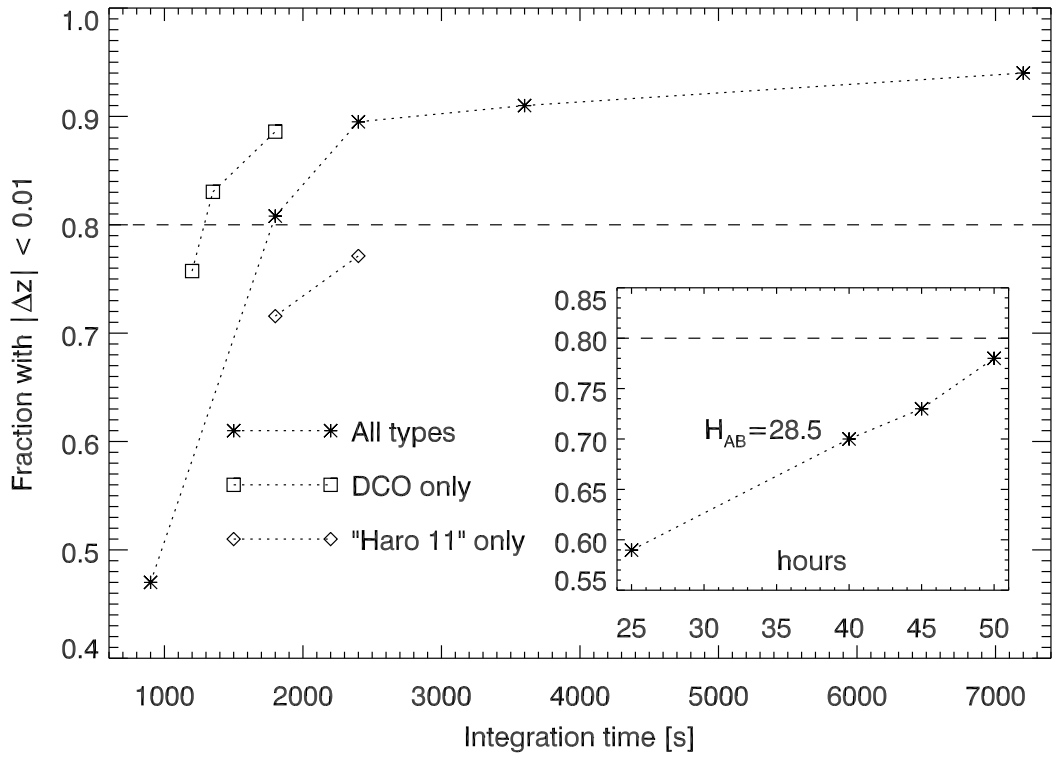

Figure 17.9: Fraction of galaxies with $H_{\rm AB} = 26$ mag with a recovered redshift within 0.01 of the input redshift as a function of integration time for various "intrinsic" parameter values. The inset shows the exposure time needed for galaxies that have $H_{\rm AB} = 28.5$ mag. At a given exposure time, the fraction depends on morphological type (Fig. 17.2). While "clumpy" galaxies follow the general trend, the DCOs are more easily confirmed (reducing the integration time by $\sim$ 25–30%), as the light of these objects is concentrated in the central few spaxels. On the other hand, for galaxies with a more extended morphology, like that of Haro 11, the exposure time needs to be increased by >35%. Most of this dependence on morphological type vanishes if the exact light distribution of the object is known (see Section 17.4.2.5).

17.5.4 Sensitivity to input parameters

To find out how the results of our simulations depend on our "technical" and "site" input parameters (described in more detail in Section 17.4.3.2), we compared the results of various simulation runs in which we changed some of the input parameters (see Tables 17.3 and 17.4 for an overview). Unless stated otherwise, the input parameters are set to the ones of simulation run s1_1800, the reference case (see Section 17.4.1).

Instrument: The type and characteristics of the instrument used had a lot of influence on the results of the simulations. In the case of an IFU instrument, the pixel size is of importance. If the pixel scale is much smaller than the reference size of 37.5 mas pixel^{-1}, the read-out noise noise (and dark current) start to become non-negligible in the regions between the sky lines, especially when simulating bright galaxies which require only short exposures. Also, for pixels with sizes around 10 mas pixel^{-1}, the dark current contributes up to 20% of the sky background. Obviously, for higher spectral resolutions, the dark current become even more important if using small pixels such as 10 mas.

In Fig. 17.11, we compare the results of simulations using different instrument types. An IFU is up to a factor of 6 more efficient as compared to a fibre-fed spectrograph (ignoring any additional light losses in the fibres). This is mainly due to the IFU taking advantage of the high resolution afforded by the MCAO PSF and because this correction and sampling maximises the contrast between object and sky and thus minimises the noise contribution from the night sky background. As $z > 7$ galaxies are (relatively) small ($\ll 1''$, see Fig. 17.2), a fibre with a diameter of $\sim$1$''$ will catch most of the light of the galaxy, independent

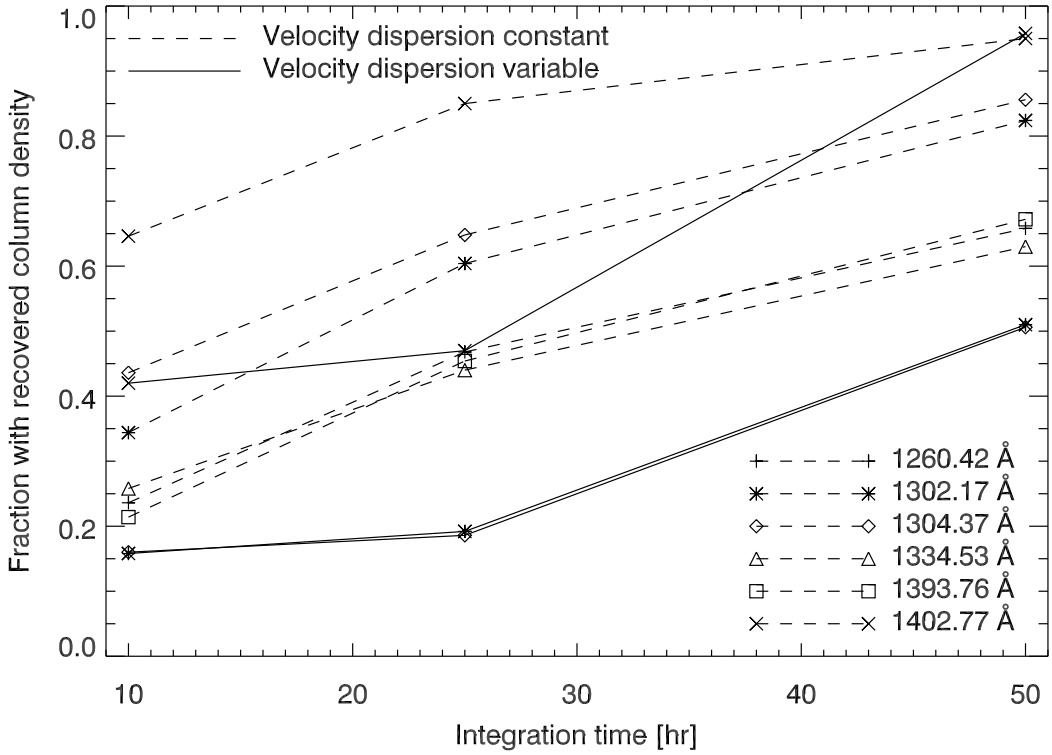

Figure 17.10: Fraction of galaxies with $H_{\rm AB} = 27$ mag for which the absorption line column density could be recovered within a factor of 3 of the input value as a function of exposure time. The different symbols represent the six absorption lines visible in a J-band spectrum of a $z = 8$ galaxy. The resolution of the spectra was 10 000. The dotted lines show the column density measured with the velocity dispersion fixed to the input value. The solid lines show the same but with the velocity dispersion a free parameter.

of the AO correction (see the similar value of encircled energy at large aperture radii for several types of AO correction in Fig. 17.5). Although a fibre-fed spectrograph requires more observing time to accurately measure redshifts, the results can be obtained even without any AO correction, as long as the diameter of the fibre is close to the seeing width (see inset of Fig. 17.11). Of course, if it is not, but much larger, the performance of the fibre-fed spectrometer will be that much worse.

AO correction: As mentioned in the previous paragraph, an IFU will take advantage of good AO corrections, since high-resolution PSFs and the relative small sizes of high-redshift galaxies combine to minimise the amount of background photons collected and hence the noise. As described in Section 17.4.3.2, the AO correction assumed in the reference case is quite optimistic (see also Fig. 17.5). In Fig. 17.12, we compare the exposure times necessary to reach the goal of the first science case for different AO corrections. The specific PSFs used in the simulations have a large influence on the required integration time. Increasing the seeing from 0.6″ to 0.8″ doubles the exposure time needed, while using the LTAO PSF triples the integration time needed. Note that using GLAO will require an even longer integration time than with LTAO.

Site and coating: Both the site and the coating of the telescope have a significant influence on the results of the simulations. The High & Dry site has lower atmospheric absorption, compared to the Paranal-like site, which increased the fraction of the galaxies with a well determined redshift. The protected silver/aluminium coating has a higher reflectivity than the bare aluminium coating, and less integration time is needed to reach the same S/N in the spectra of the galaxies. If the telescope is located on the High & Dry site and the mirrors are coated with protected silver/aluminium, then the exposure times needed to reach to goals of the science case are 20–25% lower than for the reference site and coating.

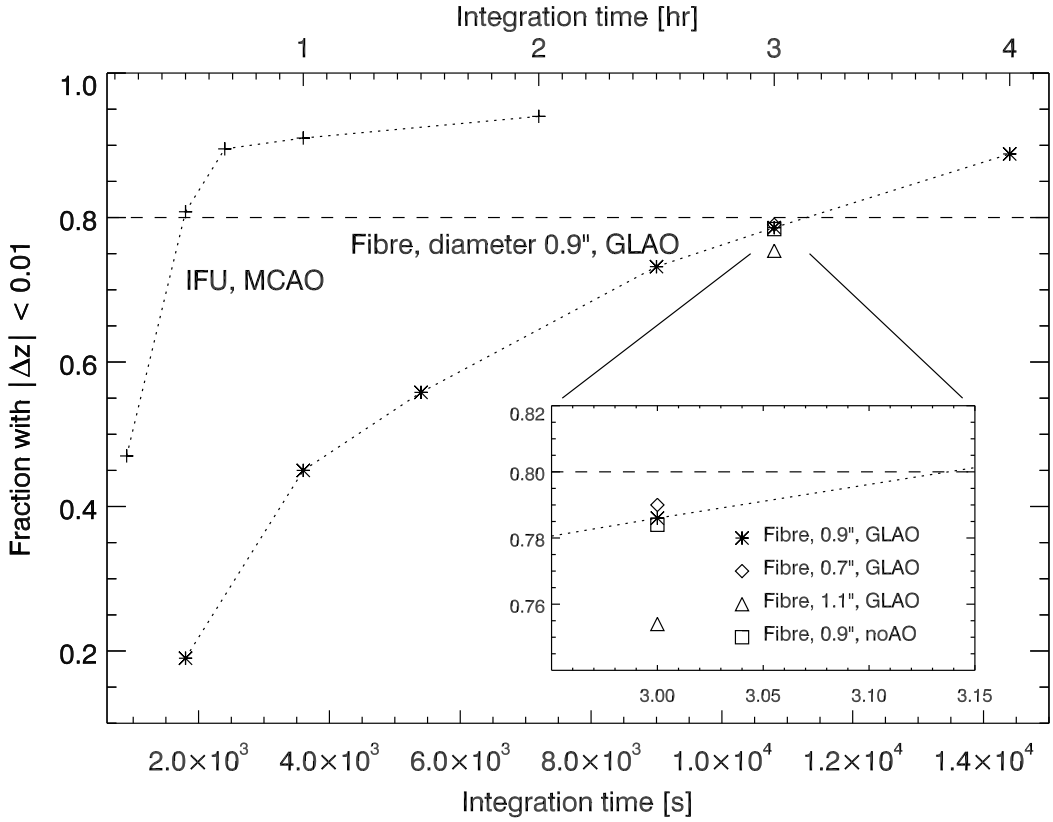

Figure 17.11: Success rate of accurately recovering the redshift for simulations with different instrument concepts. The two dotted lines show the simulations with an IFU-like instrument (top) and with a fibre-fed spectrograph (bottom). The inset shows the results for different fibre diameters. With a fibre all spatial information is lost, and the integration time needed to confirm 80% of the galaxies is of order 3 hr. This is up to 6 times longer than the IFU case (see also Fig. 17.12), although with the fibre the influence of AO correction is minimal, hence observations can be done in natural seeing as well (see inset).

Sky background: We explored the impact of the sky background on the outcome of our simulations by varying both the OH emission line strength and the (pseudo-)continuum. As noted before, the thermal emission from the atmosphere and telescope did not have an influence as we are observing in the J-band (see Fig. 17.4). Changing the strength of the atmospheric emission lines had a negligible effect on the results. This can be explained by the way we modelled the emission lines, namely as Gaussians. As we did not use regions in the spectra affected by the sky emission lines in our analysis, changing the strength of the lines had no consequence. In contrast, reducing the level of the sky continuum considerably improved our results. Simulations with the sky continuum level decreased by a factor of 5 reduced the integration time needed to reach the science goals by a factor $\sim$3. It is possible that the sky continuum level as measured by various groups is (in part) determined by the Lorentzian wings of the sky emission lines (Ellis & Bland-Hawthorn 2008). Suppressing the OH lines in the instrument (e.g. by using OH suppression fibres) could greatly reduce the sky continuum level and consequently reduce the integration times.

17.5.5 Calibration requirements

Flat-fielding: In the simulations using an IFU-type instrument, we assumed perfect flat-fielding. To obtain a decent and scientifically usable flat-field, it is very important that the wavelength range and scale of the flat-field and science frames are the same and that the illumination is the same. Any imperfections in the flat-fielding will increase the noise in the spectrum and in consequence will reduce the accuracy with which the redshift and absorption line profiles can be measured.

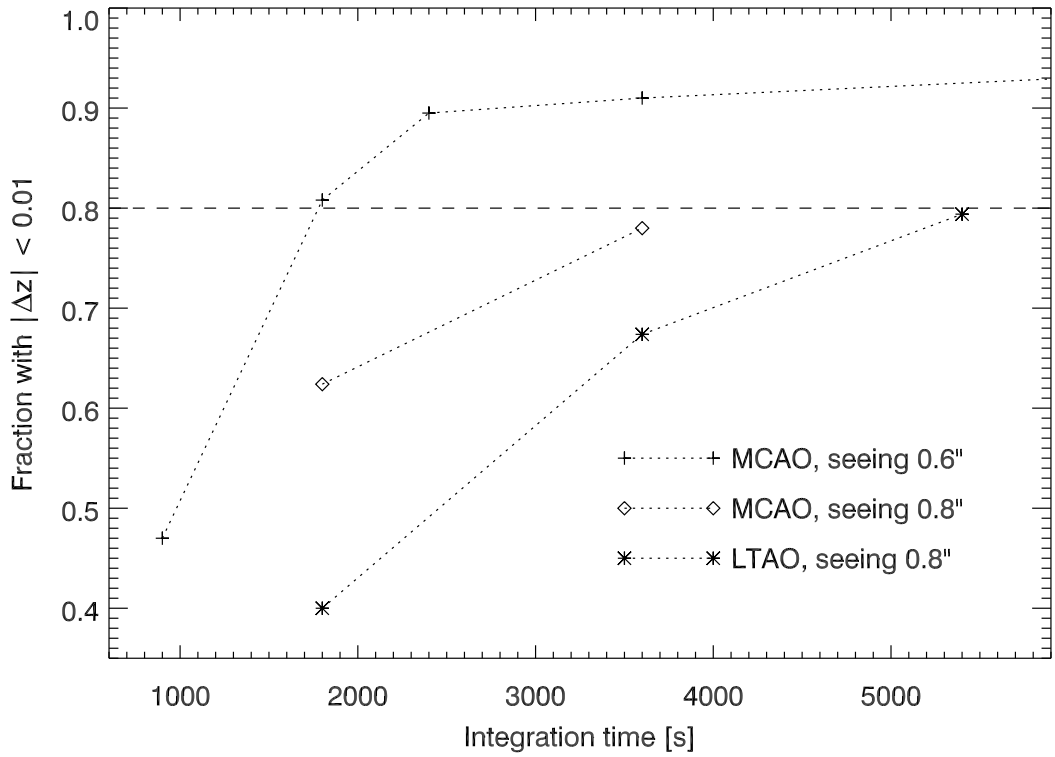

Figure 17.12: Integration time needed to reach the first goal of the science case as a function of AO correction. The type of AO correction has a large influence on the result. As shown in Fig. 17.5 the encircled energy differs greatly between the different AO modes. Using LTAO, the time needed to reach an 80% success rate of recovering the redshifts of $z > 7$ galaxies triples compared to the MCAO case.

Sky subtraction: For the case of a fibre-fed spectrograph, it is necessary to get an estimate of the sky level from somewhere else, e.g. another fibre. This requires accurate flat-fielding and an understanding of the differences in transmission efficiency and the spectral distortions.

Target acquisition: To optimally extract the spectra, it is essential to place the object on the IFU accurately. Especially when trying to extract the spectrum based on prior knowledge of the light distribution of the galaxy ("optimal" extraction, see Section 17.4.2.5), it is imperative to know the position of the galaxy on the IFU very precisely. We have not modelled this tolerance, but it seems logical that it must be a small fraction of a resolution element given the compactness of distant galaxies.

17.5.6 Limitations

The main limitations of the simulations done for this DRM science case are the somewhat simplistic scientific input data and various unrealistic steps in the pipeline reduction. Possible improvements include, e.g., varying the absorption line strength, adding velocity offsets to the absorption lines and the continuum slope of the input galaxies, and taking into account flat-fielding and wavelength calibration errors.

17.6 Concluding remarks

In this DRM case, we looked at several possible types of scientific programmes that the astrophysical community might be interested in. Although we mostly modelled galaxies with rather unrealistic magnitudes, $H_{\rm AB}$=26 and 27 mag, we can easily scale our results to fainter objects since all data are background or detector noise limited. Some of the simulations had galaxy magnitudes of $H_{\rm AB}$=28.5 mag. This is much more realistic in the sense that down to this magnitude, the surface density of $z \sim 7$–8 LBGs becomes

significant, $\sim$1 arcmin^{-2}. Hence, even though the surface density of distant LBGs is not well known, there will undoubtedly be several dozen LBGs within the E-ELT FoV. Even a spectrograph of moderate multiplex, say 20, will have sufficient targets to investigate.

One of the questions investigated here is how accurately we can recover redshifts of the objects using the UV absorption lines and how long of an integration time this requires. We found that if $\sim$80% of simulated galaxies have $\Delta z < 0.01$, the median $|\Delta z|$ is 0.0013, or about 40 km s^{-1}. In other words, 50% of the galaxies have a redshift that is within 40 km s^{-1} of the true redshift, which is much better than what can be achieved using the location of the break in the spectrum ($|\Delta z| \sim 800$ km s^{-1} for the galaxies simulated in the reference case) and/or the Lyman-α emission line, which can be systematically offset from the true redshift (e.g. Shapley et al. 2003). Such accurate redshifts will be important to study the physical clustering of and possible interactions between $z > 7$ galaxies. Measurements this accurate would enable a large number of studies on the dynamics of galaxies and (proto-)clusters and to find bona fide associations of physically related galaxies. And, depending on the exact type of spectrograph used, could be done in integration times of several 10s of hours per field.

In addition, we investigated how difficult it is to recover the physical characteristics of various UV absorption lines. These lines probe the physical conditions of the warm/hot ionised medium in the ISM and the stellar photospheres of stars in distant galaxies and probe a range of physical conditions ($T_e \sim 10^{4-6}$ K) and perhaps associated absorption from H I. At resolutions above a few thousand, we may expect the individual UV absorption lines to be resolved. If this is so, we can then estimate the column densities and velocity dispersions (as well as velocity offsets between lines but this does not require the lines to be resolved). This will tell use the covering fraction, outflow or inflow speeds, and perhaps abundance ratios of various ions that may be used to constrain the metallicity of these galaxies and hence what types of stars and stellar populations enriched the ISM of these distant galaxies. Moreover, the photospheric lines will tell us about the UV emitting stellar population in these galaxies and give us information about the initial mass function. While such studies will be very time consuming for the faintest LBGs currently known, $H_{\rm AB}$=28.5 mag, we have demonstrated that for the brighter distant galaxies, such a programme is very feasible for the large aperture of the E-ELT. Through a spectral stacking analysis, for which robust and accurate redshifts must be known, we may be able to recover a great deal of information about the "statistical faint LBG" at high redshift. This would provide a detailed look at the ISM and population of massive stars in galaxies within a few hundred Myr of the Big Bang – an exciting prospect indeed!

References

Becker G.D., Rauch M., Sargent W.L.W., 2009, ApJ, 698, 1010
Bouwens R.J., et al., 2010, ApJL, 709, L133
Bromm V., Yoshida N., Hernquist L., 2003, ApJL, 596, L135
Bunker A., et al., 2010, MNRAS, in press, arXiv:0909.2255
Ellis S.C., Bland-Hawthorn J., 2008, MNRAS, 386, 47
Greif T.H., Johnson J.L., Bromm V., Klessen R.S., 2007, ApJ, 670, 1
Heckman T.M., et al., 2005, ApJL, 619, L35
Kereš D., Katz N., Weinberg D.H., Davé R., 2005, MNRAS, 363, 2
Mac Low M., Ferrara A., 1999, ApJ, 513, 142
Madau P., Ferrara A., Rees M.J., 2001, ApJ, 555, 92
Maihara T., et al., 1993, PASP, 105, 940
McLure R.J., et al., 2010, MNRAS, 403, 960
Overzier R.A,. et al., 2009, ApJ, 706, 203
Pettini M., Rix S.A., Steidel C.C., Adelberger K.L., Hunt M.P., Shapley A.E., 2002, ApJ, 569, 742
Puech M., et al., 2008, MNRAS, 390, 1089
Ricotti M., Gnedin N.Y., Shull J.M., 2002, ApJ, 575, 49
Shapley A.E., Steidel C.C., Pettini M., Adelberger K.L., 2003, ApJ, 588, 65
Steidel C.C., Giavalisco M., Pettini M., Dickinson M., Adelberger K.L., 1996, ApJL, 462, L17
Swinbank A.M., et al., 2009, MNRAS, 400, 1121
Wise J.H., Abel T., 2008, ApJ, 685, 40

18 C2: A dynamical measurement of the expansion history of the Universe

Authors: J. Liske, M. Haehnelt, L. Pasquini

18.1 The science case

We propose to perform a direct observation of the expansion history of the Universe by obtaining the first measurement of the so-called 'redshift drift' (Sandage 1962). The redshift drift is a very small, but systematic drift as a function of time that is expected to affect the redshifts of all cosmologically distant sources partaking in the Hubble flow. This effect is induced by the evolution of the Hubble expansion, i.e. by its de- or acceleration, and it is a direct measure of the difference between the expansion rate today and the expansion rate at the redshift of the measurement. Hence, a redshift drift experiment would be unique among cosmological observations in that it is entirely model-independent, and directly probes the global dynamics of the Robertson-Walker metric, thereby offering a novel and truly independent path to the expansion history of the Universe.

It has been established that the multitude of absorption lines seen in the spectra of high-redshift QSOs offer the best opportunity for a redshift drift measurement. The Lyman α (Lyα) forest is of particular interest in this respect (Loeb 1998). We therefore propose to observe the redshift drift by monitoring the redshifts of Lyα forest and other absorption lines towards a sample of very bright QSOs in the redshift range $2 < z < 5$ using the ultra-stable, high-resolution optical spectrograph on the E-ELT over a period of 10–20 yr.

By monitoring the redshift drift of these absorption lines over a wide redshift range we will measure the instantaneous expansion rate of the Universe today and the expansion rate at high redshift. The measurement of the expansion rate and its evolution at high redshift will be an important test of General Relativity. The measurement of the instantaneous expansion rate will test whether the Universe expands today at the rate expected from other observations which, generally, measure the expansion rate averaged over a hundred Myr or more.

18.2 Goals of the DRM simulations

The main goal of the DRM simulations is to investigate the feasibility of a redshift drift (or $\dot{z}$) experiment with the E-ELT. The redshift drift is a tiny effect (see Fig. 18.1 for reference) and so the key issue is obviously the accuracy to which one can determine redshifts (or, equivalently, radial velocities).

The simulations should assume that the accuracy to which a radial velocity shift between two spectra of the same object can be determined is only limited by photon-noise down to the cm/s level. In the absence of systematic instrumental effects this accuracy depends only on the intrinsic sharpness of the spectral features used, the number of features available and the signal-to-noise ratio (S/N) at which they are recorded (assuming that the features are resolved). In a photon-noise limited experiment the latter in turn depends only on the flux density of the source(s), the size of the telescope, the total combined telescope/instrument efficiency and the integration time. This parameter space should be explored in order to ascertain how well we can expect the E-ELT to measure the redshift drift, and hence to constrain the cosmic expansion history.

Specifically, this work should:

1. determine how the properties of the Lyα forest (the number and sharpness of the absorption features) translate to an accuracy, σ_v, to which one can determine a radial velocity shift, and how σ_v scales with the S/N of the spectra;

2. determine by how much σ_v can be improved by including other parts of a QSO's absorption spectrum in the analysis – such as the Lyβ forest and metal absorption lines;

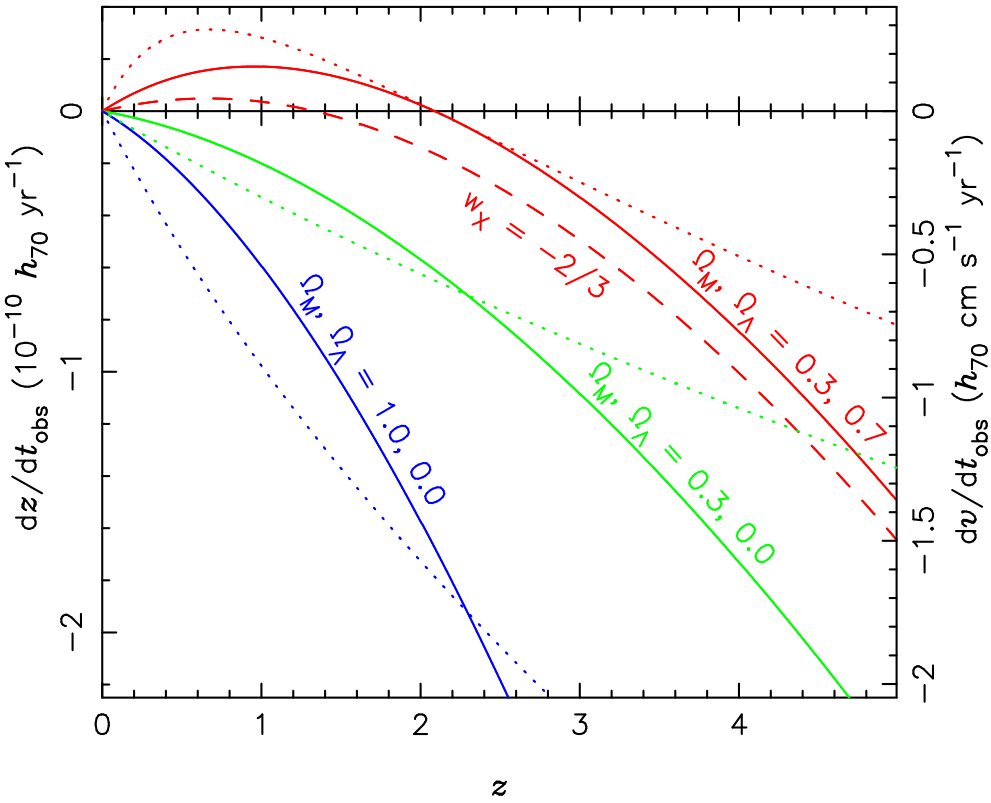

Figure 18.1: The solid (dotted) lines and left (right) axis show the redshift drift $\dot{z}$ ($\dot{v} = c\dot{z}/[1+z]$) as a function of redshift for various combinations of $\Omega_{\rm M}$ and Ω_Λ as indicated. The dashed line shows $\dot{z}$ for the case of dark energy having a constant equation of state parameter $w_X = -\frac{2}{3}$ (and $\Omega_{\rm M}, \Omega_X = 0.3, 0.7$).

3. quantify the impact of a realistic observing time distribution over the duration of the experiment on the ability to measure $\dot{z}$.

4. assuming the known population of QSOs and using the results from the above points predict the overall accuracy of a $\dot{z}$ experiment as a function of telescope diameter, telescope/instrument efficiency and total integration time; predict the significance of a non-zero detection of the redshift drift;

5. predict the possible constraints on cosmological parameters.

18.3 Metrics / figures of merit

The metric with which to compare the results of different $\dot{z}$ experiments with different parameters shall be the value of the total, overall radial velocity accuracy of the experiment, $\sigma_v^{\rm tot}$. This quantity is defined in detail in Section 18.5.3.

We now formulate a (somewhat arbitrary) cut-off in terms of this metric which will be used to decide whether the goals of this science case can be met.

At $z \approx 2$–4 the difference between the $\dot{z}(z)$ curves for $(\Omega_{\rm M}, \Omega_\Lambda) = (0.3, 0.7)$ and $(0.3, 0)$ is about 0.6 cm s^{-1} yr^{-1} (see Fig. 18.1). We wish to achieve a 3σ detection of this difference in 15 yr. Hence this requires a radial velocity accuracy of $0.6/3 \times 15 = 3$ cm s^{-1}. We also impose the requirements that the experiment encompasses no more than 2500 h total observing time and includes at least 20 different QSOs.

18.4 DRM simulations

18.4.1 Methodology

We will investigate the issues raised in Section 18.2 in two main steps. First, addressing points 1–3, we require a quantitative relation that translates the properties of a QSO absorption spectrum (i.e. the number, sharpness and S/N of its spectral features) into a quantity that measures the accuracy with which radial velocity shifts can be measured. We do not attempt to derive this relation analytically but instead use simulations of QSO absorption spectra. Secondly, addressing points 4–5, we use this relation together with a list of known potential QSO targets to predict the size of the error bars on realistic $\dot{z}$ measurements, and how these constrain cosmological parameters. Most of the remainder of this section describes the implementation of the simulations. Step 1 of the above programme is executed in Section 18.5.2 while the results pertaining to step 2 are described in Section 18.5.3.

18.4.2 Simulating QSO absorption spectra

18.4.2.1 Overview

High-resolution observations have demonstrated that, to first approximation, all absorption spectra can be decomposed into a collection of individual absorption lines (e.g. Kim et al. 2001). These are usually taken to be Voigt profiles and so each line is characterised by three parameters: redshift, z, column density, N, and velocity width, b. Here, we will reverse this decomposition process and generate (normalised) spectra with the desired instrumental characteristics from given lists of absorption lines.

The above approach allows us to quickly generate large amounts of spectra with realistic characteristics (as opposed to using proper hydrodynamical or semi-numerical simulations of the IGM). Note, however, that we make no assumptions regarding the underlying physics of the IGM in which the absorption occurs. We simply use the observational fact that QSO absorption spectra can be well represented as collections of Voigt profiles.

Given points 1 and 2 of Section 18.2 we will consider three different 'types' of spectra: (i) Lyα forest only; (ii) Lyα + Lyβ forest; (iii) Lyα + Lyβ forest + metal lines.

When referring to the Lyα forest we mean the plethora of absorption lines observed in the spectra of all QSOs in the spectral range between the QSO's Lyα and Lyβ emission lines. Almost all of these lines arise in the intervening intergalactic H I between us and the QSO. By construction the Lyα forest region cannot contain any H I transitions that are of higher order than Lyα, and it covers an absorption redshift range of $\lambda_\beta/\lambda_\alpha(1+z_{\rm QSO})-1 < z < z_{\rm QSO}$, where $z_{\rm QSO}$ is the QSO's redshift and $\lambda_\alpha = 1215.67$ Å and $\lambda_\beta = 1025.72$ Å are the rest wavelengths of the H I Lyα and Lyβ transitions, respectively.

Of course, each H I absorption line of the Lyα forest has corresponding counterparts at shorter wavelengths that result from the higher order transitions of the Lyman series. Obviously, any arguments concerning the suitability of the Lyα forest for a $\dot{z}$ measurement also apply to these higher order lines, and so, in principle, they should be almost as useful as the Lyα lines (where the qualifier 'almost' is owed to the decreasing optical depth with increasing order). However, in practice one has to contend with (i) confusion due to the overlap of low-redshift, low-order lines with high-redshift, high-order lines, and (ii) with an increased uncertainty in the placement of the QSO's continuum. Both of these difficulties are aggravated as one proceeds up the series towards shorter wavelengths. For these reasons we will not consider any H I transitions that are of higher order than Lyβ.

When considering the Lyβ forest in addition to the Lyα forest we extend our simulated spectra by adding the region between a QSO's Lyβ and Lyγ emission lines immediately bluewards of the Lyα forest. This region extends the redshift path for Lyα lines by a factor of 1.28 towards lower redshifts and also contains Lyβ lines in the redshift range $\lambda_\gamma/\lambda_\beta(1+z_{\rm QSO})-1 < z < z_{\rm QSO}$, where $\lambda_\gamma = 972.54$ Å.

In addition to the H I absorption, every high-redshift QSO spectrum shows absorption lines from a number of other ions, such as C IV, Si IV or Mg II. Although these metal lines are far less numerous than the H I lines, they are also much narrower: their widths are of order a few km s^{-1}. In fact, many metal lines

are unresolved in current spectra and so the widths of the narrowest lines are unknown. This suggests that metal lines may considerably increase a spectrum's sensitivity to radial velocity shifts and may hence supply valuable additional constraints on $\dot{z}$.

18.4.2.2 Generating spectra

As mentioned above, we will create spectra from given lists of absorption lines. We will use two types of line lists. First, for the H I absorption lines, we will generate line lists from Monte Carlo (MC) simulations based on the statistics of the largest available samples of absorption lines. Secondly, to validate the MC simulations and to gauge the effect of metal lines, we will use 8 line lists available in the literature that have previously been derived from high-resolution observations. Both types of lists are described in detail in Section 18.4.4.1.

Given an absorption line list a normalised spectrum is generated by

$$S(\lambda) = \exp\left\{-\sum_{i}^{N_{\rm al}}\sum_{j}^{N^{i}_{\rm trans}} \tau[\lambda_{{\rm r},ij}(1+z_i), N_i, b_i]\right\}, \tag{18.1}$$

where λ is the observed wavelength, $N_{\rm al}$ is the number of absorption lines in the spectrum (defined as the number of unique sets of {ion, z, N, b}), τ is the optical depth of a Voigt profile, $\lambda_{{\rm r},ij}$ is the rest wavelength of the jth transition of the ion of the ith absorption line with column density N_i and velocity width b_i. This spectrum is then pixelised, using a pixel size of 0.0125 Å, and convolved with a Gaussian line-spread function. Unless stated otherwise, we will use a resolution element four times the pixel size, corresponding to a resolution of $R = 100\,000$ at 5000 Å. We then add random noise to the spectrum assuming Poisson statistics, i.e. we assume a purely photon-noise limited experiment. All S/N values quoted in this work refer to the S/N per pixel in the continuum.

To reduce computation time we restrict the wavelength range of a simulated spectrum to the region of interest: for spectra generated from MC line lists (H I only) the wavelength range corresponds to the redshift ranges of the Lyα or Lyα + Lyβ forest regions as described above, and are hence entirely defined by the redshift, $z_{\rm QSO}$, of the assumed background QSO. For spectra generated from real line lists we use the wavelength ranges that were used in deriving these lists from the original data (see Tables 18.1 and 18.2).

18.4.3 Pipeline

The simulation pipeline is quite straightforward and is illustrated in Fig. 18.2. First, we have to decide whether to simulate a spectrum from a MC line list or from a real line list. In the following we will refer to these as the 'MC' and 'real' cases, respectively. Then we decide which absorption lines to include in the spectrum: Lyα only or Lyα and Lyβ lines (MC case) or Lyα + Lyβ + metal lines (real case). Next, we freely choose the redshift of the putative background QSO in the MC case, or select one of the 8 QSOs listed in Tables 18.1 and 18.2 in the real case. Then we select values for the remaining free parameters of the simulation: the S/N and resolution of the spectra (and, in the MC case, whether or not to include the clustering of lines). Next we either generate a MC line list or retrieve a copy of the line list of the previously chosen real QSO which may need to be supplemented with MC lines (see Section 18.4.4.1 for details on the line lists). Since we would like to simulate a $\dot{z}$ measurement we require a second epoch 'observation' of the same line of sight: we generate a second epoch absorption line list from the original line list by simply shifting the redshift of each line according to a given cosmological model:

$$\Delta z_i = \dot{z}(z_i; H_0, \Omega_{\rm M}, \Omega_\Lambda)\, \Delta t_0, \tag{18.2}$$

where Δt_0 is the assumed time interval between the first and second epochs, and may be freely chosen. Unless stated otherwise, we use the standard, general relativistic cosmological model, assuming fiducial parameter values of $H_0 = 70\ h_{70}$ km s^{-1} Mpc^{-1}, $\Omega_{\rm M} = 0.3$ and $\Omega_\Lambda = 0.7$. Finally, we generate spectra for the line lists of both epochs as described in Section 18.4.2.

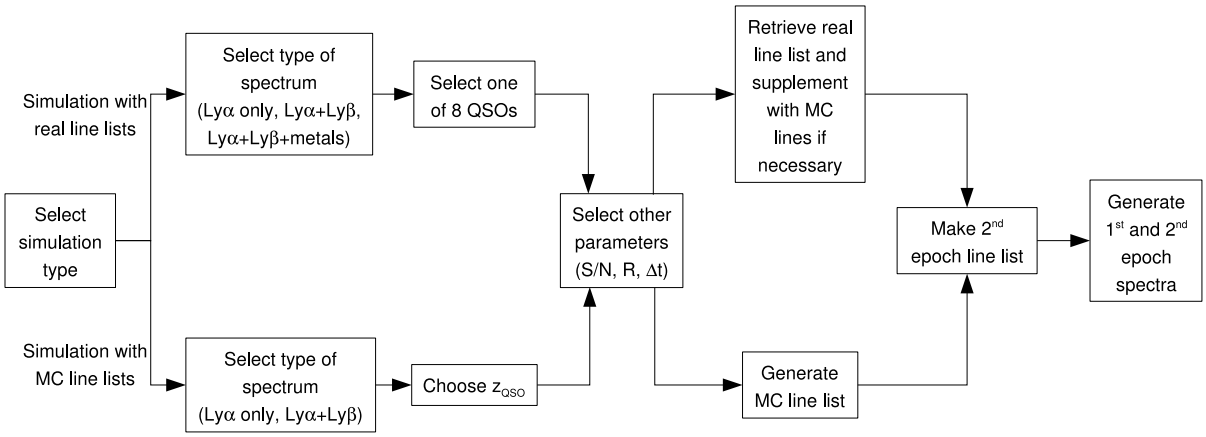

Figure 18.2: Flowchart of simulations.

The whole process described above may be repeated as often as required, where any of the parameters may be changed at any point.

All parts of the simulations are carried out using custom-built software.

18.4.4 Inputs

18.4.4.1 Scientific data

In this section we list the various scientific data that are required for the simulations and the analysis. The real absorption line lists and the QSO catalogue used in this work are available at http://www.eso.org/sci/facilities/eelt/science/drm/C2/.

Simulated absorption line lists: We form simulated MC H I line lists by simply randomly drawing values for the absorption line parameters from their observed distributions (e.g. Hu et al. 1995; Lu et al. 1996; Kirkman & Tytler 1997; Kim et al. 1997, 2001, 2002):

$$f(z, N_{\rm HI}, b) \;\propto\; (1+z)^{\gamma}\; N_{\rm HI}^{-\beta}\; \exp\left[-\frac{(b-\bar{b})^2}{2\sigma_b^2}\right], \tag{18.3}$$

where $\gamma = 2.2$, $\beta = 1.5$, $\bar{b} = 30$ km s^{-1} and $\sigma_b = 8$ km s^{-1}. We impose limits of $15 < b < 100$ km s^{-1} and also restrict $N_{\rm HI}$ to the classical Lyα forest regime, excluding Lyman limit and DLA systems: $12 < \log N_{\rm HI}({\rm cm}^{-2}) < 16$. Redshift limits are chosen as described in Section 18.4.2. The above distribution is normalised to give 10^2 lines with $13.64 < \log N_{\rm HI}({\rm cm}^{-2}) < 16$ per unit redshift at $z = 2$ (Kim et al. 2001). The actual number of absorption lines in a given line list is drawn from a Poisson distribution with a mean determined by the normalisation.

The most significant difference between our random MC line lists and the real Lyα forest is clustering: real H I lines are not randomly distributed in redshift but show significant redshift-space correlations on scales of at least 100 km s^{-1} (e.g. Cristiani et al. 1995; Fernández-Soto et al. 1996; Liske et al. 2000). We will investigate the effect of line clustering (see Section 18.5.2.2) by generating a separate set of MC line lists which incorporate the following toy clustering scheme: first, we randomly draw the positions of 'cluster' centres from the H I absorption line redshift distribution. We then populate each 'cluster' with $n \geq 0$ absorbers, where n is drawn from a Borel distribution (Saslaw 1989). Since n can be 0, 1 or > 1 this process generates voids, single 'field' absorbers, as well as groups and clusters of lines. Finally, absorption lines are distributed around their 'host' clusters according to a Gaussian distribution with FWHM $= 120$ km s^{-1}.

Table 18.1: Observed Lyα forest line lists from the literature.

QSO	$z_{\rm QSO}$	λ range[a]	$N^b_{{\rm Ly}\alpha}$	Reference
Q1101−264	2.145	3226−3810	290	1
J2233−606	2.238	3400−3850	226	2
HE1122−1648	2.400	3500−4091	354	1
HE2217−2818	2.413	3550−4050	262	2
HE1347−2457	2.617	3760−4335	362	1
Q0302−003	3.281	4808−5150	223	1
Q0055−269	3.655	4852−5598	535	1
Q0000−26	4.127	5380−6242	431	3

References: 1 = Kim et al. (2002), 2 = Kim et al. (2001), 3 = Lu et al. (1996).
[a]Wavelength range covered by the Lyα forest line lists in Å.
[b]Number of Lyα forest lines.

Equation (18.3) gave a succinct parameterisation of the properties of the H I absorption lines. Unfortunately, equivalents for all the various metal line species do not exist. Hence we are unable to gauge the effects of metal lines using MC simulations.

Real absorption line lists: We have collected 8 H I absorption line lists from the literature (see Table 18.1). These were derived from UVES/VLT (7 objects) or HIRES/Keck data (1 object). All spectra have a resolution of FWHM ≈ 7 km s^{-1}, while the typical S/N per pixel varies from $\sim$10 for Q0000−26 to $\sim$50 in the case of HE2217−2818. In all cases the absorption line lists were derived by fitting Voigt profiles to all features found in the Lyα forest regions of the spectra using VPFIT.[8] Details of the data acquisition and reduction, as well as of the line fitting and identification processes are given by Kim et al. (2001, 2002) and Lu et al. (1996).

We note that two of the spectra, those of Q1101−264 and Q0000−26, contain DLAs, which effectively block out parts of theLyα forest regions of the spectra. However, this is only expected to have a significant effect at high redshift, where the Lyα forest line density is high, and so we have excluded the affected spectral region only in the case of Q0000−26.

Note that these H I line lists do not extend below the QSOs' Lyβ emission lines (cf. Table 18.1). Hence, when generating a spectrum including the Lyβ forest region, we are forced to supplement the real H I line list with additional lines at lower redshifts drawn from the MC simulations. In other words the spectral region below the background QSO's Lyβ emission line contains the Lyβ transitions from the real H I lines as well as the Lyα transitions from lower redshift MC lines.

The authors of the real line lists also searched their spectra for metal lines. Unfortunately, they only published measured parameters for lines lying in the Lyα forest region. Although these lines will also have transitions elsewhere (which were probably used in the identification process) the published line lists do not provide a complete view of the metal line population outside of this region.

To improve on this situation we went back to the spectra of two of our QSOs, Q1101−264 and Q0000−26, and derived our own metal line lists covering the entire spectral range available to us. In the case of Q1101−264 the spectrum was the same as that used by Kim et al. (2002), whereas for Q0000−26 we used a UVES/VLT spectrum of similar quality as the HIRES/Keck spectrum used by Lu et al. (1996). In the rest of this work we will refer to these two line lists as the 'complete' metal line lists. Table 18.2 summarises the wavelength ranges that were searched for metal lines and the total number of metal absorbers that were found in each QSO spectrum.

High-redshift QSOs: For step 2 of the programme outlined in Section 18.4.1 (i.e. to address points 4 and 5 of Section 18.2) we will require a realistic list of potential targets for a $\dot{z}$ experiment. Here we do not wish to speculate on possible future discoveries of QSOs and hence we will restrict ourselves to those already known today. In the following we will hence extract a list of potential target QSOs from existing

[8]by R.F. Carswell et al., see http://www.ast.cam.ac.uk/∼rfc/vpfit.html.

Table 18.2: Observed lists of metal absorbers.

Name	λ range[a]	$N_{\rm m}^{b}$	Reference
Q1101−264	3050−5765, 5834−8530	225	this work
J2233−606	3400−3850	49	2
HE1122−1648	3500−4091	18	1
HE2217−2818	3550−4050	59	2
HE1347−2457	3760−4335	48	1
Q0302−003	4808−5150	5	1
Q0055−269	4852−5598	14	1
Q0000−26	4300−6450, 7065−8590, 8820−9300	100	this work

References: 1 = Kim et al. (2002), 2 = Kim et al. (2001).
[a]Wavelength range(s) searched for metal lines in Å.
[b]Number of metal absorbers, defined as the number of unique sets of {ion, z, N, b}.

QSO catalogues. For each candidate target QSO we will need a reliable magnitude that can be used to estimate its photon flux, as well as its redshift.

The largest QSO catalogue with reliable, homogeneous photometry and redshifts currently available is the fourth edition of the SDSS Quasar Catalog (Schneider et al. 2007). Being based on the fifth data release of the SDSS, it yields $16\,913$ QSOs with $z_{\rm QSO} \geq 2$. The catalogue provides PSF magnitudes in the $ugriz$ bands which we do *not* correct for Galactic extinction (as is appropriate for S/N calculations). Since we are interested in the continuum flux we will use, for each QSO, the magnitude of the bluest filter that still lies entirely redwards of the QSO's Lyα emission line. Specifically, for objects with $z_{\rm QSO} < 2.2$ we will use the g-band magnitude; for $2.2 \leq z_{\rm QSO} < 3.47$ the r-band; for $3.47 \leq z_{\rm QSO} < 4.61$ the i-band; and for $4.61 \leq z_{\rm QSO}$ the z-band. We then apply a small correction to the selected magnitude to transform the observed flux to that expected at the centre of the Lyα forest assuming a power-law spectral shape of the form $f_\nu \propto \nu^{-0.5}$ (Francis 1993).

Unfortunately, the SDSS catalogue does not cover the whole sky. The largest collection of QSOs covering the *entire* sky is the 12th edition of the Catalogue of Quasars and Active Nuclei recently published by Véron-Cetty & Véron (2006), which contains many additional bright QSOs not included in the SDSS catalogue. However, since the Véron catalogue is a compilation of data from many different sources its photometry is very inhomogeneous and cannot readily be converted to a photon flux. We will dispense with this inconvenience by matching the Véron catalogue to the all-sky SuperCOSMOS Sky Survey (SSS; Hambly et al. 2001). Although the photographic SSS photometry is not endowed with reliable absolute calibrations either, at least it is homogeneous and covers three bands ($B_{\rm J}$, R and I), allowing us to synthesise approximate SDSS magnitudes.

We proceed by first selecting all QSOs from the Véron catalogue with $z_{\rm QSO} \geq 2$ and not flagged as unreliable, resulting in $21\,895$ objects. For each of these we then identify the nearest SSS object, allowing a maximum separation of 5 arcsec, and retrieve the corresponding SSS catalogue data. 938 Véron objects have no SSS counterpart, presumably because their coordinates are unreliable. We then use those $11\,744$ objects that are in common to the SDSS and combined Véron-SSS catalogues to derive linear relations between the SDSS and SSS magnitudes, allowing for zeropoint offsets and colour terms. Such relations are reasonable representations of the data and we find that the distributions of the residuals have r.m.s. values of less than 0.3 mag in all cases. Finally, we purge the common objects from the Véron-SSS catalogue and use the above relations to synthesise an SDSS magnitude in the appropriate band (see above) for each of the remaining QSOs.

For those QSOs in the initial Véron catalogue that have no match in the SSS, or which are missing an SSS band needed to synthesise the required SDSS magnitude, we will simply use the V or R-band magnitude as listed in the Véron catalogue, provided it is non-photographic.

In summary, the final combined sample of $25\,974$ QSOs is constructed from three sub-sets: (i) SDSS; (ii) objects with redshifts from the Véron catalogue and photometry from the SSS (converted to the SDSS system); and (iii) objects where both the redshifts and the photometry are taken from the Véron catalogue. We note that the quality and reliability of the photometry decreases rapidly from (i) to (iii).

18.4.4.2 Technical data

For step 1 of our analysis and the simulation of the QSO absorption spectra we only need to assume some basic spectrograph characteristics:

We choose a sampling of 0.0125 Å and a resolution element four times the pixel size, corresponding to a resolution of $R = 100\,000$ at 5000 Å. This differs from the resolution specified in the CODEX concept study ($= 135\,000$) but we note that the results of this work do not depend at all on these choices, as long as $R \gtrsim 40\,000$. The reason is that the radial velocity accuracy is independent of spectral resolution as long as the spectral features used are resolved (cf. Fig. 18.8). In practice, a $\dot{z}$ experiment will gain from higher resolution because of the improved abilities to resolve the very narrowest metal lines (whose widths are currently unknown) and to deal with wavelength calibration issues. However, neither of these effects is included in the present simulations.

For the simulations we did not impose any wavelength limits, but rather simulated and used the full spectral region of interest as explained in Section 18.4.2. The reason is that at the time the core of this work was carried out the wavelength range in the CODEX concept study had not been decided on, and we did not wish to over-restrict the parameter space to be explored. We did, however, restrict ourselves to $2 \leq z_{\rm QSO} \leq 5$. The shortest and longest wavelengths to occur anywhere in our simulations are 2915 and 9300 Å. We will discuss the implications of a smaller available wavelength range in Section 18.5.5. We also assume that there are no gaps in the wavelength coverage.

For step 2 of our analysis we need to be able to calculate S/N values from QSO magnitudes and integration times. Since we are working in the photon-noise limited case we will use the following:

$$\frac{\rm S}{\rm N} = 700 \left[\frac{Z_X}{Z_r}\ 10^{0.4(16-m_X)}\ \left(\frac{D}{42\,{\rm m}}\right)^2\ \frac{t_{\rm int}}{10\,{\rm h}}\ \frac{\epsilon}{0.25} \right]^{\frac{1}{2}}, \tag{18.4}$$

where D, $t_{\rm int}$ and ϵ are the telescope diameter, total integration time and total efficiency, Z_X and m_X are the zeropoint and apparent magnitude of the source in the X-band, respectively, and $Z_r = (8.88 \times 10^{10})$ s^{-1} m^{-2} μm^{-1} is the AB zeropoint (Oke 1974, [RD1]) for an effective wavelength of 6170 Å (corresponding to the SDSS r-band). The normalisation of the above equation assumes a pixel size of 0.0125 Å (see above) and a central obscuration of the telescope's primary collecting area of $10\,\%$ which is slightly larger than the $8\,\%$ assumed in [RD1].

By 'total efficiency' we mean the ratio of the number of detected photo-electrons to the number of source photons at the top of the atmosphere, i.e. it comprises atmospheric absorption and all losses occurring in the combined telescope/instrument system, including entrance aperture losses and the detector's quantum efficiency. We chose not to split the total efficiency into its constituent factors because it is only the total that is relevant (and because the individual factors were largely not available at the time the core of this work was carried out). Nevertheless, let us write ϵ as

$$\epsilon = \epsilon_{\rm atm} \times \epsilon_{\rm tel} \times \epsilon_{\rm ent} \times \epsilon_{\rm inst} \times \epsilon_{\rm det}, \tag{18.5}$$

where the factors represent the atmospheric absorption, the throughput of the telescope, the losses at the entrance of the spectrograph, the throughput of the instrument itself and the detector's quantum efficiency, respectively. We now neglect the wavelength dependence of all of these factors and simply use the values at 5000 Å of [RD1], assuming a Paranal-like site, a protected silver/aluminium coating of the telescope mirrors, a circular entrance aperture of 400 mas radius and a GLAO PSF. We find:

$$\epsilon = 0.9 \times 0.86 \times 0.6 \times \epsilon_{\rm inst} \times 0.9 = 0.42 \times \epsilon_{\rm inst}, \tag{18.6}$$

where $\epsilon_{\rm inst}$ is the only factor lacking an estimate in [RD1]. Thus we see that a total efficiency of 0.25 implies an instrument throughput (without entrance losses and detector QE) of 0.6.

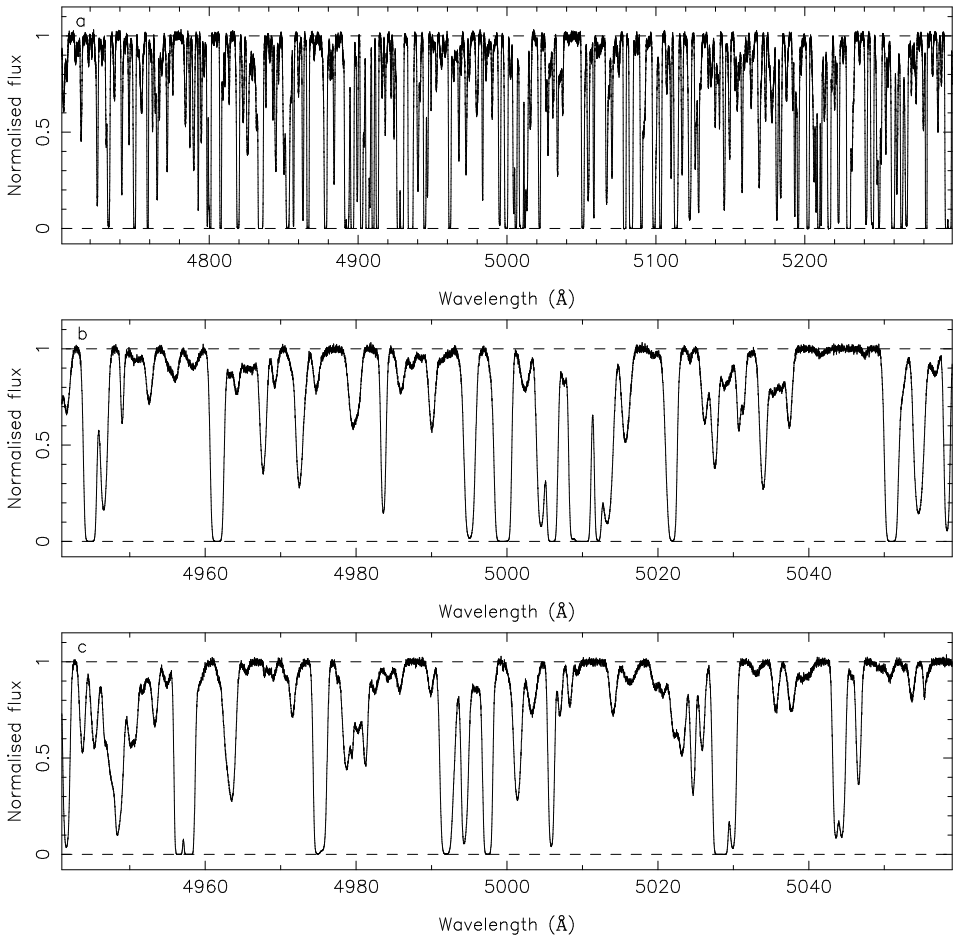

Figure 18.3: (a) Example of an artificial Lyα forest spectrum at $z \approx 3$ generated from a Monte Carlo absorption line list. (b) Close-up of the region around 5000 Å. (c) Artificial spectrum generated from the observed line list of Q0302−003. Both spectra have S/N $= 100$.

18.4.5 Outputs

The output of a single simulation run consists of one line list and one corresponding spectrum for each of two epochs. The spectra are generated in simple, multi-column ASCII files, containing wavelength, normalised flux and 1σ error arrays. Header files recording the parameters of the simulation are also generated. In Fig. 18.3 we show examples of artificial spectra generated from both MC (panels a and b) and real line lists (panel c). The expected flux difference between two spectra of the same absorption lines, taken a decade apart, is shown in Fig. 18.4.

Some example data are available at http://www.eso.org/sci/facilities/eelt/science/drm/C2/.

18.5 Results of simulations

18.5.1 Simulation runs

The various simulation runs and their parameters are summarised in Table 18.3. For each run we only list the values of those input parameters that were changed from their default values. Default parameter values are $z_{\rm QSO} = 4$, S/N$= 13\,000$, $R = 100\,000$, $\Delta t_0 = 10$ yr. Each simulation run is described in more detail in Section 18.5.2.

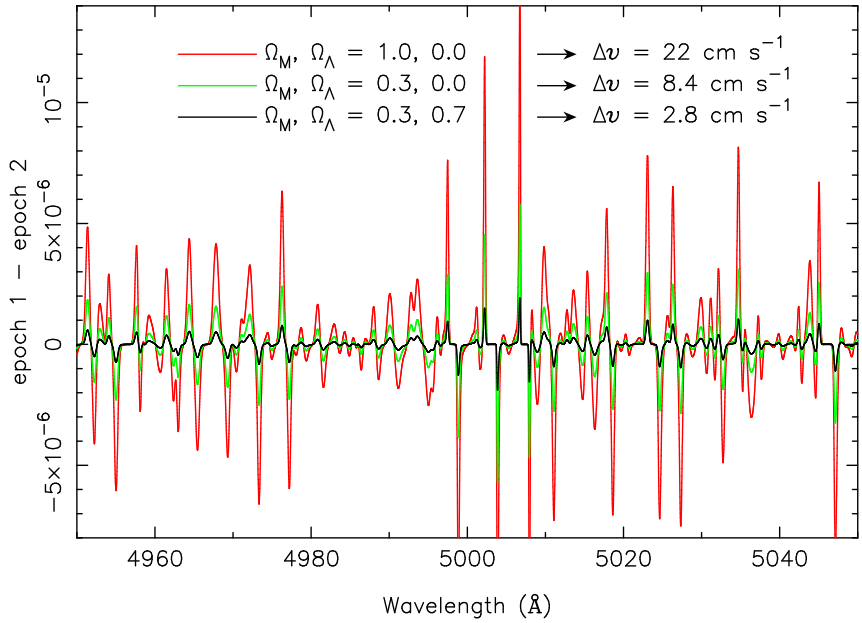

Figure 18.4: Flux difference between two artificial, noiseless spectra of the same Lyα forest at $z \approx 3$ simulated for two observing epochs separated by $\Delta t_0 = 10$ yr and for various combinations of $\Omega_{\rm M}$ and Ω_Λ as indicated. The redshift drift implied by these parameters is also given.

Table 18.3: Simulation runs and their parameters.

Run	Type of sim.[a]	Type of spectra	No. of spectra[b]	Non-default parameter
1	MCLL	Lyα	10	S/N$=100$–$30\,000$
2	MCLL	Lyα	10	$f = 0.1$–0.9[c]
3	MCLL	Lyα	10	$z = 1$–5
4	MCLL	Lyα	10	as 3 & $\gamma = 0$[d]
5	RLL	Lyα	1	
6	MCLL	Lyα	10	as 3 & clustering
7	MCLL	Lyα	10	$R = 2500$–$133\,000$
8	MCLL	Lyα + β	10	$z = 1$–5
9	RLL	Lyα + β	1	
10	RLL	Lyα + β + metals	1	

[a]MCLL = Monte Carlo line list, RLL = real line list.
[b]Number of repeat simulations at each value of the parameter that is being varied.
[c]f = fraction of total integration time spent on 1st epoch observation.
[d]γ = evolutionary index in equation (18.3).

18.5.2 Analysis I

18.5.2.1 Defining σ_v

In principle, a $\dot{z}$ determination will involve the measurement of radial velocity differences between the corresponding features of a pair of spectra of the same object taken several years apart. In order to analyse our simulated spectra and to execute step 1 of the programme outlined in Section 18.4.1 we first of all need a method to estimate the accuracy to which these differences can be determined. This requires knowledge of how exactly the measurement will be performed. However, a priori it is not at all obvious what the optimal signal extraction method might be. To proceed nevertheless we choose to base our analysis on the generic concept of the total radial velocity information content of a spectrum, which was developed by Bouchy, Pepe & Queloz (2001) in the context of optimising radial velocity searches for extra-solar planets.

Following Bouchy et al. (2001) we begin by expressing the flux observed in pixel i at the second epoch as a small perturbation on the first epoch flux in the same pixel:

$$S_{2i} = S_{1i} + \frac{\mathrm{d}S_i}{\mathrm{d}\lambda}\,\frac{\Delta v_i}{c}\,\lambda_i, \tag{18.7}$$

which defines a small velocity shift Δv_i for each pixel. λ_i is the observed wavelength of the ith pixel and $\mathrm{d}S_i/\mathrm{d}\lambda$ is the spectral slope of the flux at that pixel. To first order the slope does not change between the two epochs and hence it carries no epoch designation. Averaging the velocity shift over all pixels in a spectrum, using weights w_i, we have:

$$\Delta v = \frac{\sum_i \Delta v_i\, w_i}{\sum_i w_i}. \tag{18.8}$$

Clearly, the weight for the ith pixel should be chosen as the inverse variance of Δv_i. In calculating this variance we must differ from Bouchy et al. (2001). In the case of stars one of the spectra can be assumed to be a perfect, noiseless template, essentially because additional information on the same type of star can be used to define it. However, since every Lyα forest spectrum is unique we cannot make the same assumption here, so that in our case both spectra have noise. Hence we find:

$$\sigma^2_{v_i} = \left(\frac{c}{\lambda_i \frac{\mathrm{d}S_i}{\mathrm{d}\lambda}}\right)^2 \left[\sigma^2_{1i} + \sigma^2_{2i} \;+\; \frac{(S_{2i}-S_{1i})^2}{\left(\frac{\mathrm{d}S_i}{\mathrm{d}\lambda}\right)^2}\,\sigma^2_{S'_i}\right], \tag{18.9}$$

where σ_{1i} and σ_{2i} are the flux errors in the ith pixel of the first and second epoch spectra, respectively, and $\sigma_{S'_i}$ is the error on the slope of the flux at pixel i. We can see that a low weight is assigned to noisy pixels and to those that have a small gradient, i.e. pixels in the continuum or in the troughs of saturated absorption lines. Finally, with the above choice of weights the error on Δv is given by:

$$\sigma^2_v = \frac{\sum_i \sigma^2_{v_i} w_i^2}{\left(\sum_i w_i\right)^2} = \frac{1}{\sum_i \sigma^{-2}_{v_i}}. \tag{18.10}$$

The above process has the advantage of conveniently attaching a single figure of merit to a given pair of spectra in a non-parametric, model-independent way: σ_v simply represents the fundamental photon-noise limit of the accuracy to which an overall velocity shift between the two spectra can be determined. It is essentially just a measure of the 'wiggliness' and of the S/N of the spectra. However, we point out that σ_v does not entirely capture all of the information contained in a pair of spectra with respect to a $\dot{z}$ measurement. From Fig. 18.1 it is clear that the difference between the first and second epoch spectra is not simply an overall velocity shift. The second epoch spectrum will also be slightly compressed with respect to the first epoch spectrum because the redshift drift is larger at higher redshifts (i.e. longer wavelengths) than at lower redshifts. The additional information that is contained in this small alteration of the spectrum's shape is entirely ignored by the above method [because of the simple averaging operation in equation (18.8)] and hence it is clearly a sub-optimal method of estimating the sensitivity of a pair of spectra to $\dot{z}$. Despite this shortcoming we will retain the above definition of σ_v for the sake of its simplicity.

18.5.2.2 The σ_v scaling relation

We are now ready to derive the relation between σ_v and the properties of a QSO absorption spectrum using our simulated spectra. We will start out using only the Lyα forest spectra. Note that the number density of H I absorption lines, and hence the 'wiggliness' of a spectrum, depends on redshift (cf. equation 18.3). Therefore we are mainly looking for a relation between σ_v on the one hand and redshift and S/N on the other.

We begin by making two points regarding equation (18.9). First, it is clear from equation (18.9) that σ_v scales as (S/N)$^{-1}$, as expected for a photon-noise limited experiment. Secondly, we note that for a fixed total integration time (the sum of the integration times spent observing the first and second epoch spectra) the sum $\sigma_{1i}^2 + \sigma_{2i}^2$ in equation (18.9) takes on its minimum when the first and second epoch integration times are equal. Hence, the smallest possible σ_v is only achieved when the spectra of both epochs have the same S/N.

We now validate our simulations by checking that they reproduce these expected relations. For various values of S/N in the range $100 \leq$ S/N $\leq 30\,000$ we have generated 10 pairs of Lyα forest spectra from MC line lists as described in Section 18.4.3, each at $z_{\rm QSO} = 4$ (run 1 in Table 18.3). We then measured each pair's σ_v according to equation (18.10). The result is shown as blue dots in the left panel of Fig. 18.5, where each point and error bar is the mean and ± 1 r.m.s. of the 10 individual measurements at each S/N. We stress that we are plotting the expected accuracy of a velocity shift measurement performed on a single pair of spectra at a given S/N of a single target at $z_{\rm QSO} = 4$. We are *not* plotting the *combined* accuracy of 10 such pairs.

In the right panel of Fig. 18.5 we show our σ_v measurements from simulations at $z_{\rm QSO} = 4$ as a function of the fraction, f, of the fixed total integration time spent on any one of the two epochs (run 2 in Table 18.3). The total integration time is assumed to give S/N of $13\,000$ for both spectra when $f = 0.5$. Clearly, in both cases the simulations are behaving in exactly the way expected on the basis of equation (18.10), hence validating our procedures.

From Fig. 18.5 we may also read off our first result, namely the normalisation of the relation between σ_v on the one hand and $z_{\rm QSO}$ and S/N on the other: at $z_{\rm QSO} = 4$ the properties of the Lyα forest are such that they allow a radial velocity shift between 2 spectra of the same line of sight to be measured with an accuracy of 2 cm s^{-1}, provided that the 2 spectra have been recorded with an equal S/N of $13\,000$ each. Note that from now on we will always assume that the spectra of both epochs have the same S/N.

What happens when the spectra of multiple QSOs are considered? Consider a set of $N_{\rm QSO}$ targets that all lie at the same redshift, each of which has been observed at two epochs such that all $2N_{\rm QSO}$ spectra have the same S/N. Again, since we are considering a purely photon-noise limited experiment, σ_v should scale as $N_{\rm pix}^{-1/2}$ (where $N_{\rm pix}$ is the total number of independent data points in the sample) and hence also as $N_{\rm QSO}^{-1/2}$. Furthermore, in this ideal case, it is irrelevant how the *total* S/N is divided among the targets. Hence, for simplicity we will continue to use $N_{\rm QSO} = 1$ in the following.

We now examine the behaviour of σ_v as a function of redshift using the MC absorption line lists. Similar to the procedure described above, we have generated 10 pairs of Lyα forest line lists and spectra at each of various QSO redshifts in the range $2 \leq z_{\rm QSO} \leq 5$, where each spectrum was given a S/N of $13\,000$ (run 3 in Table 18.3). The result is shown as blue dots in Fig. 18.6, where each point and error bar again shows the mean and ± 1 r.m.s. of the 10 individual measurements at each redshift.

From Fig. 18.6 we can see that the radial velocity sensitivity improves rather rapidly with redshift for $z_{\rm QSO} < 4$, but the decrease is somewhat shallower at $z_{\rm QSO} > 4$. Overall, σ_v improves by a factor of almost 3 when moving from $z_{\rm QSO} = 2$ to 5. Specifically, we find:

$$\sigma_v \propto \begin{cases} (1+z_{\rm QSO})^{-1.7} & z_{\rm QSO} < 4 \\ (1+z_{\rm QSO})^{-0.9} & z_{\rm QSO} > 4 \end{cases} \tag{18.11}$$

The behaviour of σ_v as a function of redshift is due to the combination of several factors. The first is the redshift evolution of the Lyα forest line density (cf. equation 18.3). At higher redshift more spectral features are available for determining a velocity shift and so σ_v decreases. However, from $z \approx 4$ the absorption lines severely blanket each other and the number of sharp spectral features does not increase

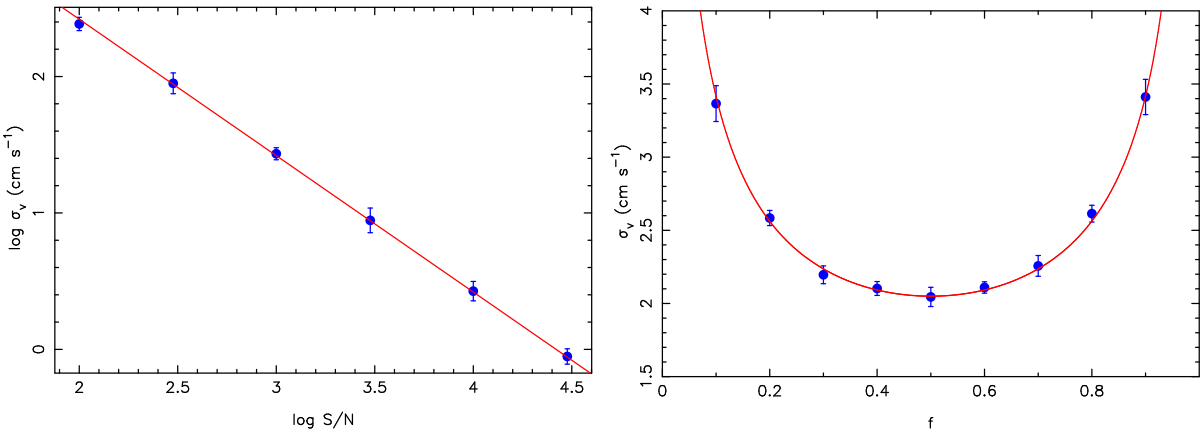

Figure 18.5: Left panel: the blue dots with error bars show the accuracy with which a radial velocity shift can be determined from a Lyα forest spectrum as a function of S/N. The red line shows the expected $\sigma_v \propto (\mathrm{S/N})^{-1}$ scaling. Each point is the mean σ_v measured from 10 pairs of artificial spectra at $z_{\mathrm{QSO}} = 4$, generated from simulated MC absorption line lists. The error bars show the ± 1 r.m.s. range of the 10 simulations. Right panel: σ_v as a function of the fraction, f, of the total integration time spent on observing the 1st (or 2nd) epoch spectrum, and its expected behaviour. Each point and error bar is the mean and ± 1 r.m.s. of 10 simulations at $z_{\mathrm{QSO}} = 4$ and a S/N that is equivalent to S/N $= 13\,000$ at $f = 0.5$.

as rapidly anymore, causing the flattening of σ_v at $z_{\mathrm{QSO}} > 4$. The green triangles in Fig. 18.6 show the σ_v measurements that result from simulations where the redshift evolution of the Lyα forest has been switched off (run 4 in Table 18.3), i.e. where the evolutionary index γ has been set to 0 (cf. equation 18.3). Indeed, we can see that in this case there is no evidence of a break.

Secondly, we recall that each Lyα forest spectrum covers the entire region between the QSO's Lyα and Lyβ emission lines. The redshift path length of this region is given by $\Delta z = 0.156(1 + z_{\mathrm{QSO}})$. Hence the number of independent pixels per spectrum also increases as $(1 + z_{\mathrm{QSO}})$. Since the S/N per pixel is kept constant this implies a larger number of photons per spectrum and hence an improved sensitivity to radial velocity shifts. The effect of this can be seen by comparing the green triangles in Fig. 18.6 with the orange crosses which are the result of using a constant redshift path length of $\Delta z = 0.4$ for each σ_v measurement, as well as $\gamma = 0$.

Finally, with $\gamma = 0$ and $\Delta z =$ const, the sensitivity to *wavelength* shifts should be constant as a function of z_{QSO}, and so the sensitivity to *velocity* shifts should go as $(1 + z_{\mathrm{QSO}})^{-1}$ (as can be seen from equation 18.9). In fact, the orange crosses in Fig. 18.6 decrease more slowly than this because the widths of the absorption lines in wavelength space increase as $(1 + z_{\mathrm{QSO}})$, making the edges of the lines less steep and hence slightly decreasing the spectrum's sensitivity to wavelength shifts.

In Fig. 18.7 we compare these results derived from the MC line lists to those from the real line lists. From Table 18.1 we can see that the real line lists (and hence the corresponding spectra) do not cover the full Lyα forest regions, with differently sized pieces missing both at the low and high redshift ends. Therefore we must correct the σ_v values derived from the real line lists in order to make them directly comparable to the values from the simulated lists. The correction is achieved by first assigning a new, slightly different QSO redshift to each spectrum, such that the 'missing' low and high redshift parts of the Lyα forest region are equally large. We then decrease the measured σ_v by a factor $(\Delta z_{\mathrm{obs}}/\Delta z)^{1/2}$, where Δz_{obs} is the redshift path length covered by the observed line list and Δz is the redshift path length of the full Lyα forest region at the new z_{QSO}. The correction factors range from 0.56 to 0.99.

The red stars in Fig. 18.7 show the corrected σ_v values derived from single pairs of spectra generated from the real absorption line lists with S/N $= 13\,000$ (run 5 in Table 18.3), while the blue dots show the measurements from the MC line lists (same as in Fig. 18.6). Overall the agreement between the results from the MC and real line lists is very good, particularly at high redshift. At $z_{\mathrm{QSO}} \approx 2.4$ the σ_v values from the real line lists are $\sim 15\,\%$ higher than those from the MC lists. By far the most significant deviation occurs at the lowest redshift where the σ_v of Q1101$-$264 is higher than expected by $47\,\%$. However, this

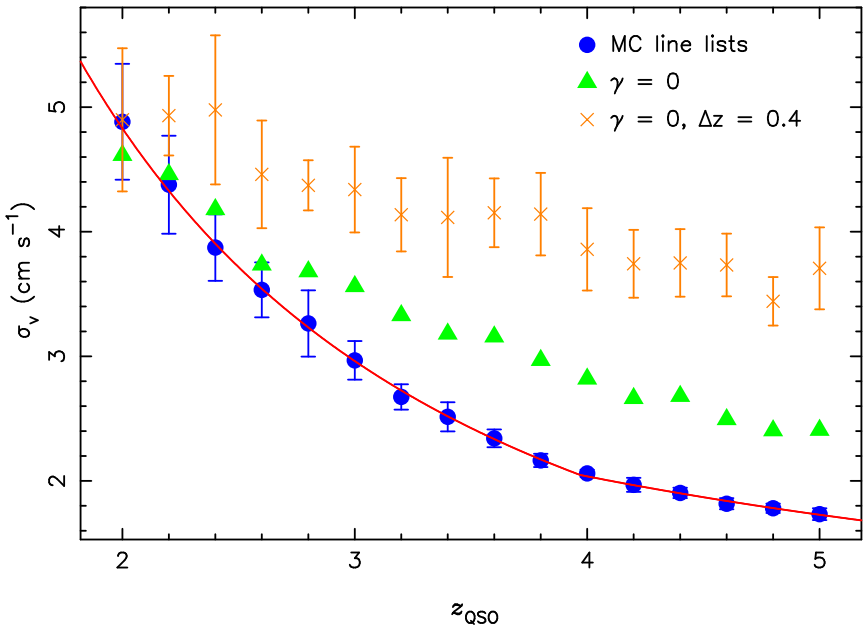

Figure 18.6: The blue dots with error bars show the accuracy with which a radial velocity shift can be determined from a Lyα forest spectrum as a function of QSO redshift. The red line parameterises the redshift dependence as in equation (18.11). Each point is the mean σ_v measured from 10 pairs of artificial spectra with S/N $= 13\,000$, generated from simulated MC absorption line lists. The error bars show the ± 1 r.m.s. range of the 10 simulations. The green triangles show the results for simulations where the redshift evolution of the Lyα forest has been switched off, i.e. where $\gamma = 0$ (cf. equation 18.3). The orange crosses show the result of also restricting the σ_v measurement in each spectrum to a redshift path of constant length $\Delta z = 0.4$, as opposed to using the full Lyα forest region between the QSO's Lyα and Lyβ emission lines.

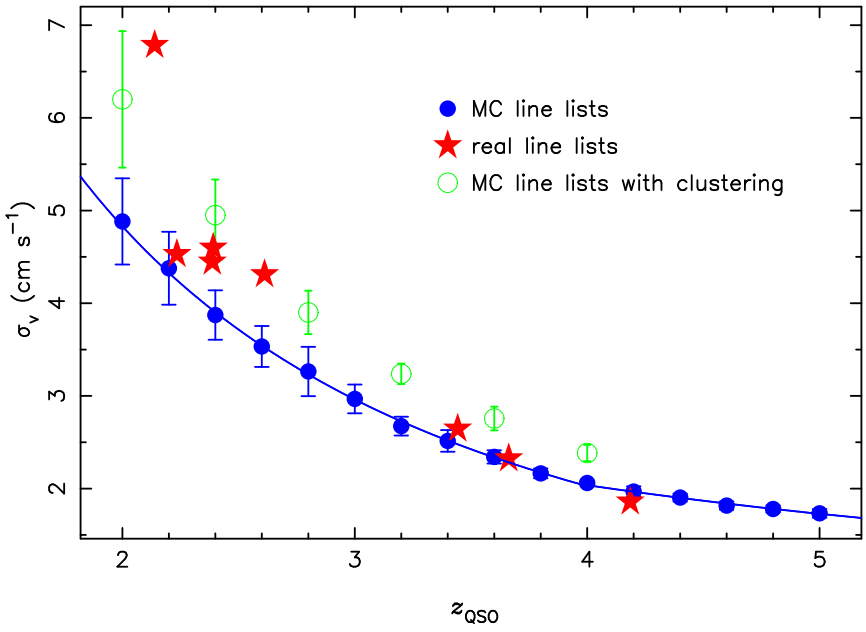

Figure 18.7: Comparison of σ_v measurements derived from simulated MC line lists (blue dots and solid line, same as in Fig. 18.6) and from real line lists (red stars). The green circles show the results derived from simulated MC line lists that include a simple scheme for clustering absorption lines in redshift space (see Section 18.4.4.1 for details).

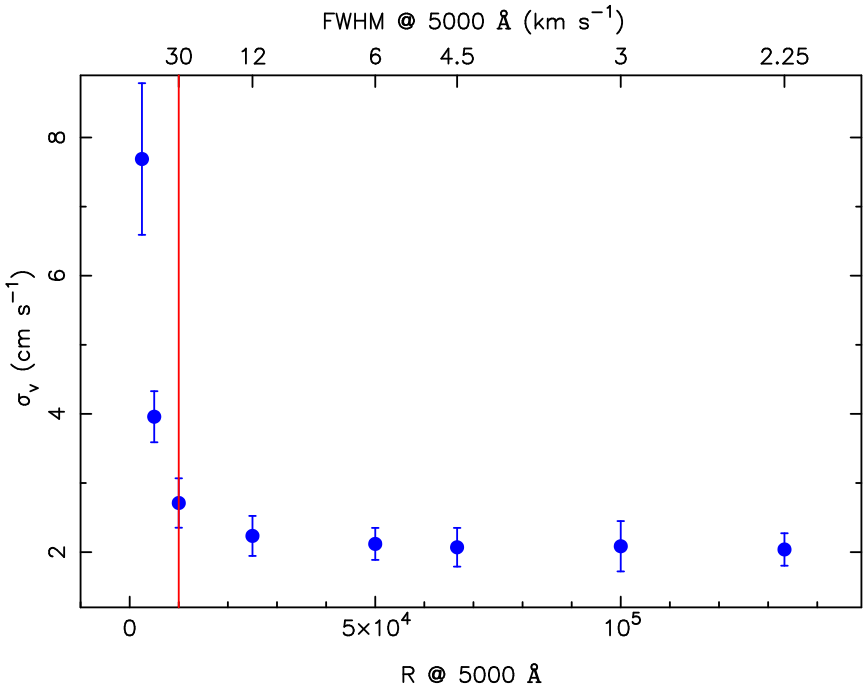

Figure 18.8: Radial velocity accuracy as a function of spectral resolution. Each point and error bar is the mean and ± 1 r.m.s. of 10 simulations at $z_{\rm QSO} = 4$ and S/N $= 13\,000$. The pixel size is kept constant and is chosen such that a resolution element is sampled by 3 pixels at the highest resolution considered. The labels along the bottom axis denote the resolving power at 5000 Å, while the labels along the top axis show the equivalent FWHM of an unresolved line. The vertical red line marks the mode of the absorption lines' b parameter distribution (cf. equation 18.3).

is not too surprising as the line of sight towards Q1101−264 is known to pass through an unusually low number of absorbers with $\log N_{\rm HI}({\rm cm}^{-2}) > 14$ (Kim et al. 2002).

We believe that the small differences at $z_{\rm QSO} \approx 2.4$ are mainly due to clustering of real absorption lines in redshift space. Clustering has the effect of reducing the number of spectral features because it increases line blanketing. However, at high redshift line blanketing is already severe because of the high line density and so clustering has a relatively smaller effect at high redshift than at low redshift. We demonstrate that clustering can explain the observed differences by using the toy clustering scheme described in Section 18.4.4.1. The σ_v values that result from simulations incorporating this scheme (run 6 in Table 18.3) are shown as open green circles in Fig. 18.7. The increase of σ_v compared to the unclustered simulations is clearly very similar to that observed for the real line lists and we conclude that clustering can indeed explain the small difference between the results obtained from the observed and simulated line lists at $z_{\rm QSO} \approx 2.4$. In any case, the near coincidence of the σ_v value of J2233−606 at $z_{\rm QSO} = 2.23$ with the expected value demonstrates that not all lines of sight are adversely affected by clustering. In the following we will assume that such sightlines can be pre-selected and hence we will ignore the effects of clustering in the rest of this section.

We now turn to the behaviour of σ_v as a function of spectral resolution. For various resolving powers in the range $2500 \leq R \leq 1.33 \times 10^5$ we have generated 10 pairs of line lists and spectra with $z_{\rm QSO} = 4$ and S/N $= 13\,000$ (run 7 in Table 18.3), and measured their σ_v values as before. The result is presented in Fig. 18.8, where we show the resolving power along the bottom axis and the corresponding FWHM of an unresolved line along the top axis. We stress that the pixel size was the same for all spectra ($= 0.0125$ Å) and that it was chosen such that a resolution element was well sampled even at the highest resolution considered. Hence the strong increase of σ_v towards lower R in Fig. 18.8 is not due to different numbers of photons per spectrum or sampling issues. Instead, it is simply due to the loss of information caused by convolving the spectrum with a line-spread function that is broader than the typical intrinsic absorption

line width (marked by the vertical line in Fig. 18.8). Indeed, at $R \gtrsim 30\,000$ the Lyα forest is fully resolved and in this regime σ_v is independent of R.

Summarising all of the above we find that the accuracy with which a radial velocity shift can be determined from the Lyα forest scales as:

$$\sigma_v = 2 \left(\frac{\text{S/N}}{2370}\right)^{-1} \left(\frac{N_{\rm QSO}}{30}\right)^{-\frac{1}{2}} \left(\frac{1+z_{\rm QSO}}{5}\right)^{-1.7} \text{cm s}^{-1}, \tag{18.12}$$

where the last exponent changes to -0.9 at $z_{\rm QSO} > 4$, and where the same S/N (per pixel) is assumed for all $N_{\rm QSO}$ spectra at both epochs.

18.5.2.3 Sensitivity gains from other spectral regions

Above we have investigated the sensitivity of the Lyα forest to radial velocity shifts, using only the H I Lyα transition in the region between a QSO's Lyα and Lyβ emission lines. However, modern echelle spectrographs are capable of covering a much wider spectral range in a single exposure and so the question arises whether other spectral regions, containing absorption lines from other ions or other H I transitions, can expediently contribute towards a $\dot{z}$ measurement.

The Lyβ forest: In Fig. 18.9 we compare the σ_v measurements derived from the extended Lyα + Lyβ MC line lists (blue dots, run 8 in Table 18.3) to the Lyα-only results (solid line, same as in Figs. 18.6 and 18.7). We find that including the Lyβ forest improves σ_v by a factor of 0.88 ± 0.006. The corresponding comparison for the real line lists (open and solid stars, run 9 in Table 18.3) yields improvement factors of 0.82 to 0.92, with a mean of 0.88, in very good agreement with the MC results.

Metal lines: Fig. 18.9 also shows the effect of the metal lines on σ_v. The green open squares show the result of only using the Lyα+Lyβ forest region as in the previous section but adding in all the available metal lines in this region (run 10 in Table 18.3. Comparing this to the H I-only results of the previous paragraph (open red stars) we find that the metal lines improve σ_v by factors of 0.58 to 0.99, with a mean of 0.85. We point out that, strictly speaking, the derived σ_v values are only upper limits for the six line lists taken from the literature because of their incomplete coverage of the Lyβ forest region. Indeed, the best improvement is achieved for one of the complete line lists, Q1101$-$264 (the lowest redshift QSO), which has a particularly rich metal absorption spectrum (cf. Table 18.2).

Finally, for our two complete line lists we show as solid green triangles the effect of also adding in the accessible spectral regions redwards of the Lyα forest. The additional metal lines further improve the σ_v values of Q1101$-$264 and Q0000$-$26 by factors of 0.55 and 0.90, respectively. With only two values it is obviously impossible to draw a firm conclusion regarding the average improvement offered by the metal lines redwards of the Lyα forest. We therefore choose to be conservative and adopt the larger of the two as a typical value.

Summarising the results of this section, the above experiments have shown that the normalisation of equation (18.12) can be reduced by a factor of $0.88 \times 0.85 \times 0.9 = 0.67$ by considering not just the Lyα forest but all available absorption lines, including metal lines, over the entire accessible optical wavelength range down to a QSO's Lyγ emission line. Hence we now obtain:

$$\sigma_v = 1.35 \left(\frac{\text{S/N}}{2370}\right)^{-1} \left(\frac{N_{\rm QSO}}{30}\right)^{-\frac{1}{2}} \left(\frac{1+z_{\rm QSO}}{5}\right)^{-1.7} \text{cm s}^{-1}. \tag{18.13}$$

18.5.2.4 Multiple epochs

Fundamentally, a redshift drift experiment consists of simply measuring the velocity shift between two spectra of the same QSO(s) taken at two distinct epochs separated by some time interval Δt_0. This is the view we have taken so far in the previous sections, where we determined the fundamental photon-noise limit of the accuracy with which this shift can be measured, and its scaling behaviour. However, in practice this notion is too simplistic. First of all, it implicitly assumes that the total integration time, $t_{\rm int}$,

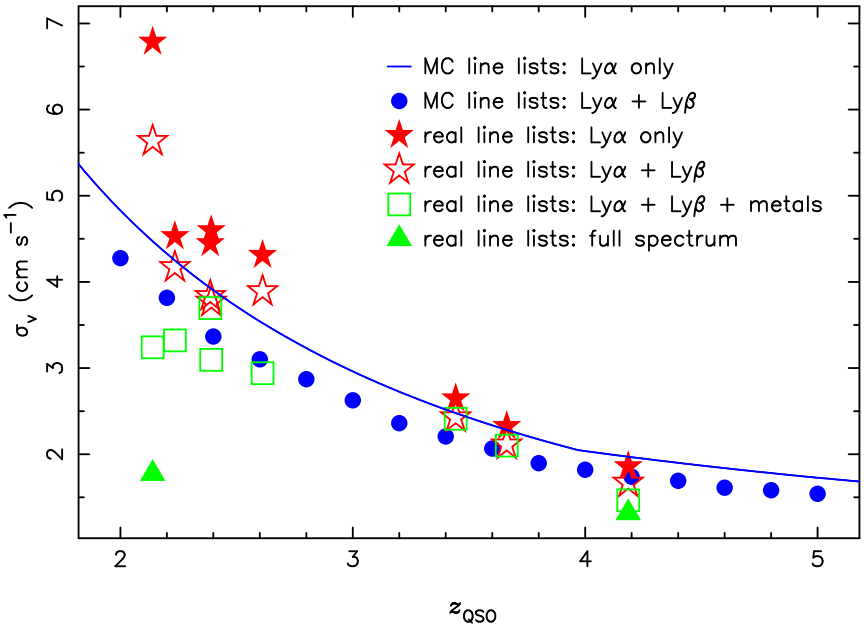

Figure 18.9: Comparison of σ_v measurements derived from Lyα-only line lists and various extended line lists covering additional spectral regions and including other absorption lines. The solid line and solid stars show, respectively, the results from the Lyα-only MC and real line lists as in Fig. 18.7. The blue dots and open stars show the corresponding improved σ_v values that result from the addition of the Lyβ forest. For the real line lists, the open squares show the effect of further adding the available metal lines in the Lyα+Lyβ forest region. Finally, for two of these QSOs, we show as solid triangles the outcome of using all lines accessible in existing spectra.

required to achieve the necessary S/N would be negligible compared to Δt_0, so that the two epochs of observation are well-defined. As we will see below this assumption is not valid. Secondly, for a variety of reasons it may be desirable to spread the observations more evenly over the whole period Δt_0 instead of concentrating them in just two more or less well-defined epochs at the interval's endpoints. Hence the question arises how the accuracy of a redshift drift experiment is affected by distributing the total available $t_{\rm int}$ over multiple observing epochs within the interval Δt_0.

Let us assume then that observations take place at N_e different epochs, where the jth epoch is separated from the first by Δt_j, so that $\Delta t_1 = 0$ and $\Delta t_{N_e} = \Delta t_0$. We can straightforwardly generalise the framework developed in Section 18.5.2.1 by turning equation (18.7) into a continuous equation for the expected normalised flux in the ith pixel at time Δt:

$$S_i(\Delta t) = S_{0i} + \frac{{\rm d}S_i}{{\rm d}\lambda}\,\lambda_i\,\frac{\dot{v}_i}{c}\,\Delta t \equiv S_{0i} + m_i \Delta t. \tag{18.14}$$

The idea is now to fit this linear relation to the observed fluxes S_{ji} at times Δt_j with errors σ_{ji}, yielding an estimate of the slope m_i and hence of $\dot{v}_i$ for each pixel. (Note that S_{0i} is a nuisance parameter. It represents the 'true' flux at the first epoch as opposed to the observed value S_{1i}.) The maximum likelihood estimator for m_i is

$$m_i = \frac{\overline{S_i \Delta t} - \overline{S_i}\;\overline{\Delta t}}{\overline{\Delta t^2} - \overline{\Delta t}^2}, \tag{18.15}$$

where the bar denotes the weighted average over all epochs:

$$\overline{x} = \frac{\sum_{j=1}^{N_e} x\;\sigma_{ji}^{-2}}{\sum_{j=1}^{N_e} \sigma_{ji}^{-2}}. \tag{18.16}$$

The variance of m_i is given by

$$\sigma^2_{m_i} = \left[\sum_j \sigma_{ji}^{-2} \left(\overline{\Delta t^2} - \overline{\Delta t}^2 \right) \right]^{-1} . \tag{18.17}$$

With m_i and its variance in place we can write down the equivalent of equation (18.9):

$$\sigma^2_{\dot{v}_i} = \left(\frac{c}{\lambda_i \frac{\mathrm{d}S_i}{\mathrm{d}\lambda}} \right)^2 \left[\sigma^2_{m_i} + \frac{m_i^2}{\left(\frac{\mathrm{d}S_i}{\mathrm{d}\lambda} \right)^2} \sigma^2_{S'_i} \right], \tag{18.18}$$

which in turn allows us to compute $\dot{v}$ averaged over all pixels and its error, $\sigma_{\dot{v}}$, corresponding to equations (18.8) and (18.10). Finally, we re-define $\sigma_v \equiv \sigma_{\dot{v}} \, \Delta t_0$. This new version of σ_v now includes the effect of multiple observing epochs and an arbitrary distribution of the total integration time among them. It is straightforward to show that for $N_e = 2$ one recovers exactly the original σ_v of Section 18.5.2.1.

We are now in a position to amend equations (18.12) and (18.13) to include the effect of multiple epochs. Recall that these scaling relations were derived for the case of $N_e = 2$ and for equal S/N in the spectra of both epochs. Since the variance of the normalised flux in pixel i scales as the inverse of the integration time we now write σ_{ji} as

$$\sigma^2_{ji} = \sigma^2_i \, \frac{0.5}{f_j}, \tag{18.19}$$

where f_j is the fraction of the total $t_{\rm int}$ used at the jth epoch ($\sum_j f_j = 1$), and σ_i denotes the flux error (in the ith pixel) that one would obtain if half of the total $t_{\rm int}$ were used. Further defining $\Delta t_j \equiv h_j \Delta t_0$, we can re-write equation (18.17) as:

$$\begin{aligned} \sigma^2_{m_i} &= \frac{2\sigma_i^2}{\Delta t_0^2} \left\{ 4 \left[\sum_j h_j^2 f_j - \left(\sum_j h_j f_j \right)^2 \right] \right\}^{-1} \\ &= \sigma^2_{m_i}(N_e = 2, f_1 = f_2 = 0.5) \; g^2(N_e, h_{1\ldots N_e}, f_{1\ldots N_e}). \end{aligned} \tag{18.20}$$

The first term above is just the variance of m_i that one obtains in the case of $N_e = 2$ and equal splitting of $t_{\rm int}$. The second term is a 'form factor' that only depends on the distribution of $t_{\rm int}$ within Δt_0. Again, it is straightforward to show that $g(N_e = 2, f_1 = f_2 = 0.5) = 1$. Since the form factor is the same for all pixels, and since $\sigma^2_{m_i}$ is the dominant term in equation (18.18), the sought-after modification of the σ_v scaling relation amounts to simply applying the form factor:

$$\sigma_v = 1.35 \; g(N_e, f_{1\ldots N_e}) \left(\frac{\mathrm{S/N}}{3350} \right)^{-1} \left(\frac{N_{\rm QSO}}{30} \right)^{-\frac{1}{2}} \left(\frac{1 + z_{\rm QSO}}{5} \right)^{-1.7} \text{cm s}^{-1}. \tag{18.21}$$

Note that the symbol 'S/N' now refers to the *total* S/N per object accumulated over all epochs, in contrast to equations (18.12) and (18.13) where it referred to the S/N achieved in each of two epochs.

Note also that we have dropped the dependence of the form factor on $h_{1\ldots N_e}$ by considering every night within the period Δt_0 as a potential epoch. This fixes N_e and $h_j = (j-1)/(N_e - 1)$, while f_j is constrained to lie in the range $0 \leq f_j \leq l/t_{\rm int}$, where l is the length of a night (which we will assume to be 9 h on average). Thus we find:

$$g(N_e, f_{1\ldots N_e}) = \frac{N_e - 1}{2} \left[\sum_{j=1}^{N_e - 1} j^2 f_j - \left(\sum_{j=1}^{N_e - 1} j f_j \right)^2 \right]^{-\frac{1}{2}} . \tag{18.22}$$

Clearly, f_j will be 0 for most nights. Nevertheless, there are obviously a large number of different possible distributions for the f_js. The best distributions are those that are symmetric and peaked towards the endpoints of Δt_0. A flat distribution, with equal observations taking place on n equally spaced nights,

results in $g = \sqrt{3\,(n-1)/(n+1)} \approx 1.7$ for $n \gg 1$. Thus the otherwise quite desirable arrangement of observing at a more or less constant rate throughout the period Δt_0 comes with a rather severe penalty attached. A priori, it is difficult to estimate the best g value that can be realistically achieved in practice. From now on we will assume, perhaps somewhat arbitrarily, that all observations occur as much as possible towards the beginning and end of Δt_0 with the constraint that the observing rate averaged over some intermediate time-scale of, say, a month cannot exceed $1/3$, i.e. that no more than a third of any month's telescope time is used for the redshift drift experiment. Depending on the ratio of $t_{\rm int}$ and Δt_0 this results in g values of $\sim$1.1. Essentially, this configuration simply shortens the effective length of the experiment by the amount of time it takes to complete the observations at either end of Δt_0.

18.5.3 Analysis II

18.5.3.1 Achievable radial velocity accuracy

In the previous sections we learnt what S/N ratio is required to achieve a given sensitivity to radial velocity shifts using QSO absorption spectra. The question now arises whether the S/N required for a meaningful $\dot{z}$ measurement can in fact be achieved (step 2 of the programme outlined in Section 18.4.1). In a photon-noise limited experiment the attainable S/N depends only on four quantities: the brightness of the source, the size of the telescope's collecting area, the total integration time and the total efficiency. By 'total efficiency' we mean the ratio of the number of detected photo-electrons to the number of source photons at the top of the atmosphere, i.e. it comprises atmospheric absorption and all losses occurring in the combined telescope/instrument system, including entrance aperture losses and the detector's quantum efficiency. We will now investigate in detail the 5-dimensional parameter space that is spanned by the above four quantities and redshift, in order to determine whether a feasible combination exists that would allow a meaningful $\dot{z}$ measurement.

In Fig. 18.10 we plot the QSO sample described in Section 18.4.4.1, split by sub-sets, in the $N_{\rm phot}$-$z_{\rm QSO}$ plane, where $N_{\rm phot}$ is a QSO's photon flux at the top of the atmosphere and at the centre of the QSO's Lyα forest, as implied by the appropriate magnitude described above. Using equations (18.21) and (18.4) and assuming values for D, ϵ and $t_{\rm int}$ we can calculate, for any given combination of $N_{\rm phot}$ and $z_{\rm QSO}$, the value of σ_v that would be achieved if *all* of the time $t_{\rm int}$ were invested into observing a single QSO with the given values of $N_{\rm phot}$ and $z_{\rm QSO}$. The background colour image and solid contours in Fig. 18.10 show the result of this calculation, where we have assumed $D = 42$ m, $\epsilon = 0.25$ and $t_{\rm int} = 2000$ h. Note that we have included both the improvement of σ_v afforded by the Lyβ forest and the metal lines as well as the deterioration caused by spreading $t_{\rm int}$ over a 0.9 yr period at either end of a $\Delta t_0 = 20$ yr interval.

From Fig. 18.10 we can see that, although challenging, a reasonable measurement of $\dot{z}(z)$ is within reach of a 42-m telescope. There exist a number of QSOs that are bright enough and/or lie at a high enough redshift to provide reasonable values of σ_v. We find 18 objects at $\sigma_v < 4$ cm s^{-1} and 5 objects at $\sigma_v < 3$ cm s^{-1}, with good coverage of the redshift range 2–4. One object even gives $\sigma_v = 1.8$ cm s^{-1}. However, for a smaller telescope with $D = 35$ m the number of objects with $\sigma_v < 4$ cm s^{-1} reduces to only 7 (cf. dotted contours).

In Fig. 18.11 we show more comprehensively how the number of QSOs with σ_v smaller than a given value depends on the telescope parameters and integration time, which we summarise into a single 'normalised observational setup parameter', O:

$$O = \left(\frac{D}{42\,\mathrm{m}}\right)^2 \frac{\epsilon}{0.25} \frac{t_{\rm int}}{2000\,\mathrm{h}} \tag{18.23}$$

(cf. equation 18.4). For a given value of O the colour image and contours in Fig. 18.11 show the number of QSOs that would give a σ_v equal to or smaller than the value along the ordinate if all of $t_{\rm int}$ was spent on any one of them. For example, if we wanted to be able to choose our targets from 30 QSOs bright enough and/or at a high enough redshift to be able to achieve $\sigma_v = 3$ cm s^{-1} or better on each object individually, then we would require $O \approx 2.1$. Note however, that the ordinate of Fig. 18.11 does *not* give the overall value of σ_v for a $\dot{z}$ experiment using the best $N_{\rm QSO}$ targets and setup O. The reason is of course that the total value of σ_v of such an experiment depends on how the total integration time is split up among the $N_{\rm QSO}$ targets (see below).

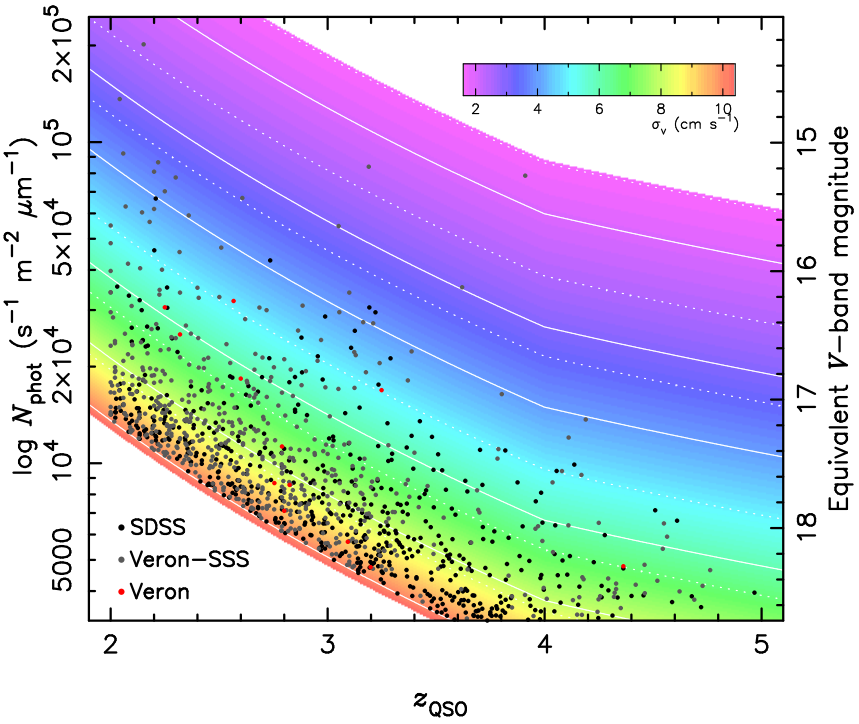

Figure 18.10: The dots show the known, bright, high-redshift QSO population (separated by sub-sets as indicated, see text) as a function of redshift and estimated photon flux at the centre of the Lyα forest. Along the right-hand vertical axis we have converted the photon flux to a corresponding Johnson V-band magnitude. The background colour image and solid contours show the value of σ_v that can be achieved for a given photon flux and redshift, assuming $D = 42$ m, $\epsilon = 0.25$ and $t_{\rm int} = 2000$ h. The contour levels are at $\sigma_v = 2, 3, 4, 6, 8$ and 10 cm s^{-1}. The dotted contours show the same as the solid ones, but for $D = 35$ m or, equivalently, for $\epsilon = 0.17$ or $t_{\rm int} = 1389$ h.

It is quite likely that a $\dot{z}$ experiment would target not just one, but several QSOs. This would be desirable for several practical reasons, including the ability to observe for as large a fraction of the total available telescope time as possible (to ease scheduling), and to be able to identify any potential directional systematics caused by the atmosphere, the telescope, the instrument, the calibration procedures or the transformation of redshifts to the cosmological reference frame. Obviously though, the more objects are included into the experiment the worse the final result will be (in the absence of systematics) because observing time will have to be redistributed from the 'best' object to the less suited ones.

So how does one select the 'best' targets for a $\dot{z}$ experiment from our list of QSOs, and, given a target list, what is the value of the total, combined σ_v for this list of objects and a given setup O?

The selection strategy of targets for a $\dot{z}$ experiment depends on the precise goals of that experiment. However, the most straightforward thing to do is to select targets by the value of their achievable σ_v, which can be read off Fig. 18.10 for each object. For a fixed number of QSOs, $N_{\rm QSO}$, and a fixed setup O this selection will result in the best possible total, overall value of σ_v, which we compute as:

$$\sigma_v^{\rm tot} = \left[\sum_i^{N_{\rm QSO}} \sigma_v^{-2}(z_{{\rm QSO},i}, m_i, f_{{\rm obj},i}\, O) \right]^{-\frac{1}{2}}, \qquad (18.24)$$

where $f_{{\rm obj},i}$ is the fraction of the total integration time allocated to the ith object. In Fig. 18.12 we plot $\sigma_v^{\rm tot}$ as a function of $N_{\rm QSO}$ and O, where we have chosen to distribute the time equally among all objects (i.e. $f_{{\rm obj},i} = N_{\rm QSO}^{-1}$). Distributing the time to give equal S/N instead (which may be operationally desirable) leads to slightly better values of $\sigma_v^{\rm tot}$ when only a few QSOs are targeted. The reason is of course that the object with the best σ_v is only the seventh brightest target (cf. Fig. 18.10) and so it is allocated relatively

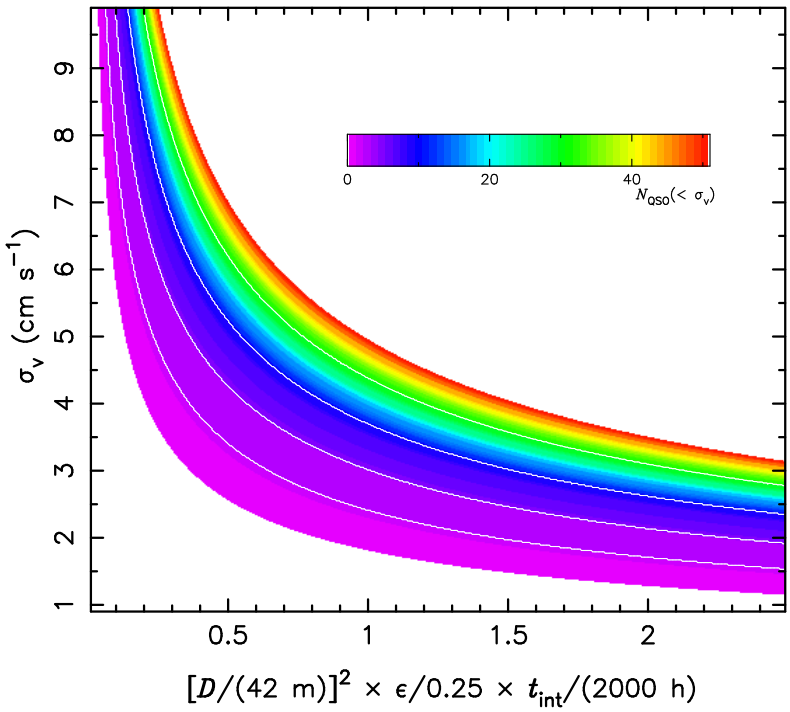

Figure 18.11: The colour image and the contours show the number of QSOs for which the σ_v value on the ordinate or better can be achieved for a given combination of telescope size, efficiency and integration time. The contour levels are at $N_{\rm QSO} = 3, 5, 10$ and 30.

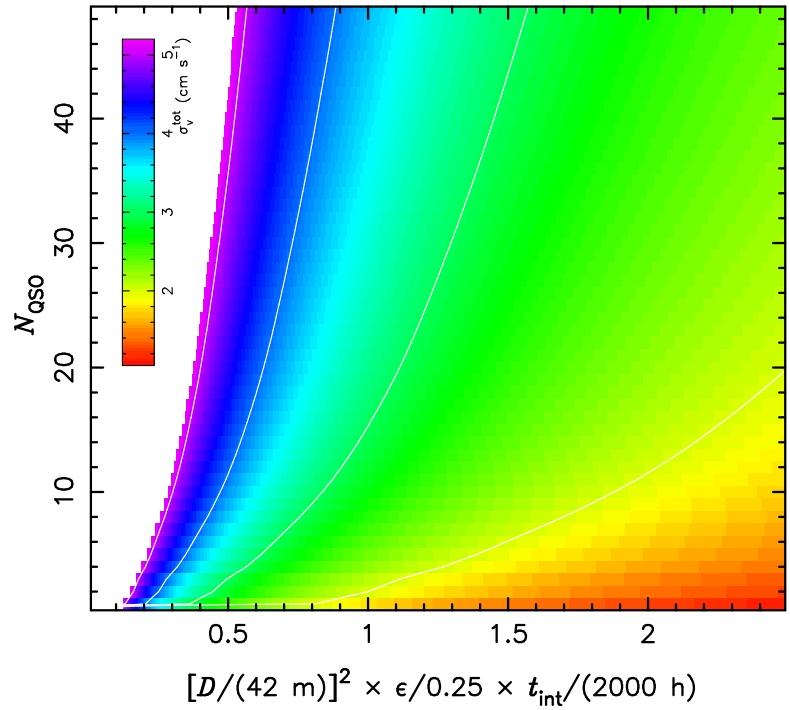

Figure 18.12: The colour image and the contours show the final, overall value of σ_v achieved by targeting the $N_{\rm QSO}$ objects with the best individual σ_v values (cf. Fig. 18.10) and by employing a given combination of telescope size, efficiency and total integration time. The total integration time is split equally among the targets. The contour levels are at $\sigma_v^{\rm tot} = 2, 3, 4$ and 5 cm s^{-1}.

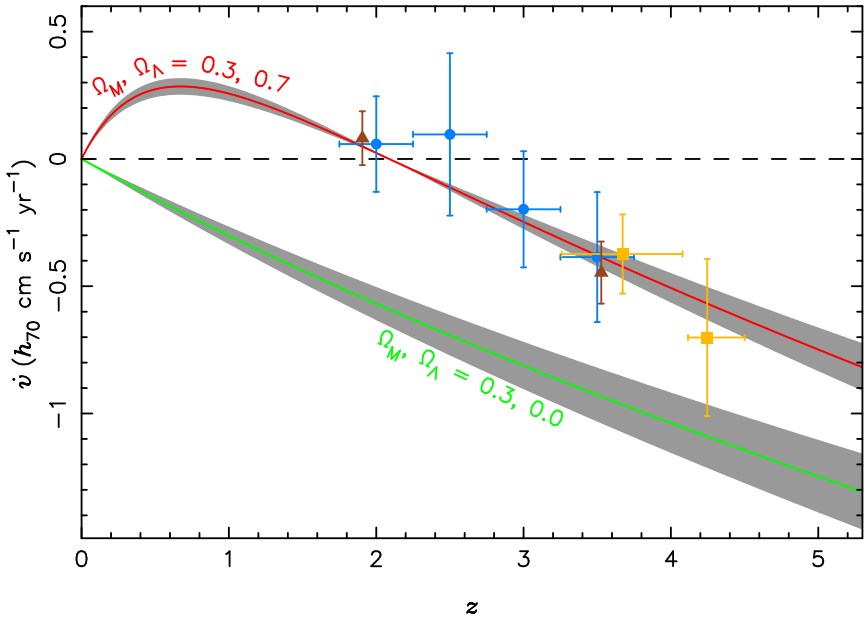

Figure 18.13: The three sets of 'data' points show MC simulations of the redshift drift experiment using the three different example implementations discussed in the text. In each case we have assumed an observational setup of $O = 2$ and we plot as 'data' points the 'observed' values and error bars of the velocity drift $\dot{v}$, expected for a total experiment duration of $\Delta t_0 = 20$ yr and for standard cosmological parameters ($h_{70} = 1$, $\Omega_{\rm M} = 0.3$, $\Omega_\Lambda = 0.7$). For a given QSO we use the centre of the Lyα forest as the redshift of the $\dot{v}$ measurement. Blue dots: selection by σ_v, $N_{\rm QSO} = 20$ (binned into four redshift bins), equal time allocation. Yellow squares: selection by $|\dot{v}|/\sigma_{\dot{v}}$, $N_{\rm QSO} = 10$ (in two redshift bins), equal time allocation. Brown triangles: selection by best combined constraint on Ω_Λ, $N_{\rm QSO} = 2$, optimal time distribution. The solid lines show the expected redshift drift for different parameters as indicated, and $h_{70} = 1$. The grey shaded areas result from varying H_0 by ± 8 km s^{-1} Mpc^{-1}.

more time. However, for $N_{\rm QSO} \gtrsim 15$ this advantage is lost and equal time allocation produces $\sigma_v^{\rm tot}$ values that are smaller by $\sim$10 % compared to equal S/N allocation.

From Fig. 18.12 we can see that an overall value of $\sigma_v^{\rm tot} \approx 2$–$3$ cm s^{-1} is well within reach of the E-ELT, even when tens of objects are targeted for the experiment. The steepness of the contours also indicate that $\sigma_v^{\rm tot}$ does not depend very sensitively on $N_{\rm QSO}$, at least for $N_{\rm QSO} \gtrsim 10$.

For further illustration we now perform a MC simulation of a redshift drift experiment using this target selection strategy. We (arbitrarily) choose $N_{\rm QSO} = 20$ and $O = 2$ so that the overall accuracy of the experiment is $\sigma_v^{\rm tot} = 2.34$ cm s^{-1} (cf. Fig. 18.12). Furthermore we assume that the observations span a time interval of $\Delta t_0 = 20$ yr and that the true $\dot{z}(z)$ is given by our standard cosmological model. The result is presented in Fig. 18.13 where we show as blue dots one realisation of the 'observed' velocity drifts and their errors along with the input model (red solid line). Since the selected 20 targets cover the redshift range $2.04 \leq z_{\rm QSO} \leq 3.91$ quite homogeneously we have binned the measurements into four equally sized redshift bins.

By construction these points represent the most precise measurement of $\dot{z}$ that is possible with a set of 20 QSOs (using equal time allocation) and $O = 2$. However, since many of the selected QSOs lie near the redshift where $\dot{z} = 0$ (for the assumed model) the redshift drift is only detected with an overall significance of $\mathcal{S} = \overline{|\dot{v}|}/\sigma_{\dot{v}}^{\rm tot} = 1.4$, where $\overline{|\dot{v}|}$ is the weighted mean of the absolute values of the expected velocity drifts.

This can be improved upon by choosing a different approach and selecting targets by the largest value of $|\dot{v}|/\sigma_{\dot{v}}$. This quantity is a strongly increasing function of redshift and this selection strategy results in quite a different set of objects: the best 20 targets according to this criterion include only 3 of the objects previously selected. For $O = 2$, $\Delta t_0 = 20$ yr and equal time allocation as before, the combined overall

significance of this sample is $\mathcal{S} = 2.7$ (assuming our standard model of $\dot{z}$ is correct). The yellow squares in Fig. 18.13 show the result of a MC simulation using this setup and implementation, except that we use $N_{\rm QSO} = 10$ (in two redshift bins) which gives a slightly better significance of $\mathcal{S} = 3.1$. To reach $\mathcal{S} \geq 4$ we need to further reduce $N_{\rm QSO}$ to 3, or instead increase O to 3.4 or Δt_0 to 25.4 yr.

18.5.3.2 Constraints on cosmological parameters

Finally, we turn to the question of how to best select targets to constrain the acceleration of the expansion and what can be achieved in this respect with our sample of known QSOs. As in the previous case the answer will depend on what to expect for the expansion history and in particular for the acceleration. For the purpose of the following discussion we will again assume our standard cosmological model.

Ideally, we would like to detect the existence of the acceleration by measuring $\dot{z} > 0$ with the highest possible significance. This implies (i) that we need a $\dot{v}$ measurement at $z < z_0$, where z_0 is defined by $\dot{v}(z_0) = 0$, and (ii) that target selection should proceed by the largest value of $\dot{v}/\sigma_v$, which indeed favours the lowest available redshifts. However, even if we use only the single best object by this criterion and assume a generous $\Delta t_0 = 25$ yr then a $2\sigma_v$ detection of $\dot{v} > 0$ would still require an unfeasible $O = 9.5$. The reason for not being able to do better is of course our inability to access the Lyα forest at $z \approx 0.7$ where $\dot{v}$ is the largest.

Under these circumstances the best thing to do is to combine a $\dot{z}$ measurement at the lowest possible redshift (and with the smallest possible error bar) with a second measurement at the highest possible redshift, thereby gaining the best possible constraint on the slope of $\dot{z}(z)$. A significant detection of a non-zero, negative slope at $z \gtrsim 2$ combined with a $\dot{z} = 0$ measurement at $z \approx 2$ leads to the inevitable conclusion that $\dot{z}$ must be positive at $z \lesssim 2$, thereby proving the existence of past acceleration and constraining Ω_Λ. Indeed, it can be shown (Liske et al. 2008) that in the parameter space of our cosmological model this strategy corresponds to obtaining two lines of constraint in the Ω_Λ-$\Omega_{\rm M}$ plane that make as large an angle as possible and therefore provides the strongest constraint on Ω_Λ.

However, when trying to turn this approach into an actual procedure for selecting the best targets for the low and high-z measurements, respectively, one realises that the selection of the two groups cannot in fact be separated, making an analytical procedure impractical. Hence we resort to the simplest case of only one object per group and simply try all possible combinations of two objects in our sample. For any given combination we also determine empirically the optimal split of the total integration time among the two objects. In this way we find that the best lower limit on Ω_Λ is achieved for the two very bright objects in Fig. 18.10 at $z_{\rm QSO} = 2.15$ and 3.91, with the latter object receiving 0.3 of the total integration time. Mock $\dot{v}$ measurements for these objects are shown as brown triangles in Fig. 18.13.

In Fig. 18.14 we present the expected constraints from these two objects in the Ω_Λ-$\Omega_{\rm M}$ plane, assuming $O = 2$ and $\Delta t_0 = 20$ yr. The red and blue solid lines show the individual constraint lines for a fixed $h_{70} = 1$ and the grey shaded areas around them are the error bands that correspond to the respective values of σ_v for each object. Note that the optimal time split (in the sense of giving the best lower limit on Ω_Λ) is such that the two bands are of roughly equal width.

The coloured ellipses in Fig. 18.14 show the joint 68 and 90 % confidence regions that result from the combination of the two $\dot{v}$ measurements. Here we have assumed that an external constraint places 1σ limits of ± 8 km s^{-1} Mpc^{-1} on H_0 (Freedman et al. 2001) before marginalising over this parameter. By comparing the error bars on the brown triangles in Fig. 18.13 with the grey shaded region around the red line we can see that this uncertainty in H_0 is small relative to the errors on $\dot{v}$. Hence the ellipses in Fig. 18.14 are almost those that would be obtained for a fixed H_0.

Projecting the joint constraints onto the vertical axis we find that the 95 % lower limit on the cosmological constant is $\Omega_\Lambda > 0.16$ (indicated by the hashed region in Fig. 18.14), while $\Omega_\Lambda = 0$ is excluded at the 98.2 % confidence level. Note that we can also exclude $q_0 = 0$ at a similar level of 97.6 %. These numbers do not depend sensitively on the adopted prior on H_0. Doubling the error on H_0 to ± 16 km s^{-1} Mpc^{-1} has almost no effect at all. Even if we use a flat prior and simply marginalise over the range $0 \leq H_0 \leq 140$ km s^{-1} Mpc^{-1} we can still reject $\Omega_\Lambda = 0$ at 98.1 % confidence, although the 95 % lower limit now

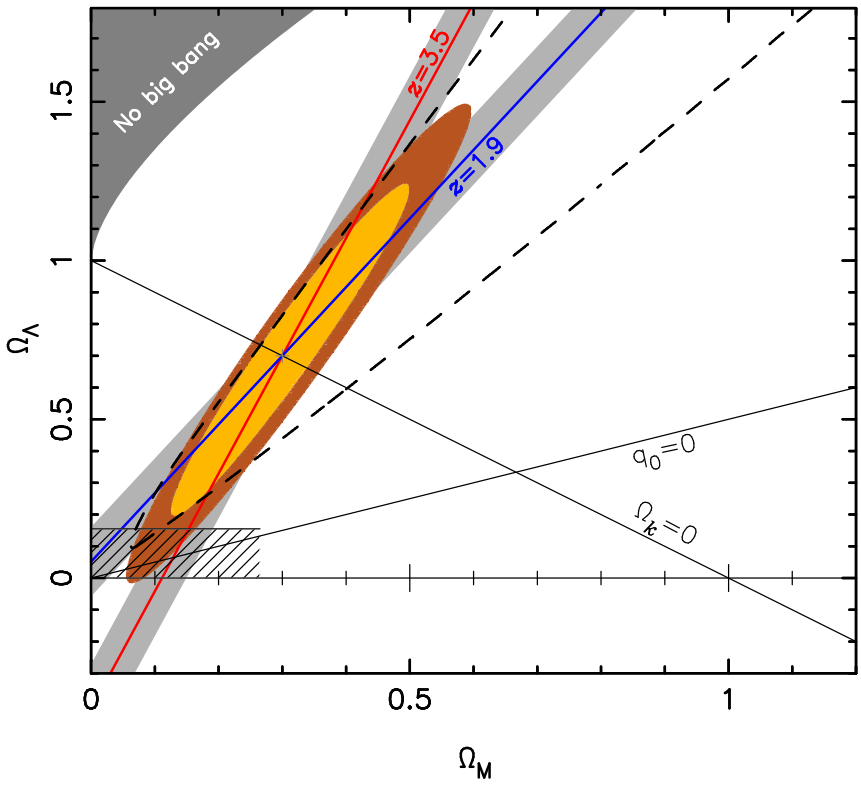

Figure 18.14: Expected constraints in the Ω_Λ-Ω_M plane from two measurements of the redshift drift at two different redshifts as indicated, assuming $O = 2$ and $\Delta t_0 = 20$ yr. The two targets and the split of the observing time among them were chosen to give the best lower limit on Ω_Λ. The red and blue solid lines and the grey shaded bands show the individual constraints provided by each of the two objects assuming a fixed $h_{70} = 1$. The coloured ellipses show the joint 68 and $90\,\%$ confidence regions that result from combining the two measurements, marginalising over H_0 using an external prior of $H_0 = (70 \pm 8)$ km s^{-1} Mpc^{-1}. The hashed region indicates the $95\,\%$ lower limit on Ω_Λ. The dashed contour shows the $68\,\%$ confidence region that results from using a flat prior on H_0 and just marginalising over the range $0 \leq H_0 \leq 140$ km s^{-1} Mpc^{-1}. Flat cosmologies and the boundary between current de- and acceleration are marked by solid black lines. The dark shaded region in the upper left corner designates the regime of 'bouncing universe' cosmologies which have no big bang in the past.

reduces to 0.08. For a shorter experiment duration of $\Delta t_0 = 15$ yr (and the flat H_0 prior) we can still exclude $\Omega_\Lambda = 0$ at $93.5\,\%$ confidence.

Combining the constraints of a $\dot{z}$ experiment with those from other cosmological observations would of course lead to significantly stronger results. A detailed examination of this issue is, however, beyond the scope of this work. We will let it suffice to make two general points: (i) the $\dot{z}$ constraints will be unique and therefore new in the sense that they cannot be provided by any other cosmological observation. $\dot{z}$ probes $H(z)$ in a redshift range that is inaccessible to other methods (such as SNIa, weak lensing, BAO) and this will remain true even in the ELT era. In terms of characterising $H(z)$, $\dot{z}$ measurements at $z \gtrsim 2$ will nicely complement the data at $z \lesssim 2$ provided by SNIa surveys. In the Ω_Λ-Ω_M plane these datasets offer similar but nevertheless distinct constraints. (ii) The datasets that are most complementary to the redshift drift in the Ω_Λ-Ω_M plane are those that constrain the geometry of the Universe, such as the fluctuation power spectrum of the CMB, which can break the remaining degeneracy in the $\dot{z}$ constraints. In fact, the WMAP 3-year data constrain the quantity $\Omega_\Lambda + 0.72\,\Omega_\mathrm{M}$ (Spergel et al. 2007), so that the CMB degeneracy line is almost exactly orthogonal to the $\dot{z}$ constraints shown in Fig. 18.14. Combining these constraints would lead to individual 2σ errors in Ω_Λ and Ω_M of ~ 0.08 and ~ 0.06, respectively.

Finally, let us consider models in which the universe is *exactly* flat ($\Omega_k = 0$). The parameter space of these models is 2-dimensional and in Fig. 18.15 we show the constraints in the Ω_Λ-H_0 plane offered by the two

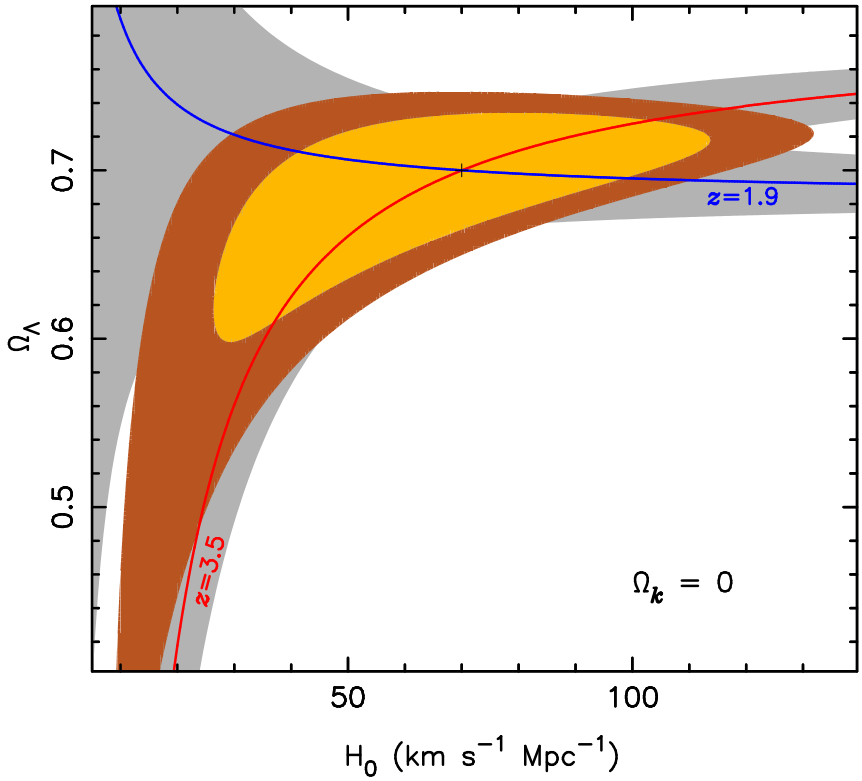

Figure 18.15: Expected constraints in the Ω_Λ-H_0 plane for flat cosmologies assuming $O = 2$ and $\Delta t_0 = 20$ yr and using the same two redshift drift measurements as in Fig. 18.14. The red and blue solid lines and the grey shaded bands show the individual constraints provided by each of the two objects. The coloured ellipses show the joint 68 and 90 % confidence regions that result from combining the two measurements.

$\dot{v}$ measurements considered above ($O = 2$, $\Delta t_0 = 20$ yr). Clearly, H_0 is not constrained in any useful way and both substantially higher and lower redshift measurements of similar or better quality would be required to do so. On the other hand, Ω_Λ is constrained quite well. Marginalising over H_0 we find a 2σ range of $[0.42, 0.74]$, and $\Omega_\Lambda = 0$ is excluded at 98.7 % confidence. For a shorter experiment duration of $\Delta t_0 = 15$ yr we can still exclude $\Omega_\Lambda = 0$ at 94.7 % confidence. Note that these latter values are not very much higher than in the general case above. The difference is of course that in the flat case the best fit along the $\Omega_\Lambda = 0$ axis is obtained for an unfeasibly low H_0 value of $h_{70} = 0.1$, while in the general case this value is 1. Hence, for flat cosmologies any external H_0 prior has a strong impact on the lower limit one can place on Ω_Λ, in contrast to the general case. For example, applying the relatively weak prior that H_0 is known to within ± 16 km s^{-1} Mpc^{-1} results in the rejection of $\Omega_\Lambda = 0$ at > 99.99 % confidence for any $\Delta t_0 > 9$ yr or, alternatively, for $\Delta t_0 = 15$ yr and any $O > 0.5$.

18.5.4 Compliance with figures of merit

Fig. 18.12 succinctly summarises the results of this work in terms of the metric defined in Section 18.3. It plots the figure of merit as a function of the main parameters of a $\dot{z}$ experiment. The goal formulated in Section 18.3 was to reach $\sigma_v^{\rm tot} = 3$ cm s^{-1} in an observing time of 2500 h using at least 20 QSOS. From Fig. 18.12 we read off a value of $\sigma_v^{\rm tot} = 2.86$ cm s^{-1} for $N_{\rm QSO} = 20$ and $O = 1.25$. However, this figure was calculated assuming a total experiment duration of $\Delta t_0 = 20$ yr. Recall that σ_v depends weakly on Δt_0 because of the presence of the form factor in equation (18.21). Using the required $\Delta t_0 = 15$ yr we find $\sigma_v^{\rm tot} = 2.92$ cm s^{-1}. Thus we conclude that that the goals of this DRM science case can indeed be met.

18.5.5 Sensitivity to input parameters

Fig. 18.12 shows the quantitative dependence of the figure of merit on the main 'input' parameters. Clearly, the photon collecting power of the combined telescope/instrument system, $D^2\epsilon$, is the most crucial parameter for a $\dot{z}$ experiment. Assuming a fixed telescope diameter and recalling the discussion of Section 18.4.4.2, we conclude that the main parameter that can and must be optimised is the throughput of the spectrograph (including entrance losses).

Recall that our simulations and subsequent analysis did not account for any wavelength limits imposed by the spectrograph. However, from Figs. 18.13 and 18.14 it is clear that the lower wavelength limit is quite important, as it determines the lowest redshift to which $\dot{z}$ can be probed. This is important because a $\dot{z}$ measurement near $z \approx 2$ is needed to unambiguously prove the existence of accelerated expansion (assuming that the predictions of our current cosmological model are correct and $\dot{z}$ really does change sign at $z \approx 2$). Full coverage of the Lyα forest at $z = 2$ would require a lower wavelength limit of 3340 Å. Conversely, a lower limit of 3700 Å (as is currently being discussed for CODEX) would exclude Lyβ forest observations at $z < 2.6$; would start to impinge on the Lyα forest for $z < 2.3$; but would allow at least some Lyα forest observations down to $z = 2$.

We point out, however, that the precise sensitivity of a $\dot{z}$ experiment to the lower wavelength limit of the spectrograph depends on the target sample. For the exact target sample used in this work a lower limit of 3700 Å would be very detrimental as it would all but exclude the very bright QSO at $z_{\rm QSO} = 2.15$ in Fig. 18.10, which is the third best object in the whole sample and pivotal to our derived constraints on Ω_Λ. However, the addition of just a single, similarly bright, but slightly higher redshift QSO to the sample would again remove this dramatic effect.

18.5.6 Calibration requirements

The quest for photon-noise limited radial velocity shift measurements places rather exacting requirements on the spectrograph to be used for a redshift drift experiment, in particular on the wavelength calibration: (i) the error on a radial velocity shift measurement induced by the random error on the wavelength calibration must not exceed that induced by photon noise for any single observation. What this precisely means depends on the details of the signal extraction method but essentially it means that the error bar on a position measurement (of a single feature? of an entire spectrum?) at a given epoch (a single observation?) must be dominated by photon noise. (ii) There must be no unrecognised (and hence uncorrected) long-term drift in the wavelength calibration exceeding the final, overall radial velocity precision of the experiment, i.e. any unrecognised long-term drift must be $\ll 3$ cm s^{-1}.

Depending on the method used to extract the redshift drift signal from the data, it is likely that additional requirements, e.g. on the accuracy of the flat-fielding and the stray light correction, may arise in the future.

18.5.7 Limitations

The most obvious limitation of this work is that all results were derived under the assumption that the uncertainty on radial velocity shift measurements will be dominated by photon noise. The next step would be to understand to what extent this assumption can be realised in practice. To this end one would need to (i) draw up a list of all potential sources of error, both astrophysical and instrumental in nature, (ii) quantitatively assess their impact on a $\dot{z}$ measurement, and (iii) where necessary and possible identify appropriate hardware solutions, calibration and/or analysis techniques that will reduce their impact below the level of the photon noise. Although some of the most obvious astrophysical sources of noise have already been addressed elsewhere (Liske et al. 2008) a realistic and comprehensive assessment of the total error budget is a non-trivial task and beyond the scope of this work.

The main reason why the above is not possible at the moment is that we have not yet identified the 'best' method to extract the redshift drift signal from the data. There are several reasons why the $\dot{z}$ measurement method described in Section 18.5.2.1 (whose accuracy is measured by σ_v) is sub-optimal. At the end of that section we had already pointed out that this method does not capture the information that is contained

in the $\dot{z}$-induced change of the shape of a spectrum. In the case of a Lyα-only spectrum this change of shape consisted of a compression of the spectrum. However, if we include additional transitions with different rest wavelengths in the analysis then we no longer have a one-to-one correspondence between absorption redshift and wavelength, and so the redshift drift will in general induce a much more complex change of the shape of a spectrum. Hence it is clear that any attempt to harness this additional information must involve the complete identification and modelling of all absorption features used in the analysis.

Another problem is the fact that QSO continua are to some extent variable. Our current method of extracting the redshift drift signal is essentially a flux-difference method and it is clearly very sensitive to any errors in determining the change of the continuum. Similarly, it is also quite susceptible to any residual uncertainties in the sky subtraction, flat-fielding and stray light correction. Clearly, a flux-difference type of approach would not be the method of choice for a real-life $\dot{z}$ experiment. However, other methods of extracting the redshift drift signal are likely to be less sensitive to these issues. As suggested above we believe that the most promising approach is to simultaneously model the absorption in all spectra of a given object with Voigt profiles, including the $\dot{z}$ of each component (or a model thereof) as an additional free parameter. The drawback of this method is that a complex absorption feature can usually be modelled in more than one way, none of which necessarily has to correspond to reality, and this could represent an additional source of error for a $\dot{z}$ measurement. The issue of an ideal $\dot{z}$ extraction method remains to be investigated.

In deriving our results we have also used the magnitudes and redshifts of the high-redshift QSO population as it is known today. First, we point out the significant uncertainties regarding the photometry of this sample: we had to rely on photographic photometry from the SuperCOSMOS Sky Survey for most of the brightest objects in Fig. 18.10. Secondly, note that we have used *all* known QSOs, irrespective of their declination. In exchange, so to speak, we have refrained from speculating on future discoveries of bright, high-redshift QSOs. It is clear that existing QSO catalogues do not yet cover the entire sky uniformly with the same completeness, even at the bright magnitudes we are interested in here. Indeed, only one third of the objects shown in Fig. 18.10 lie in the south. Hence it is not unreasonable to expect that a sample similar to the one used here will present itself to the E-ELT no matter where it will be located, once uniform all-sky surveys have discovered all exisiting bright QSOs. If anything we believe that we have erred on the side of caution in using our sample: since colour selection is one of the most popular methods of QSO candidate identification it is likely that our current sample is incomplete in the redshift range $2.4 \lesssim z \lesssim 3$ where the colours of QSOs are very similar to those of main sequence stars (e.g. Richards et al. 2002). Considering that a $\dot{z}$ experiment would benefit greatly from *any* new discoveries (see also Section 18.5.5) we feel that our results are more likely to be conservative than optimistic. In any case, uncertainties about the exact target sample available for a future $\dot{z}$ experiment are clearly a limitation of the current work.

Finally, we point out that we have neglected to include any wavelength dependence of the efficiency ϵ in our simulations.

18.6 Concluding remarks

In this work we have appraised the prospects of using the E-ELT to detect and characterise the cosmological redshift drift, $\dot{z}$, by means of a $\sim$two decade-long spectroscopic monitoring campaign of high redshift QSO absorption lines. We summarise our results as follows:

1. Under the assumption of photon-noise limited observations the accuracy, σ_v, to which a putative radial velocity shift between two spectra of the same object can be determined depends chiefly on the number and sharpness of the relevant spectral features and the spectra's S/N. The number density of the absorption lines that constitute the Lyα forest depends strongly on redshift. Using extensive Monte Carlo simulations we have hence derived a quantitative relation between the σ_v of a pair of Lyα forest spectra on the one hand, and the spectra's S/N and the background QSO's redshift on the other hand (equation 18.12).

2. Apart from the Lyα forest, all QSO spectra also display higher order H I lines and a variety of metal absorption lines. Using the H I Lyβ lines in addition to the Lyα lines, as well as all available metal

lines within and to the red of the Lyα forest significantly improves a given spectrum's sensitivity to radial velocity shifts by a factor of 0.67 (equation 18.13).

3. In practice it will not be possible to measure the redshift drift by simply comparing just two observations of some set of objects obtained at two well-defined epochs separated by some Δt_0. In reality, the total observing time required for the experiment will be spread over many nights (= epochs) whose distribution within Δt_0 (now defined as the interval between the first and last observations) will be subject to many practical constraints. In Section 18.5.2.4 we have calculated the relevant 'form factor' by which σ_v increases (or, equivalently, by which the effective Δt_0 decreases) as a result of a given observing time distribution. For realistic distributions we find factors of ~ 1.1 to 1.7.

4. We use a comprehensive list of high redshift QSOs *already known today* together with our final σ_v scaling relation to predict the E-ELT's capability to measure $\dot{z}$. For example, selecting those 20 targets that deliver the greatest measurement accuracy we find that a 42-m E-ELT could achieve a total, overall accuracy of 2.9 cm s^{-1} in 2500 h of observing time (Fig. 18.12). This would allow a 3σ distinction between $(\Omega_{\rm M}, \Omega_\Lambda) = (0.3, 0.7)$ and $(0.3, 0.0)$ in 15 yr.

   The E-ELT will be able to provide unequivocal and direct proof of the existence of past acceleration by obtaining an upper limit on $\dot{z}$ at high redshift and a lower limit at $z \approx 2$ where the transition from $\dot{z} < 0$ to > 0 is expected to occur (cf. Fig. 18.13). For a total observing time of 4000 h and an experiment duration of 20 yr these limits will be strong enough to prove that $\dot{z}$ must become positive below $z \approx 2$. The validity of this proof only requires that gravity can be described by a metric theory and that the Universe is homogeneous and isotropic on large scales. It is entirely independent of any specific theory of gravity (such as general relativity) and does not require any assumptions regarding the spatial geometry of the Universe. It is also independent of other cosmological observations and does not rely on any other astrophysical assumptions whatsoever. Hence, a $\dot{z}$ experiment carried out with the E-ELT will arguably provide the most direct evidence of acceleration possible.

References

Bouchy F., Pepe F., Queloz D., 2001, A&A, 374, 733
Cristiani S., D'Odorico S., Fontana A., Giallongo E., Savaglio S., 1995, MNRAS, 273, 1016
Fernández-Soto A., Lanzetta K.M., Barcons X., Carswell R.F., Webb J.K., Yahil A., 1996, ApJ, 460, L85
Francis P.J., 1993, ApJ, 407, 519
Freedman W.L., et al., 2001, ApJ, 553, 47
Hambly N.C., et al., 2001, MNRAS, 326, 1279
Hu E.M., Kim T.-S., Cowie L.L., Songaila A., Rauch M., 1995, AJ, 110, 1526
Kim T.-S., Carswell R.F., Cristiani S., D'Odorico S., Giallongo E., 2002, MNRAS, 335, 555
Kim T.-S., Cristiani S., D'Odorico S., 2001, A&A, 373, 757
Kim T.-S., Hu E.M., Cowie L.L., Songaila A., 1997, AJ, 114, 1
Kirkman D., Tytler D., 1997, ApJ, 484, 672
Liske J., et al., 2008, MNRAS, 386, 1192
Liske J., Webb J.K., Williger G.M., Fernández-Soto A., Carswell R.F., 2000, MNRAS, 311, 657
Loeb A., 1998, ApJ, 499, L111
Lu L., Sargent W.L.W., Womble D.S., Takada-Hidai M., 1996, ApJ, 472, 509
Oke J.B., 1974, ApJS, 27, 21
Richards G.T., et al., 2002, AJ, 123, 2945
Sandage A., 1962, ApJ, 136, 319
Saslaw W.C., 1989, ApJ, 341, 588
Schneider D.P., et al., 2007, AJ, 134, 102
Spergel D.N., et al., 2007, ApJS, 170, 377
Véron-Cetty M.-P., Véron P., 2006, A&A, 455, 773

19 Conclusions

19.1 Summary of the simulation results

In this section we summarise the results from each of the DRM cases described in the previous sections. The summary is structured along the lines of the three primary questions that were to be answered for each science case by the DRM (see Section 4): What can be achieved and how much observing time does this require? Which properties of the E-ELT system are the limiting factors? What are the key requirements?

For the benefit of the discussion below we now define the following symbols: D = telescope diameter, $t_{\rm exp}$ = exposure time, ε = total system throughput (including atmosphere, telescope, instrument and detector; i.e. ε denotes the ratio of detected electrons to photons impinging at the top of the atmosphere; for the sake of brevity we will ignore the fact that, in general, ε is different for object and background photons [e.g. thermal emission from the telescope mirrors does not suffer atmospheric extinction]), B = background surface brightness (see [RD1] for a detailed discussion of the various components contributing to the background), $\mathrm{EE}_{\rm DLC}$ = fraction of energy contained in the diffraction-limited core of the PSF. Note that in the discussion below we will generally consider B a property of the E-ELT system because, depending on the wavelength, B may depend on a number of system parameters such as the site, the telescope's mirror coatings, etc. The parameter combination $D^2\,\varepsilon$ will also be referred to as 'photon collecting power'.

19.1.1 S3-1: Direct imaging of terrestrial and giant exoplanets

This DRM case investigated which types of planets the E-ELT will be able to detect in reflected light using high-contrast imaging techniques.

What can be achieved? For a G-type star at a distance of 10 pc, $t_{\rm exp}$ = 10 h, and an apodised pupil Lyot coronagraph we expect to be able to reach a $5\,\sigma$ contrast with respect to the peak of the PSF of 10^{-9} at 0.5″ separation from the star (see Fig. 5.6), enabling the detection of a Jupiter-like planet. Under the same assumptions, the detection of an Earth-like planet, on the other hand, is not feasible: at a separation of 0.1″ we expect to be able to achieve a contrast of only $\sim 2\times10^{-8}$ instead of the required contrast of $\sim 6\times10^{-10}$. However, at least in principle the detection of an Earth-like planet is possible with the E-ELT: for a 'perfect' coronagraph without any chromaticity and with instrumental post-AO aberrations of 10 nm rms the expected contrast decreases to just about the required level (see the left panel of Fig. 5.14).

Limiting factor: For this DRM case there is not just one, but three nearly equally important parameters that limit the performance of the E-ELT system: coronagraph chromaticity, instrumental post-AO aberrations and resolution (i.e. D).

Key requirements: $D \geq$ 42 m, XAO-fed integral field spectrograph operating in the wavelength range 0.9–1.7 μm equipped with a coronagraph with as little chromaticity as possible, instrumental post-AO aberrations $\leq$ 10 nm.

19.1.2 S3-2: Earth twins in the habitable zone of solar-type stars

This DRM case pursued the question of whether the E-ELT will be able to uncover Earth-like planets by detecting and measuring the reflex motion of the host stars using the radial velocity method.

What can be achieved? The radial velocity 'noise' generated by stellar activity, granulation and oscillations can be largely overcome by adopting a suitable observing strategy and by averaging observations over suitable timescales. The best strategy is to obtain three $t_{\rm exp}$ = 10 min observations (separated by 2 h) of a given target star every three nights. For an inactive K1V star this strategy brings down the stellar noise to a level enabling the detection of a $\sim$2 $M_\oplus$ planet orbiting at the inner edge of its host star's habitable zone (see Fig. 6.10). A radial velocity survey operating under this strategy would thus require

5 h of observing time per target star per month. Pushing the detection limit down to $\sim$1 $M_{\oplus}$ would require tripling the number of observing epochs, i.e. a given target star would have to be observed three times *every* night.

Limiting factor: This case is limited by the achievable S/N for point source spectroscopy. The noise is in turn dominated by object photon noise. In addition, it is assumed that the spectrograph will be operated in seeing-limited mode. Hence the results are only limited by the E-ELT's photon collecting power (i.e. $D^2\,\varepsilon$). Note, however, the unusual way in which this limitation operates: for a fixed total amount of observing time, increasing the photon collecting power would allow one to decrease the exposure time of individual observations (while still remaining in the regime where the radial velocity error is dominated by the intrinsic stellar noise), thus freeing up time which could be invested into additional observing epochs, thereby decreasing the detection limit. An additional (non E-ELT related) limitation arises from the requirement to include only relatively quiet stars in the target sample.

Key requirements: $D \geq$ 42 m, $\varepsilon \gtrsim$ 0.2, high-resolution optical spectrograph with $R >$ 100 000 and with long-term radial velocity stability of $<$ 2 cm s^{-1}.

19.1.3 S9: Imaging the planet-forming regions of circumstellar disks

This DRM case examined the E-ELT's capability of studying the process of planet formation by investigating the spatial structure of young, low-mass circumstellar disks via direct near and mid-infrared imaging.

What can be achieved? A young, low-mass disk surrounding a T Tauri star can easily be detected out to distances of at least 1 kpc in as little as $t_{\rm exp}$ = 10 s, even without the use of a coronagraph. Detecting planetary formation features, on the other hand, is much more challenging. For example, a 2 AU-wide gap at a radius of 5 AU in a circumstellar disk at a distance of 400 pc corresponds to a $\sim$1.5 pixel-wide flux reduction of a few per cent at a distance of less than 4 pixel from the PSF centre. Nevertheless, the detection of such gaps at radii of 5–10 AU in disks at the distances of the nearest star-forming regions ($\sim$140 pc) appear feasible. The application of a coronagraph would presumably provide significantly better performance.

Limiting factor: The E-ELT's ability to detect planetary formation features is limited by the telescope's resolution (i.e. D) and the accuracy with which its PSF can be calibrated.

Key requirements: $D \geq$ 42 m, diffraction-limited imager operating in the K and L bands, ideally equipped with a coronagraph, PSF calibration accurate to a few per cent.

19.1.4 S5-1: Characterising the lowest mass freely floating objects in star forming regions

The aim of this DRM case was to understand how well the E-ELT will be able to spectroscopically characterise the low-mass end of the sub-stellar initial mass function in the nearest star forming regions.

What can be achieved? At the distances of the nearest star forming regions it will be possible to distinguish the young, low-mass objects that this case seeks to identify from much more evolved and more massive T dwarfs of the same effective temperature down to $T_{\rm eff} \approx$ 1000 K, despite the fact that the exposure times will be limited to $<$ 2 h by the rotation period these young objects. For an assumed age of 5 Myr this limit in $T_{\rm eff}$ corresponds to a mass of 3 $M_{\rm Jup}$. Note, however, that the proposal for this case stipulated that this programme would be executed in a survey-like manner in non-optimal conditions. Hence, the simulations were carried out under the assumption of an integral field spectrograph with 40 mas spaxel size and GLAO correction. Since this case does not require a large field of view a diffraction-limited spectrograph could be used instead, which would deliver significantly better results. In fact, since spatially resolved spectroscopy or high multiplexity are not required either, a possibly more efficient slit spectrograph could be used.

Limiting factor: This case is limited by the achievable S/N for point source spectroscopy. The noise is in turn dominated by the background photon noise. Hence, in principle this case is limited by the parameter combination $D^4\,\varepsilon\,\mathrm{EE}_{\mathrm{DLC}}^2\,B^{-1}$. However, for the observational set-up used in the simulations the results are limited by $D^2\,\varepsilon\,\mathrm{EE}_{\mathrm{sp}}^2\,B^{-1}$, where $\mathrm{EE}_{\mathrm{sp}}$ is the fraction of energy ensquared by one of the instrument's spaxels, and where we have assumed that, to first order, the spaxel size of a GLAO-fed instrument would not vary with D. Note that this case includes observations in the wavelength range 0.8–0.9 μm, where ε is strongly affected by the E-ELT's mirror coatings; as well as observations in the mid-infrared, where B is strongly affected by the coatings and the site.

Key requirements: $D \gtrsim$ 42 m, GLAO-fed intermediate resolution spectrograph(s) with $R \approx 4500$ operating in the wavelength range 0.6–4 μm with as much simultaneous wavelength coverage as possible.

19.1.5 S5-2: The centres of massive dense young clusters: deep E-ELT infrared imaging and 3D spectroscopy

Showing how the E-ELT will help to understand the formation and early dynamical evolution of massive stars embedded in dense, highly obscured Galactic proto-cluster clouds was the topic of this DRM case.

What can be achieved? At the distance of the Galactic centre, and using $t_{\mathrm{exp}} \approx 1$ h, up to 200 mag of visual extinction can be penetrated in the mid-infrared L and M bands in order to probe the high-mass end of the initial mass function. If K-band photometry is also required then the extinction limit decreases to 150 mag while t_{exp} has to be increased to $\sim$25 h. Similarly, intermediate-resolution K-band spectroscopy will be possible for stars down to 8 $M_{\odot}$ behind 100 mag of extinction.

Limiting factor: This case is limited by the S/N achievable for point source photometry and spectroscopy. The noise is in turn dominated by the background photon noise. Hence the results are limited by the parameter combination $D^4\,\varepsilon\,\mathrm{EE}_{\mathrm{DLC}}^2\,B^{-1}$. Note that this case includes observations in the mid-infrared, where B is strongly affected by the E-ELT's coatings and site.

Key requirements: $D \geq$ 42 m, diffraction-limited imager(s) operating in the wavelength range 2–5 μm, LTAO-fed integral field spectrograph with $R \approx 10\,000$ and field of view size $\gtrsim$ 2.5 arcsec^2, operating in the wavelength range 2.1–2.2 μm.

19.1.6 S5-3: Giant planet-mass objects in the Large Magellanic Cloud

This DRM case appraised the E-ELT's ability of constraining the low-mass end of the sub-stellar initial mass function, and possibly determining its lower mass limit, in the low-metallicity environment of the LMC through near-infrared imaging.

What can be achieved? In t_{exp} = 20–30 h the E-ELT will be capable of reaching down to magnitudes in the J, H and K bands that correspond to 10–20 M_{Jup} at the distance of the LMC. However, the exact mass limit depends on the age of the examined star forming region.

Limiting factor: This case is limited by the S/N achievable for point source photometry, given the presence of other stars in the field. For most stars the noise is dominated by the background photon noise, but some have additional noise due to the close proximity of a much brighter star. Hence, for the most part the results are limited by the parameter combination $D^4\,\varepsilon\,\mathrm{EE}_{\mathrm{DLC}}^2\,B^{-1}$. Note that the factor $\mathrm{EE}_{\mathrm{DLC}}$ is particularly affected by the fact that the LMC will only be observable at high zenith distance. A (non E-ELT related) limitation may arise from a spatially varying background nebulosity in the star forming region under study. If the variation length scale is smaller than a resolution element, then this could easily become the dominating limitation for this DRM case.

Key requirements: $D \geq$ 42 m, diffraction-limited imager operating in the wavelength range 1–2.5 μm.

19.1.7 G4-1: The resolved stellar populations of elliptical galaxies

This DRM case was concerned with the E-ELT's ability to perform photometry in extremely crowded stellar fields and to construct accurate colour-magnitude diagrams for elliptical galaxies out to the distance of the Virgo cluster.

What can be achieved? Accurate colour-magnitude diagrams will be best constructed in the I and H bands. For NGC 205 in the Local Group it will be possible to probe the main sequence turn-off over almost the entire extent of the galaxy with a photometric accuracy of no less than 0.05 mag (see Fig. 11.8). For Cen A it will be possible to probe the horizontal branch with an accuracy of 0.05 mag all the way in to the central parts of the galaxy, down to at least 0.3 effective radii (see Fig. 11.7). For M87 in the Virgo cluster the E-ELT will be able to probe the tip of the red giant branch with 0.05 mag accuracy all the way in to the very dense central parts of the galaxy, down to $\sim$0.5 effective radii (see Fig. 11.6). Achieving these results will require exposure times of a few to several tens of hours per observation.

Limiting factor: This case is limited by the S/N achievable for point source photometry. In the crowded central regions of galaxies the noise is completely dominated by not knowing precisely how many stars are 'hiding' below a given star. Thus, in this regime the results are entirely limited by the E-ELT's resolution alone, which, in turn, only depends on D for $\lambda \gtrsim 1$ μm, but also depends on the effectiveness of the AO correction for shorter wavelengths, because in this regime the AO correction is not effective enough to provide a PSF with a well-developed diffraction limited core. In the less crowded outskirts of galaxies the accuracy of the photometry is background limited, i.e. the results are limited by the parameter combination $D^4\,\varepsilon\,\mathrm{EE}^2_{\mathrm{DLC}}\,B^{-1}$ for $\lambda \gtrsim 1$ μm. In the I-band, i.e. in the intermediate regime between quasi-seeing-limited and diffraction-limited PSFs, the limiting parameter combination is $D^{2(1+\alpha)}\,\varepsilon\,\mathrm{EE}^2_{\mathrm{ref}}\,B^{-1}$, where α is in the range 0–1, and $\mathrm{EE}_{\mathrm{ref}}$ is the fraction of energy encircled by the optimal S/N reference area. Note also that in the I-band ε is strongly affected by the E-ELT's mirror coatings. An additional limitation may arise from PSF variations as a function of position within the field of view, time, star colour, etc. Depending on the accuracy of the PSF calibration this could easily become the dominant limitation.

Key requirements: $D \geq 42$ m, diffraction-limited imager(s) operating in the wavelength range 0.8–2.5 μm, PSF calibration accurate to a few per cent.

19.1.8 G4-2: The chemo-dynamical structure of galaxies

This DRM case considered the prospect of obtaining accurate metallicity distributions and kinematics from intermediate resolution spectroscopy of individual red giant branch (RGB) stars in galaxies as distant as the Virgo cluster.

What can be achieved? At the distance of NGC 205 in the Local Group it will be possible to derive metallicities accurate to $\pm$0.25 dex and velocities accurate to $\pm$3 km s^{-1} from observations of the Calcium triplet in stars 1 mag below the tip of the RGB in just 20 min of exposure time. This will even be possible at the very centre of the galaxy. For Cen A accurate velocities can be derived for stars 0.5 mag below the tip of the RGB in $t_{\mathrm{exp}} \approx 2$ h, even at just 1 effective radius from the galaxy's centre. Precise metallicities, on the other hand, will require significantly more time ($t_{\mathrm{exp}} \gtrsim 5$ h) and will be limited to the very tip of the RGB and to distances larger than 2 effective radii from the centre. Finally, at the distance of M87 in the Virgo cluster, it will be extremely challenging to obtain spectra of individual RGB stars of sufficient S/N to measure velocities and metallicities with reasonable accuracy, even when employing exposure times as large as 50 h. However, note that the simulations were carried out under the assumption of an integral field spectrograph with 50 mas spaxel size. The reason for this was the need for a relatively large field of view. Nevertheless, a smaller spaxel size, more suited to the size of the PSF should deliver somewhat better results.

Limiting factor: This case is limited by the achievable S/N for point source spectroscopy. Under the best circumstances the noise is in turn dominated by the background photon noise (i.e. in the outer regions of a galaxy where crowding can be neglected). Hence, in principle this case is limited by the

parameter combination $D^4\,\varepsilon\,\mathrm{EE}_{\mathrm{DLC}}^2\,B^{-1}$. However, the observations are carried out in the wavelength range 0.8–0.9 μm where the AO correction is not effective enough to provide a PSF with a truly diffraction limited core. In this intermediate regime between quasi-seeing-limited and diffraction-limited performance the limiting parameter combination is $D^{2(1+\alpha)}\,\varepsilon\,\mathrm{EE}_{\mathrm{sp}}^2\,B^{-1}$, where α is in the range 0–1, and $\mathrm{EE}_{\mathrm{sp}}$ is the fraction of energy ensquared by one of the instrument's spaxels. Note also that in this wavelength range ε is strongly affected by the E-ELT's mirror coatings. Finally, given the need to observe large numbers of stars, this case will also be limited by the instrument's field of view.

Key requirements: $D \geq 42$ m, integral field spectrograph with $R \approx 6000$ and field of view size $\gtrsim$ 100 arcsec^2, operating in the wavelength range 0.8–0.9 μm. The requirement on the AO is to deliver the largest possible fraction of energy ensquared by a spaxel (which is, strictly speaking, not the same thing as delivering diffraction-limited performance).

19.1.9 G4-3: First stars relics in the Milky Way and satellites

This DRM case addressed the question to what extent the E-ELT will be able to help us understand the nature and nucleosynthesis of the earliest chemical enrichment processes by obtaining detailed abundance patterns of extremely metal-poor stars in the Milky Way and other Local Group galaxies by way of high-resolution spectroscopy.

What can be achieved? The E-ELT will be able to provide a comprehensive survey of extremely metal-poor candidate stars in the halo of the Milky Way and in nearby dwarf galaxies like Sagittarius, LMC, Sculptor and Fornax. Covering 100 stars in the Milky Way as well as 100 stars in these four dwarf galaxies would require a total of $\sim$530 h of observing time.

Limiting factor: This case is limited by the achievable S/N for point source spectroscopy. The noise is in turn dominated by object photon noise. In addition, it is assumed that the spectrograph will be operated in seeing-limited mode. Hence the results are only limited by the E-ELT's photon collecting power (i.e. $D^2\,\varepsilon$). In the ultraviolet regime ε is strongly affected by the E-ELT's site and its mirror coatings.

Key requirements: $D \geq 42$ m, high-resolution optical spectrograph with $R > 70\,000$, operating down to at least 0.37 μm, although 0.31 μm is required for some aspects of this case.

19.1.10 G9: A survey of black holes in different environments

The goal of this DRM case was to quantify to what extent the E-ELT will be able to shed light on both the low and high-mass ends of the M_{BH}-σ relation through spatially resolved spectroscopy of the central regions of galaxies.

What can be achieved? The E-ELT will be able to probe the M_{BH}-σ relation over the entire mass range of supermassive black holes out to the distance of the Virgo cluster, including the low-mass end at $\sim 10^6\ M_\odot$. At larger distances the lower mass limit that can be probed increases, but the highest mass black holes of $\sim 10^{10}\ M_\odot$ will remain detectable out to a distance of $z \approx 0.3$.

Limiting factor: This case is limited by the achievable S/N for extended object spectroscopy over an area whose size is set by the sphere of influence of the black hole being studied (i.e. the size of the area is independent of D). The noise is in turn dominated by background photon noise. Hence the results are limited by the parameter combination $D^2\,\varepsilon\,B^{-1}$. An additional limitation is imposed by the resolution (i.e. D) and $\mathrm{EE}_{\mathrm{DLC}}$ because these parameters control to what extent light from stars at different positions (and hence different velocities) is mixed, thereby smearing out the effect of the black hole and making it harder to detect. Note that this case includes observations in the wavelength range 0.8–0.9 μm where ε is strongly affected by the E-ELT's mirror coatings.

Key requirements: $D \geq 42$ m, diffraction-limited integral field spectrograph(s) with $R \approx 5000$ operating in the wavelength range 0.8–2.5 μm.

19.1.11 C10-1: The physics and mass assembly of galaxies out to z∼6

This DRM case explored the E-ELT's capability of elucidating the evolution of galaxies by surveying large numbers of high-redshift galaxies with spatially resolved spectroscopy.

What can be achieved? We expect to be able to detect the [O II] emission line from galaxies of characteristic stellar mass (M_*) up to a redshift of ∼4.8 in $t_{\rm exp}$ = 24 h. In fact, at z = 4 it will be possible to recover the large scale motions of galaxies and to distinguish between rotating disks and major mergers down to a limiting stellar mass of 0.5 M_* (see Fig. 15.13). At the same redshift, even the detailed kinematics of rotating disks and their internal structure will be available down to M_*. The determination of full rotation curves of M_* galaxies, however, will be limited to z = 2. A comprehensive survey of 800 galaxies over a range of redshifts (2–5.6) and stellar mass (0.5–5 M_*) would require ∼720 h of observing time.

Limiting factor: This case is limited by the achievable S/N per spaxel for spatially resolved, extended object spectroscopy. The size of a spaxel is set by the scale of high-redshift galaxies (i.e. it is independent of D). The noise is in turn dominated by background photon noise. Hence the results are limited by the parameter combination $D^2\,\varepsilon\,B^{-1}$. However, there is also a slight dependence on $\mathrm{EE_{sp}}$ because high-redshift galaxies can be quite clumpy and/or quite small (depending on redshift and stellar mass), meaning they may be sampled by only a few spaxel.

Key requirements: $D \geq$ 42 m, MOAO-fed multi object integral field spectrograph with $R \approx$ 5000, individual IFU field of view $\gtrsim$ 4 arcsec^2, spaxel scale ∼50 mas, multiplexity $\gtrsim$ 100, patrol field of view size $\approx$ 25 arcmin^2, operating in the wavelength range 1.3–2.5 μm. Note that the spatial resolution is set by the spaxel scale. The requirement on the AO is not to degrade this resolution any further (hence excluding GLAO) and to deliver the largest possible fraction of energy ensquared by a spaxel.

19.1.12 C10-2: High resolution imaging of high redshift galaxies

This DRM case examined the role of the E-ELT in studying the morphologies of high-redshift galaxies.

What can be achieved? For galaxies of characteristic stellar mass the E-ELT will be able to provide useful diffraction limited images up to a redshift of 2.5 (see e.g. Fig. 16.11). These images will be of high enough resolution to easily resolve individual clumps frequently seen in high-redshift galaxies, and, for $t_{\rm exp}$ = 10 h, of high enough S/N to allow quantitative morphological analyses. At higher redshifts it will be necessary to give up diffraction limited resolution and to employ larger pixel sizes in order to obtain sufficient S/N. Nevertheless, this will allow the E-ELT to probe the galaxy stellar mass function down to at least the characteristic stellar mass up to a redshift of ∼7.5. Although JWST will undoubtedly be an important competitor in this regime, given its superior sensitivity, the E-ELT will still have the advantage of being able to supply a somewhat higher resolution.

Limiting factor: This case is limited by the achievable S/N per pixel for extended object imaging. The size of a pixel is set by the scale of high-redshift galaxies (i.e. it is independent of D). The noise is in turn dominated by background photon noise. Hence the results are limited by the parameter combination $D^2\,\varepsilon\,B^{-1}$. However, there is also a slight dependence on $\mathrm{EE_{pix}}$ because high-redshift galaxies can be quite clumpy and/or quite small (depending on redshift and stellar mass), meaning they may be sampled by only a few pixel.

Key requirements: $D \geq$ 42 m, MCAO-fed diffraction limited imager with field of view size $\approx$ 4 arcmin^2, operating in the J, H and K bands for $z < 2.5$ galaxies. For higher redshift galaxies an imager with pixel size $\approx$ 30 mas is required. The requirement on the AO in this case is to deliver the largest possible fraction of energy ensquared by a pixel.

19.1.13 C4: First light – the highest redshift galaxies

The goal of this DRM case was to test the E-ELT's ability to push spectroscopic studies of galaxies beyond a redshift of 7.

What can be achieved? The E-ELT will be able to measure redshifts of $z > 7$ galaxies with $H_{\rm AB}$ = 26 and 28.5 mag with an accuracy of 0.01 in $t_{\rm exp} \approx$ 1 h and 100 h, respectively, depending on the quality of the AO correction. Deriving physical parameters for an $H_{\rm AB}$ = 27 mag galaxy by measuring absorption line column densities will require $t_{\rm exp} \gtrsim$ 50 h.

Limiting factor: This case is limited by the achievable S/N for spatially integrated extended object spectroscopy. The noise is in turn dominated by the background photon noise. Hence one would expect the results to be limited by the parameter combination $D^2\,\varepsilon\,B^{-1}$. However, due to the clumpy nature of high-redshift galaxies, and the small sizes of these clumps, the targets are actually more like clusters of point sources than extended (i.e. roughly constant surface brightness) objects. Hence this case is in fact limited by the parameter combination $D^4\,\varepsilon\,{\rm EE}_{\rm sp}^2\,B^{-1}$ (assuming that the instrument's spaxel size would scale with D). In addition, since we are working between the sky lines, the results may be limited by the strength of any broad wings of the spectrograph's line spread function. Finally, a survey of high-redshift galaxies will also be limited by the size of the multi-object spectrograph's patrol field of view, and, depending on the surface density of the targets, possibly by its multiplexity.

Key requirements: $D \geq$ 42 m, multi-object integral field spectrograph with $R \approx$ 4000, operating in the wavelength range 1–1.4 μm, with multiplexity similar to the area covered by the patrol field of view, expressed in arcmin^2. The requirement on the AO is to deliver the largest possible fraction of energy ensquared by a spaxel.

19.1.14 C2: A dynamical measurement of the expansion history of the Universe

This DRM case explored the possibility that the E-ELT might be able to directly observe the evolving expansion rate of the Universe by detecting and measuring the so-called redshift drift through a long-term high-resolution spectroscopic monitoring campaign of QSO absorption lines.

What can be achieved? Using a target sample of 20 known QSOs in the redshift range 2–5 the E-ELT will be able to achieve a total, overall radial velocity accuracy of 2.9 cm s^{-1} in 2500 h of observing time, allowing a clear detection of the redshift drift and a 3σ distinction between an open and flat Universe over a timescale of 15 yr. Using a total of 4000 h of observing time over a period of 20 yr would provide a direct, purely dynamical and definitive proof of the existence of a period of accelerated expansion in the past (see Figs. 18.13 and 18.14).

Limiting factor: This case is limited by the achievable S/N for point source spectroscopy. The noise is in turn dominated by object photon noise. In addition, it is assumed that the spectrograph will be operated in seeing-limited mode. Hence the results are limited by the E-ELT's photon collecting power (i.e. $D^2\,\varepsilon$). Furthermore, this case is limited by the lifetime of the E-ELT. An additional (non E-ELT related) limitation comes from the fact that more suitable targets than currently known probably exist but remain to be discovered.

Key requirements: $D \geq$ 42 m, $\varepsilon \geq$ 0.25, high-resolution optical spectrograph with $R >$ 100 000, operating in the wavelength range 0.35–0.7 μm (preferably going down to 0.33 μm), with long-term radial velocity stability of $<$ 2 cm s^{-1}, observatory lifetime $\gtrsim$ 20 yr.

19.2 Implications for the observatory system

19.2.1 Telescope

Diameter: Almost all DRM cases are at least in part sensitivity limited. Given that the exposure time needed to achieve a given S/N scales as D^{2-4} (depending on the circumstances), and given that for D = 42 m several cases (e.g. S3-2, C10-1, C2) are already operating at the edge of feasibility in terms of total exposure time required, it seems unlikely that a smaller D could be compensated for by further increasing the exposure times. In addition, the cases S3-1, S9 and G4-1 require the diffraction limited resolution of a $D \geq$ 42 m telescope to be able to achieve their science goals. Hence a large fraction of the science represented in the DRM relies on the E-ELT having $D \geq$ 42 m for its accomplishment.

Field of view: The strongest requirements on the size of the unobscured field of view are set by the high-redshift galaxy survey cases C10-1, C10-2 and C4. These requirements are based on the surface densities of high-redshift galaxies (which are not well known for the highest redshifts) and, for the cases C10-1 and C4, are coupled to the (unknown) multiplex capabilities of the E-ELT's multi-object spectrograph(s). This results in only an approximate requirement for the field of view size of $\sim$25 arcmin^2.

Mirror coatings: The telescope's mirror coatings affect the observatory system performance in two ways: they determine the transmission of the telescope as well as the emissivity of the mirrors. The emissivity governs (along with the ambient temperature) the thermal emission from the telescope (although other components, such as the support structure, the gaps between the segments of M1 and dust on the mirrors will also contribute) which in turn dominates the background over large parts of the thermal infrared wavelength regime. The DRM explored the effect of the mirror coatings on the simulation results by considering two different coatings (see [RD1] for their detailed properties): a bare aluminium (Al) coating and a protected silver/aluminium (Ag/Al) coating. The latter is superior to the former in all respects except that only the Al coating provides access to the ultraviolet.

All sensitivity and background-limited cases that require observations in the thermal infrared (S5-1, S5-2, S5-3, G4-1, G9, C10-1, C10-2) prefer the Ag/Al coating. Its emissivity is lower than that of the Al coating by a factor of 2.9, which translates to a background brightness difference of $\sim$0.4 mag arcsec^{-2} in the K-band and 1.1 mag arcsec^{-2} in the L, M and N bands. Hence the cases S5-1 and S5-2 are particularly affected by this issue.

The reflectivity of the Ag/Al coating is higher than that of the Al coating at all wavelengths $>$ 0.37 μm. The difference is particularly large near 0.8 μm where the telescope transmission provided by the Ag/Al coating is a factor of 1.8 higher than that of the Al coating. Therefore all sensitivity-limited DRM cases operating in the I-band (S5-1, G4-1, G4-2, G9) have a strong preference for the Ag/Al coating. On the other hand, only the Al coating provides access to the ultraviolet, with the Al transmission exceeding that of the Ag/Al coating by more than a factor of 2 at $\lambda <$ 0.33 μm. However, observations in this regime are only required by the case G4-3 for a part of its science goals.

In conclusion, the trade-off between the Al and Ag/Al coatings consists of access to the ultraviolet on the one hand, and substantially improved transmission in the I-band as well as a considerably lower background in the mid-infrared on the other hand. When considering all of the science represented in the DRM this trade-off must probably be decided in favour of the Ag/Al coating.

19.2.2 Instrumentation

The following types of instruments cover most of the requirements of the science cases represented in the DRM: diffraction-limited, red-optical and near-infrared imager (S9, S5-2, S5-3, G4-1, C10-2); diffraction-limited, mid-infrared imager (S9, S5-2); diffraction-limited, intermediate resolution, red-optical and near-infrared integral field spectrograph (S5-1, S5-2, G4-2, G9); diffraction-limited, intermediate resolution, mid-infrared integral field spectrograph (S5-1, S5-2); wide-field multi-object near-infrared integral field spectrograph (C10-1, C4); high-resolution, high-stability optical spectrograph (S3-2, G4-3, C2); high-contrast, XAO imager (S3-1).

[RD4] contains a comparison of the instrumentation requirements of each DRM case with the capabilities of each of the instruments included in the E-ELT phase A instrumentation study plan. This comparison revealed that the instruments studied under this plan match the DRM requirements very well. A notable exception is the DRM case G4-3 which requires observations below 0.37 μm for some of its science goals. This would not be possible with the studied instruments. Other areas of concern are the size of the field of view required by the case C10-2, as well as the size of the patrol field of view and multiplexity required by the case C4. Nevertheless, the overall conclusion is that the instruments of the phase A studies would enable almost all of the science represented in the DRM. Thus the DRM provides no indication that any important capabilities have been missed by the phase A study plan.

19.2.3 Adaptive optics

Several DRM cases (S3-1, S9, S5-2, S5-3, G4-1, G9, C10-2) operate at the diffraction limit of the telescope, either because they require the resolution or for sensitivity reasons or both. The requirements of most of these cases can be satisfied with LTAO, except for the high-contrast imaging case S3-1, which requires XAO, and the case C10-2, which needs a larger field of view than provided by LTAO and hence requires MCAO. Although they do not operate at the diffraction limit, the cases C10-1 and C4 nevertheless need good AO correction (over a large field of view) for sensitivity reasons and thus require MCAO or, since the correction is only needed for small patches in the field of view, MOAO. GLAO satisfies the requirements of only one case, S5-1, although this case does not require a large field of view and would benefit from using LTAO instead.

19.2.4 Site

Since the site for the E-ELT (Cerro Armazones) was selected only towards the very end of the DRM process, many of the DRM simulations explored the effects of varying some of the most important site parameters. In particular, the simulations considered two different, idealised sites (see [RD1] for their detailed properties): a Paranal-like site and a higher, dryer, cooler, and therefore superior site; where Armazones must clearly be identified with the former. In general, the influence of the site on the observatory system performance can be separated into three distinct issues: image quality, the brightness of the background, and transmission. The DRM's exploration of the effect of the site on the simulation results was, however, limited to the latter two issues.

The site affects the background most strongly in the thermal infrared. The site determines the thermal emission from the atmosphere as well as the ambient temperature, which in turn governs (along with the emissivity) the thermal emission from the telescope and all other optical components that are not specifically cooled. Those sensitivity and background-limited science cases operating at the shortest thermal infrared wavelengths (second half of the K-band, L and M bands) are particularly sensitive to the ambient temperature because this region corresponds to the steep Wien part of the black body curve. Consequently, all DRM cases affected by this issue (S5-1, S5-2, S5-3, G4-1, G9, C10-1, C10-2) would prefer the High & Dry site. However, most of these cases do not require observations beyond the K-band, where the temperature difference of 15 K between the Paranal-like and High & Dry sites translates to a background brightness difference of only $\sim$0.4 mag arcsec^{-2}, and so overall this issue is not critical to the science represented in the DRM.

The site also determines the atmospheric transmission. However, almost all of the DRM cases only require observations in wavelength regions where the atmospheric transmission is close to 1, and so any site dependence of the transmission is not an issue. The only exception is the case G4-3, part of which requires observations at $\lambda < 0.37$ μm where the transmission differs by > 30 % between the two sites assumed for the DRM, and hence this case has a strong preference for the High & Dry site.

Not surprisingly then, a number of DRM cases would have benefited from a higher and cooler site than Armazones. However, the main conclusion is that the choice of Armazones for the E-ELT site does not render any of the science represented in the DRM completely unfeasible.

19.2.5 Operations

The DRM simulations have shown that several cases (S3-2, possibly G4-2, G4-3, C10-1, C4, C2) require very large amounts of total observing time, well beyond the scope of current Large Programmes on the VLT. The observatory's science policies and operations must be able to accommodate these. In addition to requiring a large amount of total observing time, the case S3-2 also requires a specific observing cadence on several timescales, with a total programme duration of several years. The case C2 even requires a total programme duration of 15–20 yr, thus placing a requirement on the observatory lifetime.

A rough, non-rigorous estimate of the total observing time required to accomplish the minimum science goals represented in the DRM is $\sim$7500 h. Including overheads this amounts to more than 3 yr.

19.3 Final remarks

Even when taking into account that the DRM cases were partly selected for being challenging, even for the E-ELT, the DRM analysis has clearly shown that the unique capabilities of the E-ELT have inspired some ambitious thinking. Although the simulations have proven that almost all of the science represented in the DRM will be achievable with a 42 m E-ELT, they have also demonstrated that many cases are operating at the edge of feasibility. Thus the overall conclusion is that the E-ELT system has neither been over-specified nor under-specified: its design is well matched to the requirements of the science cases investigated by the DRM.

20 Appendix: DRM proposals

As described in Section 4, for each science case studied in the DRM members of the SWG drafted a 'DRM proposal' which described the observations to be simulated, formulated the specific questions to be explored and provided basic scientific input data. These proposals served as the starting point for the work described in this document, and we attach them below for completeness.

European Organisation for Astronomical Research in the Southern Hemisphere

Organisation Européenne pour des Recherches Astronomiques dans l'Hémisphère Austral
Europäische Organisation für astronomische Forschung in der südlichen Hemisphäre

VISITING ASTRONOMERS DEPARTMENT • Karl-Schwarzschild-Straße 2 • D-85748 Garching bei München • e-mail: visas@eso.org • Tel. : +49-89-32 00 64 73

APPLICATION FOR OBSERVING TIME PERIOD: **80A**

Important Notice:

By submitting this proposal, the PI takes full responsibility for the content of the proposal, in particular with regard to the names of CoIs and the agreement to act according to the ESO policy and regulations, should observing time be granted

1. Title Category: **C–7**

Direct imaging of terrestrial and giant exoplanets

2. Abstract

We propose to perform a direct imaging search in the J-band (1.2 μm) for terrestrial and giant planets around 150 FGK stars within 20 pc from the Sun. Using the 42m-ELT equipped with a high-contrast coronagraph and simultaneous differential imaging we intend to achieve a contrast of 1 x 10^{-10} beyond an angular distance of 8λ/D (>50 mas) from each parent star. The observations will be conducted to ensure detection (S/N=5) of the reflected light from Earth-like planets in the habitable zone of the parent stars (i.e. at 1 AU orbital distance of the solar-type stars in the sample). This will allow a rather extensive demographic study of terrestrial and giant exoplanets over a wide range of separations leading to a better understanding of the formation mechanism of planetary systems. Second and third epoch observations will be required to determine orbital motion. Follow-up low-resolution spectroscopy (R=100) in the near IR will be pursued for the nearest detected exoplanets.

3.

Run	Period	Instrument	Time	Month	Moon	Seeing	Sky Trans.	Obs.Mode
A	80	NACO	1.5n	mar	n	$\leq 0.8''$	THN	v

4.

Number of nights/hours	Telescope(s)	Amount of time
a) already awarded to this project:	NTT	4n in 078.B-1234
b) still required to complete this project:	2.2/NTT	2n/20h

5. Special remarks:

This is a DRM proposal

6. Principal Investigator: R. Rebolo (Inst. Astrofsica de Canarias, E, rrl@iac.es)

CoI(s): I. Name2 (Leiden, NL), I. Name3 (Geneva, CH), I. Name4 (STScI, USA), I. Name5 (ESO, ESO)

7. Is this proposal linked to a PhD thesis preparation? State role of PhD student in this project

Yes / A. Student. Data important for PhD thesis / mid-course

8. Description of the proposed programme

A) Scientific Rationale: The recent discovery that at least 7% of solar-type stars host giant planets at separations of less than 5 AU has opened a new domain for research. High precision radial velocity measurements of stars and microlensing techniques provide increasing evidence that planets with terrestrial mass and radius may also be abundant (see e.g. the recent results by Udry et al. 2007). The diversity of the properties (orbital distances, eccentricities and projected mass distributions) of known giant exoplanets has already challenged traditional theories of planet formation: do these planets form via gravitational instabilities in protoplanetary disks or via accretion of planetesimals? what are the planetary environments around other stars? how typical is our Solar System? are there other Earths? how important is evolution for habitability? Characterisation via direct imaging and low-resolution spectroscopy of exoplanets in various evolutionary stages will be key to answer these questions. Direct detection will make feasible the determination of masses, radii, composition, atmospheres and temperatures both for giant and terrestrial planets at different times of evolution. This will offer unique information to understand how planets form and evolve. Direct detection of Earth-like planets in extra-solar systems may also lead to the search for bio-markers (e.g. water in the near infrared and oxygen bands in the optical far red, ozone in the mid-infrared) via low resolution spectroscopy with a sufficiently large diameter telescope or using interferometry from space.
If the frequency of terrestrial planets around solar-type stars is similar to that of giant planets (i.e. 5-7 %), we would need to survey a minimum of one hundred solar-type stars to have a significant probability of finding an Earth-like planet. We shall note, however, that according to planet formation models the frequency of terrestrial planets may indeed be significantly higher and similarly the detection rate of such a survey. Nearby FGK stars are the primary targets for this search since for a sufficiently large diameter telescope, it may be possible to resolve terrestrial planets at physical distances where liquid water can exist on the surface, in the so-called "habitable zone". The habitable zone of a star like the Sun is approximately 0.9-1.5 AU. The habitable zone around an F star occurs farther out and around K stars occurs farther in (Kasting, 1993).
There are 71 FGK dwarf stars known within 10 pc from the Sun. Out of which 27 are F and G-type and 44 are K-type stars. Up to a distance of 20 pc we may expect 8 times more FGK stars of which a large fraction (40-50 %) are likely binaries. We will consider for our search single stars and wide binaries with angular separations larger than 10 arcsec, so unwanted reflections in the optical system caused by the companion do not limit our sensitivity. Nothing prevents that terrestrial planets form around each component in these relatively wide binaries and we may gain additional insight on the planet formation mechanism including them in the sample. Removing the unwise binaries, the total number of targets is then reduced to some 400 stars out of which one half will possibly be suitable (adequate elevation) for observations from a given hemisphere. We are left with about 200 FGK single stars or members of wide binaries as potential survey targets at a distance of less than 20 pc from the Sun. Our goal is to achieve detection of the reflected light of terrestrial planets in the habitable zone around as many of these stars as possible and also detect their giant planets up to orbital distances comparable to those of the giant planets in the Solar System.
Detection of a terrestrial planet at 1 AU of solar-type stars up to a distance of 20 pc from the Sun (angular separation of 50 mas) requires a sufficiently large diameter telescope (42m or larger) that can see such a faint planet resolved at an angular distance of at least several λ/D in the presence of the stellar glare. The limiting distances of stars where an Earth can be detected will depend among other factors: on the Strehl ratio that can be achieved by the AO system, the star cancellation technique employed and the "superspeckle" treatment. Current AO systems achieve reasonable Strehl ratios in the J-band, so we will adopt here this spectral range as suitable from the technical point of view. The diffraction limit of a 42m telescope in the J-band is 6 milli-arcsec (H-band observations may also be a valid alternative albeit with a slightly worse diffraction limit). Given the high-contrast imaging needed (Earth reflects about $5\text{x}10^{-10}$ L_{sun} in the optical/near-IR), coronagraphs with differential imaging cancellation will be required. At present the highest contrast images obtained at 8-10m telescopes in the near infrared are of order 10^{-6} at 10 λ/D, but extreme-AO systems are planned which may improve contrast by two orders of magnitude in the near future. We will assume here that specially dedicated instruments (coronagraphs and differential imagers) may achieve brightness contrast of 10^{-10} at 8λ/D in the near infrared at the time of operation of the 42m ELT and that the design of the telescope does not prevent this high contrast imaging work.

B) Immediate Objective: We propose to perform a direct imaging search for planets around 150 FGK dwarf stars in the solar neighbourhood (d< 20 pc) with sufficient sensitivity for J-band detection (S/N=5) of Earth-size planets in the habitable zone of the parent stars. At a distance of 10 pc, an Earth-like planet orbiting at 1 AU of a solar-like star of absolute magnitude $M_J = 3.70$ will display a J-band magnitude of 28.5 (assuming albedo of 0.35). The search will be conducted around the nearest 80 F and G-type stars and the nearest 70 early/mid K-type stars accessible to the telescope. While F and G-type stars (absolute magnitudes M_J=2.3-3.2 and M_J=3.3-4.4, respectively) will be surveyed up to a distance of 20 pc, the intrinsic lower luminosity of K-type dwarfs and the subsequent smaller orbital distances associated to the habitable zone (0.5-0.7 AU) will restrict the search up to a distance of 15 pc and 10 pc for early K-types (M_J=4.5-5.0) and mid K-types (M_J=5-5.5), respectively. This survey will also be able to detect giant planets in a much wider range of orbital distances

8. Description of the proposed programme (continued)

from the parent stars. For instance, Jupiter-mass planets would be detected up to separations of 10-15 AU of a G-type star at a distance of 20 pc from the Sun. It will then be possible to identify planetary systems similar to ours.

We expect to detect by direct imaging more than 10 terrestrial planets and a comparable number of giants (possibly many more of each class if several terrestrial and giant planets orbit the same star as it is the case of the Solar System and also for many giant planets found in radial velocity searches). We also expect that a large fraction of the terrestrial planets will show physical conditions very close to those of Earth. The candidate planets will require second and third epoch observations for confirmation. In the long-term, follow-up of the orbital motion should allow precise determination of the orbital parameters and characterisation of the planet main physical parameters. Photometric monitoring of the giant planets as they orbit their parent stars may also show evidence for rings (the radius of the rings are expected to scale with the third power of the planet mass), and the light curve may yield the size, optical depth, the albedo and colour of these rings.

Low-resolution spectroscopy in the far red and near infrared (R=100) shall be performed in the J-H (and K-band if suitable separation from the primary) for preliminary characterisation of the atmosphere of any confirmed terrestrial planet. We remark that complementary observations in the more demanding I-band may be critical to prove the oxygen content in the atmosphere of any terrestrial planet. Any Earth-like planets detected in this survey will be crucial targets for mid-infrared spectroscopy of bio-markers (e.g., ozone) with future space interferometry missions.

Obviously, spectroscopic observations shall also be conducted with as much spectral coverage as possible for a full characterisation of any detected giant planet. Very likely, giant planets will be found at sufficiently large separations of their parent stars for AO systems to allow observations in the red part of the optical spectrum.

Specific cases: As a reference we discuss details of this search for several FGK type stars within 20 pc from the Sun where radial velocity studies have already shown the presence of giant planets. We list these cases in order of increasing distance to the Sun, i.e. from less to more demanding observations (an albedo of 0.35 is alwasy assumed):

Case 1.

a) JHK-band imaging of the giant planet in ϵ Eri. This is a K2 V star (V=3.73, J=2.3) at a distance of 3.2 pc from the Sun. Radial velocity studies (e.g. Hatzes et al. 2000, ApJ 544, L145) give evidence for a giant planet (m sin i $>$ 1.55 M_{Jup}) orbiting at 3.4 AU of the star, angular separation of 1.06 arcsec. We estimate the planet reflected light from the parent star to be $\sim$4.5 x 10^{-9} L_{sun} assuming the size of Jupiter. Contrast imaging at the level 1-2 x 10^{-9} at 175 λ/D is required to detect this giant planet with an estimated magnitude of J=23.2.

b) A super-Earth with 1.5 times the radius of the Earth (approx. 5 times the mass), in the habitable zone of this star ($\sim$0.5 AU) would display a magnitude J=23.3 but at angular separation of 26 λ/D.

Case 2.

a) JHK-band imaging of planets Upsilon Andromeda c and d. This is a F8V star (V=4.1, J=3.17) located at a distance of 13.5 pc from the Sun. The star hosts a planetary system where planet "c" has a mass of $\sim$ 2 M_{Jup} and semimajor orbital axis of 0.83 AU (angular separation of 0.06 arcsec) and planet "d" with a mass of $\sim$ 4 M_{Jup} orbits at 2.5 AU (angular separation of 0.186 arcsec). The expected J-band magnitudes of planets "c" and "d" are 21.1 and 23.5,respectively. We intend to obtain direct imaging of both planets.

Case 3.

a) JHK-band imaging of planet HD 39091b. This is a 10 M_{Jup} orbiting a solar-type star G1 IV star at a distance of 20.6 pc from the Sun (V=5.7, J=4.8). The semimajor axis of the planet's orbit is estimated at 3.3 AU, angular separation of 0.16 arcsec. Because of reflected light from the primary, we expect a J-band magnitude of J=25.8 for this rather massive planet.

b) Imaging of a super-Earth with 1.5 times the Earth radius in the habitable zone of the same star (i.e. at 1 AU) would require detection of an object with J=28.1 at 0.05 arcsec, or 8 λ/D for this band.

C) Telescope Justification: This is the only telescope able to perform the requested observations

D) Observing Mode Justification (visitor or service): visitor

E) Strategy for Data Reduction and Analysis: We will use available reduction packages.

9. Justification of requested observing time and lunar phase

Lunar Phase Justification: Provide here a careful justification of the requested lunar phase.

Time Justification: (including seeing overhead) Provide here a careful justification of the requested number of nights or hours. ESO Exposure Time Calculators exist for all Paranal and for most La Silla instruments and are available at the following web address: http://www.eso.org/observing/etc.

Calibration Request: Special Calibration - Adopt a special calibration

10. Report on the use of ESO facilities during the last 2 years

Report on the use of the ESO facilities during the last 2 years (4 observing periods). Describe the status of the data obtained and the scientific output generated.

11. Applicant's publications related to the subject of this application during the last 2 years

Name1 A., Name2 B., 2001, ApJ, 518, 567: Title of article1

Name3 A., Name4 B., 2002, A&A, 388, 17: Title of article2

Name5 A. et al., 2002, AJ, 118, 1567: Title of article3

12. List of targets proposed in this programme

Run	Target/Field	α(J2000)	δ(J2000)	ToT	Mag.	Diam.	Additional info	Reference star
A	eps Eri	03 32 55	-09 27 29	5.0	3.7		K2V	
A	HD39091	05 37 09	-80 28 08	5.0	5.7		G1IV	
A	Ups And	01 36 48	+41 24 38	5.0	4.1		F8V	

Target Notes: The planned grid of pointings around the targets listed above will be defined during the first observing night.

12b. ESO Archive - Are the data requested by this proposal in the ESO Archive (http://archive.eso.org)? If yes, explain why the need for new data.

Are the data requested in this proposal on the ESO Archive (http://archive.eso.org)? If yes, explain the need for new data.

13. Scheduling requirements

2. Specific date(s) for time critical observations:

Run	from	to	reason
A	12-nov-07	14-nov-07	parallel observation with HST

3. Unsuitable period(s) of time

Run	from	to	reason
A	15-jan-08	18-jan-08	International Conference

14. Instrument configuration

Period	Instrument	Run ID	Parameter	Value or list
80	NACO	A	IMG 54 mas/px IR-WFS	provide HERE list of filters

European Organisation for Astronomical Research in the Southern Hemisphere

Organisation Européenne pour des Recherches Astronomiques dans l'Hémisphère Austral
Europäische Organisation für astronomische Forschung in der südlichen Hemisphäre

VISITING ASTRONOMERS DEPARTMENT • Karl-Schwarzschild-Straße 2 • D-85748 Garching bei München • e-mail: visas@eso.org • Tel. : +49-89-32 00 64 73

Important Notice:

By submitting this proposal, the PI takes full responsibility for the content of the proposal, in particular with regard to the names of CoIs and the agreement to act according to the ESO policy and regulations, should observing time be granted

1. Title — Category: **C–7**

Earth twins in the habitable zone of solar-type stars

2. Abstract

Five years of sub-m s^{-1}s radial-velocity measurements of quiet stars in the HARPS-GTO planet-search programme have unveiled the tip of a large population of Neptune-mass and super-Earth planets present around ∼30 % of G and K dwarfs of the solar neighborhood, within 0.3 AU from the central star. These findings support recent results of synthetic planet-population models that also predict the existence of a large population of Earth-mass planets at all separations. We propose in this ambitious large program to directly probe the presence of Earth twins in the habitable zone in a sample of 20 close-by solar-type stars, through radial velocities obtained with the high-resolution optical spectrograph CODEX on the European Extremely Large Telescope (hereafter E-ELT). The high resolution and long term stability of CODEX coupled with the large collecting area of the E-ELT provide an unequaled facility for measuring stellar radial velocities at the few cm s^{-1} level. Simulations show that stellar noise (p-mode, granulation, activity) can be averaged down to this level for the quietest dwarf stars. The scientific goals of the proposal are 1) to detect Earth-mass planets in the habitable zone around solar-type stars and thus build-up a list of suitable targets for future space missions aiming at characterizing their atmosphere; 2) to determine the frequency of Earth twins around neighboring stars; 3) to derive statistical properties of low-mass planets, priceless constraints for planet-formation models; and 4) to characterize the multi-planet aspect of systems with Earth twins.

3.

Run	Instrument	Time	Month	Moon	Seeing	Sky Trans.	Obs.Mode
A	HRS	1400h	any	g	$\leq 0.8''$	CLR	v

4. Principal Investigator: S. Udry (Geneva Observatory, CH, stephane.udry@unige.ch)

CoI(s): C. Lovis (Geneva Observatory, CH), D. Naef (Geneva Observatory, CH), L. Pasquini (ESO, ESO), on behalf of the CODEX collaboration (Several institutes, OTHER)

5. Is this proposal linked to a PhD thesis preparation? State role of PhD student in this project

6. Description of the proposed programme

A) Scientific Rationale:

Discovering Neptunes and super-Earths with radial velocities: the HARPS experience.

From giant planets to super-Earths with radial velocities.

The discovery 14 years ago of an extra-solar planet orbiting the solar-type star 51 Peg (Mayor & Queloz 1995, Nature 378, 355) has encouraged the launch of numerous new search programs, leading now to a steadily increasing number of exoplanet detections. About 350 other planetary companions have been found to orbit dwarfs of spectral types from F to M and more massive evolved stars (Fig.1). Their properties provide important constraints for planet formation models. The majority of the exoplanets have been found through the induced Doppler spectroscopic variations on the primary star, the so-called radial-velocity (RV) technique. Because of the limitations inherent to the RV technique, most of the candidates detected so far are giant gaseous planets, similar in nature to Jupiter. However since 2004, research teams using the RV technique have been making the headlines several times for their discoveries of low-mass extrasolar planets, some of which having only a few times the mass of the Earth (Fig. 1). These discoveries were made possible thanks to the sub-m s^{-1} precision reached by HARPS-type instruments.

The HARPS high-precision programme.

Among the different surveys conducted by the HARPS consortium within the GTO time, the one aiming at finding the lightest possible exoplanets has been dedicated the largest fraction of the observing time (50 %). The "high-precision" HARPS-GTO subprogram is targeting close to 400 bright, non-active FGK dwarfs, selected from the large volume-limited CORALIE planet search sample, to which had been added the stars with already known planets. Figure 1 shows examples of low-mass planetary systems coming from these surveys: HD 69830 (Lovis et al. 2006, Nature 441, 305) and HD 40307 (Mayor et al. 2009, A&A 493, 639). In each case, at least three low-mass planets have been found with minimum masses similar to or smaller than Neptune, among them five below 10 $M_\oplus$, making them super-Earths which are very probably mainly solid (icy and/or rocky). Other examples are provided by HD 181433 with a super-Earth (7.6 $M_\oplus$) in a 3-planet system and HD 47186 with a Neptune-mass planet (22.8 $M_\oplus$), with a Saturn-mass companion (Bouchy et al. 2009, A&A 496, 527). Undoubtedly, all these exciting discoveries are important steps towards the first detection of a true Earth twin. After 5 years of observations, a lot has been learned about the RV behaviour of the stars surveyed by HARPS. The most stable ones exhibit RV dispersions at the level of 80 cm s^{-1} over several years, very close to instrumental limitations. Many others show variability at the 2-3 m s^{-1} level that often turns out to be due to orbiting low-mass planets. Most of these objects are actually found in multi-planet systems, leading to complex, low-amplitude RV curves. This is well illustrated by the HD 69830 and HD 40307 systems. The total RV dispersion before fitting any Keplerian model amounts to 3.7 m s^{-1} for HD 69830 and 2.9 m s^{-1} for HD 40307, whereas the post-fit rms dispersion of the residuals is only $\sim$0.8 m s^{-1} in both cases. This clearly demonstrates the need for the sub-m s^{-1} precision and for a large number of measurements (74 RVs for HD 69830 an 135 for HD 40307).

An emerging population of super-Earths and Neptune-mass planets.

The discovery of low-mass planets is especially important for the understanding of formation and evolution of planetary systems. Preliminary results from the HARPS program suggest that the already published discoveries only represent the tip of the iceberg. A recent census of planetary candidates among stars of the "high-precision" subprogram is revealing more than 40 possible low-mass planets ($m_2 \sin i < 30 M_\oplus$) orbiting in less than 50-60 days. Statistically, taking into account only the subsample with enough measurements to safely conclude on the existence or not of low-mass planets, this would mean that about 30 % of solar-like stars do possess such close-in ice giants and super-Earths. More than 80% of these light planets are in multi-planet systems. Although still to be confirmed, these preliminary numbers are a strong indication that a whole new population of low-mass objects is now emerging. This is supported by discoveries of similar low-mass objects at larger separations (a few AUs) using the microlensing technique (e.g. Bennet et al. 2008, ApJ 684, 663), and is also in agreement with the predictions of simulations of synthetic planet population (e.g. Mordasini et al. 2009, A&A in press, astro-ph/0904.2542; Ida & Lin 2008, ApJ 685, 584; Fig. 2). The latter theoretical approach is also predicting the existence of a large population of small-mass planets at all separations, and among them a large population of Earth-like planets at separations corresponding to the habitable zone around the star.

Limitations: instrumental, photon, and stellar noise.

The planetary minimum mass estimated from Doppler measurements is directly proportional to the amplitude of the reflex motion of the primary star. The detection of the lowest possible mass planets is thus intimately linked to the ultimate long-term precision achieved on the RV measurements of the star. Looking for the highest radial-velocity precision, we have to take into account several sources of noise. They can be classified into several broad categories: photon count, technical, and astrophysical. Each of these sources is essentially independent from the others and thus the actual precision eventually obtained on the measurements will be a quadratic combination of the different contributions. Each of them has thus to be smaller than the aimed precision (typically a few cm s^{-1}; the effect of the Earth on the Sun is 9 cm s^{-1}).

B) Immediate Objective: The main scientific goals of the proposed observations are: i) to detect

6. Description of the proposed programme (continued)

Earth-mass planets in the habitable zone around solar-type stars and thus build-up a list of suitable targets for future space missions aiming at characterizing their atmosphere; ii) to determine the frequency of Earth twins around neighboring stars; iii) to derive statistical properties of low-mass planets, priceless constraints for planet-formation models; and iv) to characterize the multi-planet aspect of systems with Earth twins. In order to achieve these goals, we propose to conduct a ultra-high radial-velocity precision survey of a sample of 20 bright Solar-type stars using the E-ELT/CODEX instrument. Investigating 20 targets, we expect to detect at least 7 planetary systems. We will fully characterize them (number of planets, orbital properties, minimum masses...). Ideally, we would prefer to survey a larger sample but 20 stars seems to be a reasonable compromise between statistics and requested telescope time. The key issue of such a challenging programme is the achieved precision. Apart from the instrument itself, the achieved radial-precision, and thus the detection capability of the survey, critically depends on the sample definition and on the observing strategy.

Definition of the sample: For the definition of the sample to be used for this program, the idea is to build up from the experience, results and data gained during the HARPS-GTO and ESPRESSO programs. The sample will consist of a set of 20 slowly-rotating, non-active, radial-velocity constant and bright G and K dwarfs. Precise activity indexes and rotational velocities are available thanks to HARPS and ESPRESSO observations allowing us to build the best possible sample for such a program. Targets with bright (less than 7-9 magnitudes fainter) close (closer than $\simeq 3$ arcsec) visual companions will be excluded from the sample in order to avoid possibly seeing-dependant contamination.

Observation strategy: As mentionned before, several error sources do affect radial-velocity measurements. From the E-ELT spectroscopic exposure time calculator, we estimate that photon noise errors below $5\,\mathrm{cm\,s^{-1}}$ can be obtained in less than 15 minutes for stars brighter than $m_V = 10$. Monte-Carlo simulations have been performed for checking our ability to damp the impact of astrophysical error sources. For example, Dumusque et al. (2009, in prep.) have simulated the impact of stellar pulsation and granulation in a realistic way (see Fig. 3). They have obtained the typical timescales and RV amplitudes of these effects. They have also determined an optimal observing strategy in terms of precision and observing time availibility that will allow us to average these effects out down to a few $\mathrm{cm\,s^{-1}}$ level. These simulations so far do not include the effect of stellar activity (i.e. spots) so these results correspond to the case of spotless stars (like the Sun at the moment). We propose to use this optimal observing strategy which can be summarized like this:

i) the exposure time of an individual velocity measurement will be 15 minutes (divided in multiple exposures if necessary to avoid saturation) in order to average out the stellar seismic noise (typical timescale: a few minutes)

ii) 3 individual velocity measurements per target per night separated by about $\simeq 2$ hours for averaging out the granulation effect (typical timescale: from 0.15 to several hours).

iii) The same strategy is repeated every second night during a 10–day observing run and one observing run per month in order to sample well the stellar rotation period (typical timescale: weeks for non-active stars and days for more active ones).

iv) We will use the entire yearly visibility period of our targets ($\simeq 7$ months per year) during at least two years in order to sample well the period interval corresponding to the habitable zone (e.g. one year for a G dwarf, less than one year for K dwarfs).

Apart from damping noise sources of astrophysical origin, this observing strategy will provide us with ultra-high precision RV data sets for every target that will have a temporal sampling sensitive to periods from less than one day to more than a year allowing the characterization of very low-mass planets at separations from 10^{-2} AU to beyond the habitable zone around their host star. The rather large number of observations per target will be sufficient for disentangling multiple low-amplitude Keplerian signals. Our experience suggests that of the order of 20 data points (epochs) per planet in the system are required for a well constrained orbital fit.

C) Telescope Justification: The presence of the Earth around the Sun changes the velocity of the Sun by about $9\,\mathrm{cm\,s^{-1}}$. In order to detect and characterize such a tiny variation on solar-type stars, we need an instrument capable of reaching such a precision ($<< 10\,\mathrm{cm\,s^{-1}}$), and stable over several years. Another important point is the signal-to-noise (S/N) required to reach such a precision. The RV precision inversely scales with the S/N of the spectra. With CODEX on the E-ELT, we reach $4\,\mathrm{cm\,s^{-1}}$ on a V=7 in 1 minute, or on a V=10 in 15 minutes (based on the E-ELT spectroscopic Exposure Time Calculator). The needs in stability, precision and large collecting area make E-ELT/CODEX the unique instrument in the world to reach the ambitious goal of this programme.

D) Strategy for Data Reduction and Analysis: The online CODEX DRS pipeline delivers calibrated high-resolution spectra and high-precision radial velocities. The optimum pipeline is developed in the frame of the CODEX project. The ultra-high precision of individual velocities will then be combined in an optimal way to efficiently average down the different contributions of the stellar limiting noise. The observational strategy had been design with similar goals (Texp long enough to dump stellar oscillation effects, several measurements per night to average granulation effects, and regular observations to cover the activity time scale).

6. Attachments (Figures)

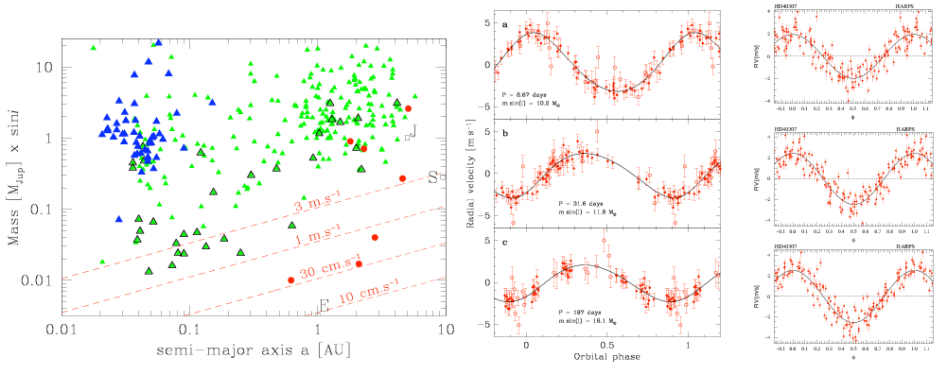

Fig. 1: **Left:** Mass-separation diagram of the more than 350 known exoplanets. The triangles refer to exoplanets found by radial velocities. The dark triangles refer to transiting exoplanets. The circles refer to exoplanets found by microlensing. The bold triangles correspond to planets discovered with HARPS. Lines of radial-velocity semi-amplitude of 3, 1, 0.3 and 0.1 ms^{-1} are shown, assuming a 1 $M_\odot$ primary star (adapted from Bouchy et al. 2009, A&A 496, 527). Examples of multi-planet systems detected with HARPS: three Neptunes in HD 69830 (Lovis et al. 2006, Nature 441, 305, **Middle**) and three super-Earths around HD 40307 (Mayor et al. 2009, A&A 493, 639, **Right**).

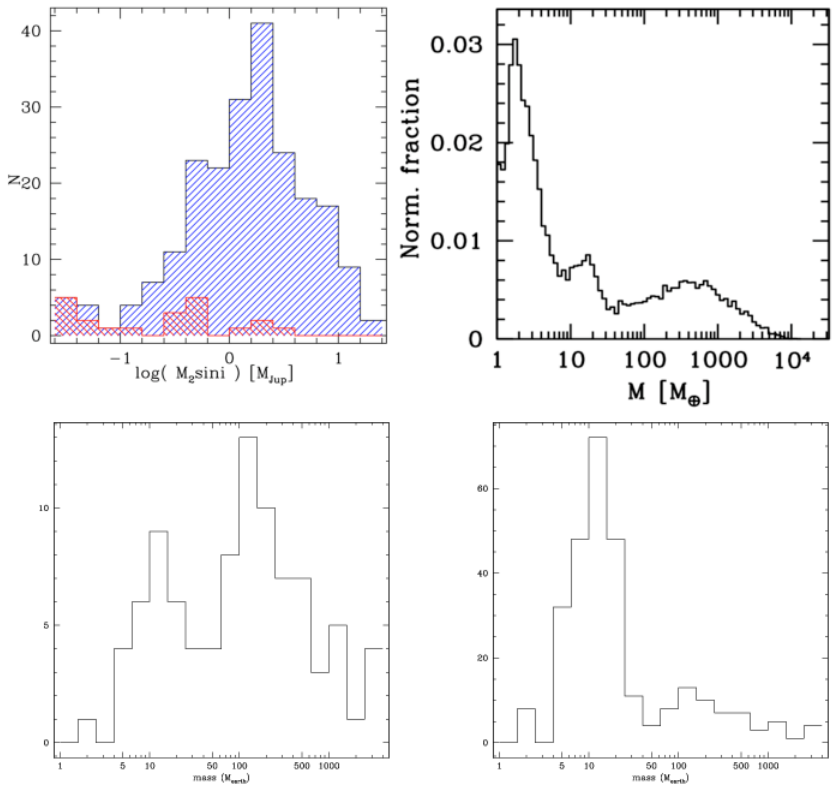

Fig. 2: **Top left:** Mass distribution of known extrasolar planets. The red-shadded is the HARPS contribution. A new low-mass population is emerging on the left side of the diagram. **Top right:** This new population is also predicted by models of synthetic planet populations (from Mordasini et al. 2009, A&A in press, astro-ph/0904.2542). **Bottom left:** Mass distribution of known exoplanets with periods shorter than 100 days. **Bottom right:** Same distribution with a normalization taking into account the fact that the sample out of which very low-mass planets are detected (i.e. the HARPS high-precision sample) is about 8 times smaller than the total number of stars in exoplanet search samples.

6. Attachments (Figures)

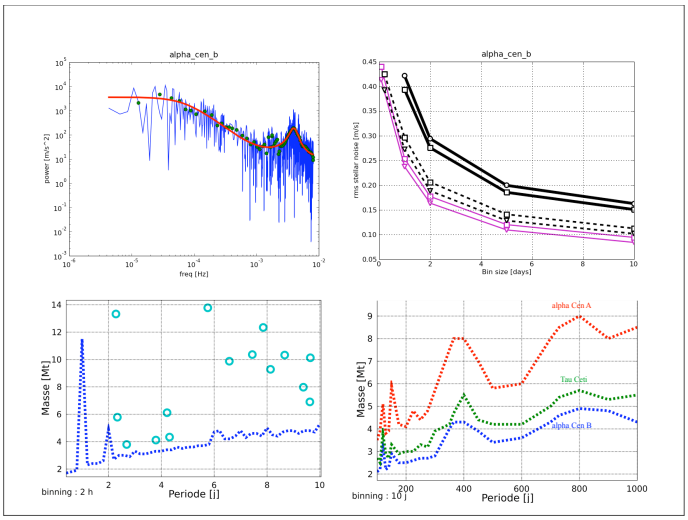

Fig. 3: Simulations aiming at optimizing the observing strategy for low-mass planets detection (from Dumusque et al. 2009, in prep.) **Top left:** Power spectrum of real α Cen B radial velocities obtained over 8 hours of consecutive observations and fitted noise model including photon noise, p-mode and granulation (bold line). **Top right:** *rms* of the synthetic RV data generated from the noise model for different observing strategies and averaged over various bins lengths in days. **Bottom:** Simulated detections limits in the planetary mass-separtion space, for optimized binning applied to a real observation calendar (the one of HD 69830), for short period planets around α Cen B (left, 2 h bins) and planets close to the habitable zone around 3 different stars (right, 10 d bins).

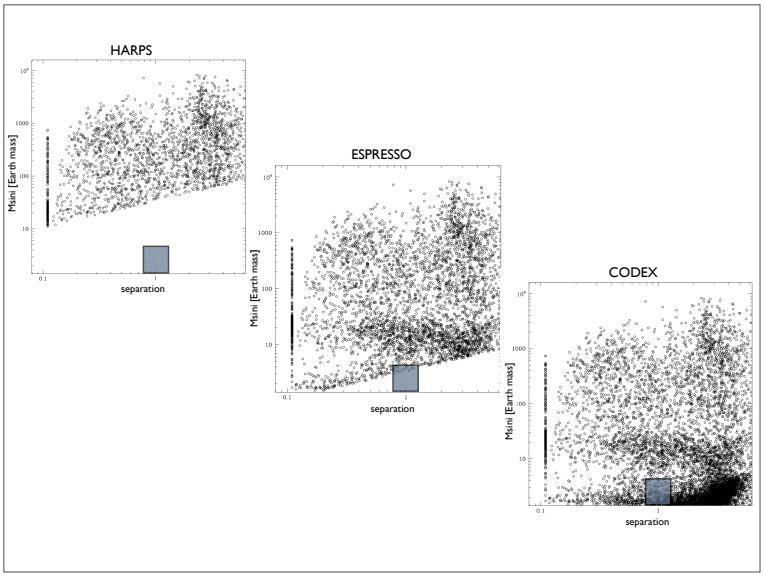

Fig .4: Expected planet population to be detected by Doppler spectroscopy with the HARPS/3.6m (precision of 1 m s^{-1}; left), ESPRESSO/VLT (precision of 10 cm s^{-1}; middle) and CODEX/E-ELT (precision of 1 cm s^{-1}; right) spectrographs, after applying observational limits for radial-velocity measurements on the predictions of planet population synthesis models with Solar-type primaries from the Bern group (Mordasini et al. 2009, in press, astro-ph/0904.2542). The detection criterium is set as the RV semi-amplitude being equal to twice the precision. Note that limitations due stellar noise are supposed to be solved by the averaging statistical approach (see text) and thus they are not taken into account in these estimates. The impact of the achieved RV precision is clearly seen.

7. Justification of requested observing time and lunar phase

Lunar Phase Justification: We have to avoid that direct or indirect Sun light reaches the detector with a contrast magnitude compared to the science target smaller than ∼9. In order to avoid significant RV errors, one would have to avoid twilight observations, avoid observations during full moon, or when target is too close to the moon, avoid observations with cloudy sky and cirri. Grey or dark time is thus required.

Time Justification: (including seeing overhead) The global need in observation time is governed by the statistics required to obtain trustable answers to the scientific questions/objectives defined for the programme (i.e. number of sample stars), and by the number of observations and their temporal sampling needed for each star to average the intrinsic stellar jitter over the different variation time scales (minutes, hours, days and weeks). The total number of data points also needs to be sufficient for convincing multi-planet orbital fits to be performed. Finally, the total observing time span has to be long enough to allow the detection of planets in the habitable zone (e.g. longer than one year in the case of a G-type primary).
Following the considerations developed in the previous sections, an optimized survey would require 3 observations of a given star per night, repeated every other night during regularly scheduled 10-night runs (once a month) over the visibility period of the star (7 months per year, i.e. 7 runs per star per year). The typical exposure time for individual integrations would be 15 minutes (as mentionned in Box 8C, this is sufficiant for obtaining a photon noise error of the order of a few $\mathrm{cm\,s^{-1}}$ for any target brighter than $m_\mathrm{V} = 10$) plus 5 minutes of overheads hence 1 hour per star per night or 5 hours per star per run or 35 hours (3.5 nights) per star per year. In order to be able to detect signals with a period close to one year, the same observing strategy has to be applied at least during two years, e.g. 7 nights (3.5 per year) per target taken during two different years. The surveyed stellar sample has to be large enough for obvious statistical reasons. We estimate that statistically useful conclusions on low-mass planet properties could be drawn from an initial sample of about 20 targets. The total number of requested nights would thus be $7 \times 20 = 140$ nights. As mentionned before, the minimal time span for a given target is two years but it would be posible to spread the total 140-night survey over more than two years. If successful, this initial exploratory sample will be extended and more observing time will be requested in order to improve the statistical significance and robustness of our findings. The positive statistical perspectives provided by the theoretical models of planet formation (Mordasini et al. 2009, A&A in press, Fig. 4) completely justify the required observational effort.

Calibration Request: Standard Calibration

8. Instrument requirements

- Extreme RV stability (2–5 $\mathrm{cm\,s^{-1}}$) over 10 yr
- Minimum energy in optical fibres: 80%
- Very stable fibre illumination
- Wavelength range required: 3800Å to 6800Å.
- Spectral resolution: $\geq 150\,000$
- FOV: a few arcsec
- Sky coverage: $\geq 90\%$

9. List of targets proposed in this programme

Run	Target/Field	α(J2000)	δ(J2000)	ToT	Mag.	Diam.	Additional info	Reference star
A	Codex1	00 00 00.00	-00 00 00.0	1.0	7.0			
A	Codex2	01 00 00.00	-00 00 00.0	1.0	7.0			
A	Codex3	02 00 00.00	-00 00 00.0	1.0	7.0			
A	Codex4	03 00 00.00	-00 00 00.0	1.0	7.0			
A	Codex5	04 00 00.00	-00 00 00.0	1.0	7.0			
A	Codex6	05 00 00.00	-00 00 00.0	1.0	7.0			
A	Codex7	06 00 00.00	-00 00 00.0	1.0	7.0			
A	Codex8	07 00 00.00	-00 00 00.0	1.0	7.0			
A	Codex9	08 00 00.00	-00 00 00.0	1.0	7.0			
A	Codex10	09 00 00.00	-00 00 00.0	1.0	7.0			
A	Codex11	10 00 00.00	-00 00 00.0	1.0	7.0			
A	Codex12	12 00 00.00	-00 00 00.0	1.0	7.0			
A	Codex13	13 00 00.00	-00 00 00.0	1.0	7.0			
A	Codex14	14 00 00.00	-00 00 00.0	1.0	7.0			
A	Codex15	16 00 00.00	-00 00 00.0	1.0	7.0			
A	Codex16	17 00 00.00	-00 00 00.0	1.0	7.0			
A	Codex17	19 00 00.00	-00 00 00.0	1.0	7.0			
A	Codex18	20 00 00.00	-00 00 00.0	1.0	7.0			
A	Codex19	21 00 00.00	-00 00 00.0	1.0	7.0			
A	Codex20	23 00 00.00	-00 00 00.0	1.0	7.0			

Target Notes: The best targets will come as output of the on-going HARPS/3.6m high-precision programmes and the future ESPRESSO/VLT surveys.

European Organisation for Astronomical Research in the Southern Hemisphere

Organisation Européenne pour des Recherches Astronomiques dans l'Hémisphère Austral
Europäische Organisation für astronomische Forschung in der südlichen Hemisphäre

VISITING ASTRONOMERS DEPARTMENT • Karl-Schwarzschild-Straße 2 • D-85748 Garching bei München • e-mail: visas@eso.org • Tel. : +49-89-32 00 64 73

APPLICATION FOR OBSERVING TIME

PERIOD: **80A**

Important Notice:

By submitting this proposal, the PI takes full responsibility for the content of the proposal, in particular with regard to the names of CoIs and the agreement to act according to the ESO policy and regulations, should observing time be granted

1. Title

Category: **C–7**

Imaging the planet-forming regions of circumstellar disks

2. Abstract

We propose to investigate the spatial structure of circumstellar disks via direct near- and mid-IR imaging with the E-ELT. The goal is to search for gaps and non-axisymmetric structure such as spirals and hotspots in the young and more mature disks indicative of ongoing or completed planet formation. In particular, it is important to reach as far as possible into the zone where putative terrestrial-mass habitable planets may be forming or have already formed. The key E-ELT requirements will be near-diffraction limited spatial resolution at 1–20 microns, with core wavelengths 2–10 micron. A diffraction-limited 42m telescope will yield a linear resolution of 2 AU at 150 parsecs, allowing access to young disks in nearby star-forming regions as seen in scattered light. Similar resolution will be achieved at 10 microns out to 30 parsecs for imaging of self-emission from dust in debris disks around nearby stars.

3.

Run	Period	Instrument	Time	Month	Moon	Seeing	Sky Trans.	Obs.Mode
A	80	NACO	100h	any	n	$\leq 0.6''$	CLR	s
B	80	VISIR	100h	any	n	$\leq 0.6''$	CLR	s

4. Number of nights/hours — Telescope(s) — Amount of time

a) already awarded to this project:

b) still required to complete this project:

5. Special remarks:

This is an E-ELT observing program, with NACO and VISIR acting as placeholders for diffraction-limited, small-field near- and mid-IR imagers, respectively.

6. Principal Investigator: M.J. McCaughrean (University of Exeter, UK, mjm@astro.ex.ac.uk)

CoI(s):

7. Is this proposal linked to a PhD thesis preparation? State role of PhD student in this project

8. Description of the proposed programme

A) Scientific Rationale: In the paradigm of the formation of low-mass stars and their planetary systems, circumstellar disks are a critical component, as they provide a conduit to channel material onto the forming star (Hartmann et al. 1997, ApJ 475 770) and supply a reservoir of material from which planets may form (Bodenheimer & Pollack 1986, Icarus, 67 391). They are also directly involved in the formation and collimation of jets and outflows, which generally thought to be implicated in the key process of angular momentum removal (e.g., Coffey et al. 2004, ApJ 604 758).

After centuries of speculation that our own solar system must have formed from a rotating disk of gas and dust, the existence of circumstellar disks around other young stellar objects (YSOs) is now firmly established (Beckwith et al. 1990, AJ 99 924; Strom et al. 1993, Protostars & Planets III, 837; Beckwith & Sargent 1996, Nature 383 139). Disks are found around $\sim$30–50% of all low-mass (0.3–3 $M_{\odot}$) young (0.3–10 Myr) stars, with radii $\sim$10–1000 AU, and masses $\sim$0.01–0.1 $M_{\odot}$. Most of these results are based on indirect measurements (e.g., spectral energy distributions, asymmetric wind profiles, polarization mapping), but in the past decade, well-resolved direct images of circumstellar disks have become available and can provide us with important information on the structure of the disks and how they are affected by their ambient environment (e.g., Burrows et al. 1996, ApJ 473 437; McCaughrean & O'Dell 1996, AJ 111 1977; Watson et al. 2007, Protostars & Planets V, 523).

Over 10–100 Myr, these optically-thick young disks lose their original gas and dust through accretion onto the central star or through agglomeration and accretion to form planets and planetesimals. The reservoir of small dust grains is then replenished by collisions between rocky planetesimals, leading to the so-called dusty "debris disks" seen around older, nearby stars such as Better spatial resolution is possible for residual, dusty "debris disks" around older, nearby stars such as Vega, β Pictoris, Fomalhaut, ε Eridani, AU Microscopii, and HD 107146. There is considerable evidence that non-axisymmetric structures in these debris disks can be used to infer the presence of planetary systems (e.g., Kalas et al. 2005, Nature, 435 1067)

The transition between young, optically-thick, gas-dominated disks and older, residual dust debris disks is crucial to our understanding of the planet formation process (e.g., Meyer et al. 2007, Protostars & Planets V, 573). In particular, high spatial resolution direct imaging of disks from 1–100 Myr in age has a crucial role to play, as protoplanets and mature planets can leave tell-tale signatures in the disks in terms of radially axisymmetric structures such as gaps and belts, and non-axisymmetric structures such as resonances and spiral density waves (Kalas et al. 2006, ApJ 637 57; Wyatt 2006, ApJ 639 1153): see also Figure 1.

However, such direct imaging studies, particularly in the crucial regions where analogue gas giants like Jupiter and terrestrial-mass planets like the Earth are expected to form. This is simply a matter of the available spatial resolution: the 5 AU radius orbit of Jupiter would subtend only 33 mas at 150 pc, the distance to the nearest low-mass star-forming regions, while a search for signs of Earth-like planets forming in a potentially habitable zone would require just 7 mas resolution at the same distance.

Matters are somewhat improved for the older, nearby debris disks, which have been probed in optical/near-IR scattered light with HST, VLT, Gemini, and Keck, and in direct thermal emission from dust grains by Spitzer and ground-based sub-mm/mm telescopes. The same fiducial distances of 5 AU and 1 AU subtend 170 mas and 35 mas at 20 pc: such resolutions are achievable at optical/near-IR wavelengths with HST and ground-based AO, but not by Spitzer or single dish sub-mm telescopes.

A fully diffraction-limited E-ELT should yield angular resolutions of 13 mas at 2μm, 60 mas at 10μm, and 120 mas at 20μm, enabling very substantial gains in disk imaging, both in terms of scattered light at the shorter wavelengths and direct thermal emission from dust at the longer wavelengths.

Images obtained of scattered light and warm (100–300 K) thermal emission from disks at near- to mid-infrared wavelengths, respectively, with the E-ELT would be directly complementary to those obtained at sub-mm/mm wavelengths with ALMA of the colder dust and gas at larger distances from the parent star, with a spatial resolution of $\sim$0.1 arcsec at 1 mm, equivalent to 15 AU at 150 pc.

Complementary spectroscopic E-ELT studies of the dynamics of gas and the chemical and physical evolution of gas and dust in the inner regions of circumstellar disks are presented in another DRM case.

B) Immediate Objective: The E-ELT will be used to obtain direct near- and mid-IR images of young and more mature circumstellar disks in search of gaps and non-axisymmetric structure such as spirals and hotspots indicative of ongoing or completed planet formation. In particular, it is important to reach as far into the zone where putative terrestrial-mass habitable planets may be forming or have already formed.

High quality, high contrast adaptive optics over a limited field (few arcsec) is required: in many cases, the parent star can be used for on-axis wavefront sensing, although a near-IR wavefront sensor would be highly recommended for more embedded sources. Post-wavefront sensing coronography will be required at short wavelengths in order to suppress the central star, with a trade-off necessary between the diameter of the occulting disk, the level of suppression, and the minimum distance from the star that can be probed. Coronography will likely not be needed at longer wavelengths, where the dust emission should far outweigh the stellar flux. Particular care should be taken to analyse the sensitivity and 'believability' of the AO system when imaging extended nebulosity with only minor structural inhomogeneities.

8. Description of the proposed programme (continued)

It would also be highly advantageous to be able to image in multiple narrow-band filters spanning a range of infrared wavelengths, in order to use emission lines in H II regions to provide high contrast images of disks seen in silhouette, as in Orion (Figure 2). Here the aim would be to image the outer edge structure of the disks to obtain insight into truncation mechanisms and the potential evidence for grain growth from the wavelength dependence of disk diameter and structure (*cf.* Throop et al. 2001, Science 292 1686).
Polarimetry may be useful for the scattered-light near-IR observations.

C) Telescope Justification: A very large, filled aperture telescope is required in order to achieve full u, v-plane imaging at the highest possible spatial resolutions. The "information content" of the images is directly proportional to the resolution (arguably even to the square of the resolution), and thus smaller telescopes such as the JWST will simply yield less detail. Interferometers such as the VLTI may have the edge in terms of raw resolution, but they are relatively poor when it comes to imaging structures on a variety of scales and any substantial degree of asymmetry.

D) Observing Mode Justification (visitor or service): Fully diffraction limited imaging is needed from 2–20μm, but covering a relatively small field only: a 100 AU diameter disk subtends just 5 arcsec at 20 pc and 0.7 arcsec at 150 pc.
(That having been said, there are disks which range up to almost 1000 AU in diameter and thus if they are nearby, as with Vega, are rather large on the sky, up to an arcminute.)

E) Strategy for Data Reduction and Analysis: Reduction and analysis of the proposed direct imaging data are relatively straightforward.

8. Attachments (Figures)

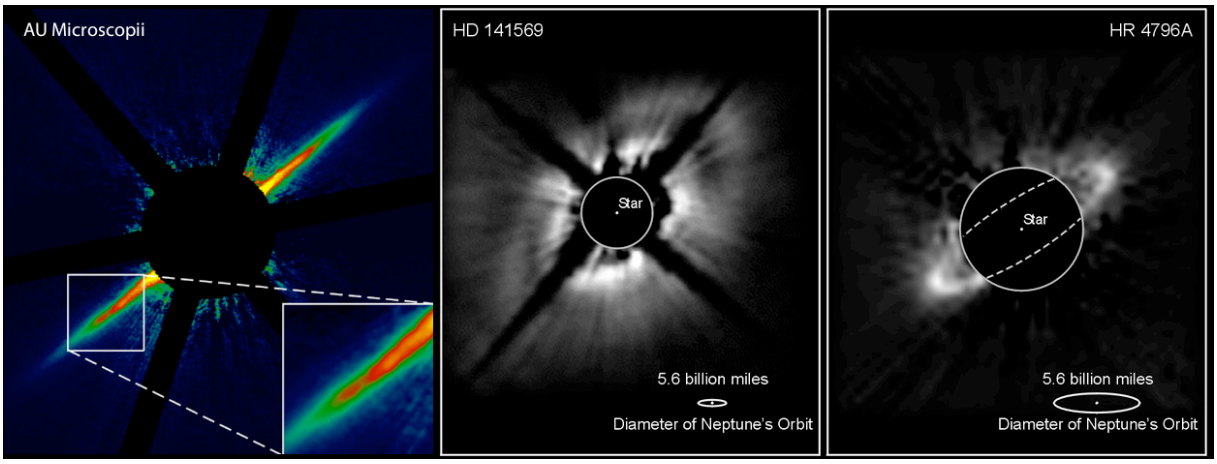

Figure 1: Near-infrared scattered light coronographic images of representative circumstellar disks. From left to right: AU Microscopii (Keck image, resolution ∼0.1 arcsec, 1 AU at 10 pc; Kalas et al. 20XX, Science 303 1990); HD 141569 (NICMOS image, resolution ∼0.2 arcsec, 20 AU at 100 pc; Weinberger et al. 1999, ApJ 525 L53); HR 4796A (NICMOS image, resolution ∼0.2 arcsec, 15 AU at 70 pc; Schneider et al. 1999, ApJ 513 127)

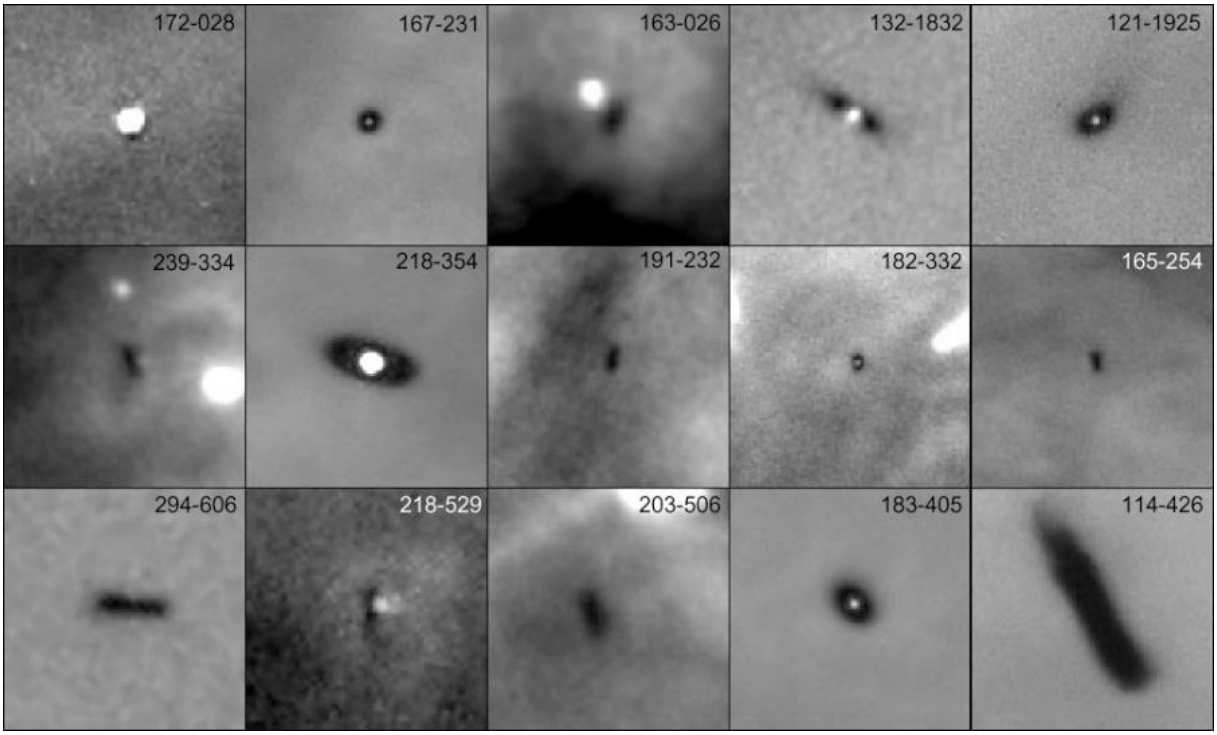

Figure 2: A gallery of young circumstellar disks seen in silhouette against the background nebulosity of the Orion Nebula H II region, as imaged at optical wavelengths with the HST. Narrow-band filters are used to obtain high contrast silhouettes against individual bright lines. Each square is 1200 AU (2.4 arcsec) across and the spatial resolution is ∼0.1 arcsec or 50 AU at the 500 pc distance to Orion. (Bally, O'Dell, & McCaughrean 2000, AJ 119 2919).

8. Attachments (Figures)

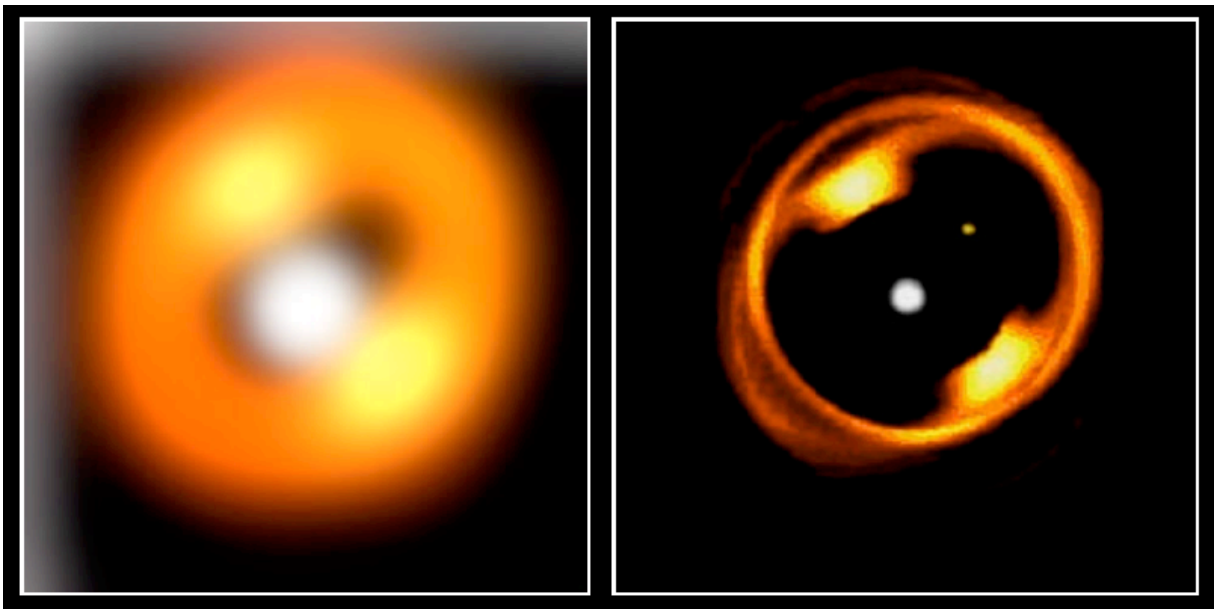

Figure 3: A representation of the gain in spatial resolution between the JWST MIRI and the E-ELT at 10μm. Both images show the simulated dust density of a Vega-like debris disk, with increased density "hotspots" in resonance with a putative planet (the yellow dot; the central star is the white dot). Although the dust density does not directly translate to mid-infrared surface brightness, it is nevertheless illustrative. The left-hand panel shows how such a system would appear to the diffraction-limited 6.5m̀ JWST if located at a distance of 40 pc: the right-hand panel shows how it would appear with 5 times higher spatial resolution, *i.e.* to a diffraction-limited telescope with a diameter in excess of 33 m. In both cases, the ring structure is roughly 5 arcsec in diameter. Adapted from Wilner et al. (2002, ApJ 569 115) via Rieke et al. (JWST MIRI Science Case).

9. Justification of requested observing time and lunar phase

Lunar Phase Justification: No constraints

Time Justification: (including seeing overhead) Circumstellar disks exhibit a very wide range of surface brightnesses within a given object, as a function of age, and as a function of wavelength. Thus it is hard to make a definitive statement re: surface brightness for input into an exposure time calculator. Some examples, however:
Vega: Surface brightness at 24 microns (from MIPS): 2 mJy/arcsec2 at 70AU radius; 0.2 mJy/arcsec2 at 200AU radius; 0.01 mJy/arcsec2 at 400AU radius
HD141569: Scattered light surface brightness: 0.05-0.15 mJy/arcsec2 at K over 1.5-5 arcsec radius (Boccaletti et al. 2003, ApJ 585 494); 0.05-0.3 mJy/arcsec2 at H (Weinberger et al. 1999, ApJ 525 L53)
GG Tau: L-band scattered light surface brightness (from Keck AO): peak 11.3 mag/arcsec2 dropping by factor of 5-10 (Duchene et al. 2004, ApJ 606 969) (see also McCabe et al. 2006, ApJ 636 932 for high spatial resolution imaging survey of T Tauri stars at 10–20μm)
HD97048: VISIR imaging at 8.6 and 11.3 micron PAH features; HAeBe star in ChamI at 180pc Doucet et al. 2007, A&A (astro-ph/0704.3678)
Potential sample sizes: there are 100s of disks within reach, both in terms of nearby mature debris disks and young disks in nearby star-forming regions

Calibration Request: Standard Calibration

10. Report on the use of ESO facilities during the last 2 years

11. Applicant's publications related to the subject of this application during the last 2 years

12. List of targets proposed in this programme

Run	Target/Field	α(J2000)	δ(J2000)	ToT	Mag.	Diam.	Additional info	Reference star
A	Vega	00 00 00	00 00 00	0	0			
A	Fomalhaut	00 00 00	00 00 00	0	0			
A	Eps Eridani	00 00 00	00 00 00	0	0			
A	HR4796A	00 00 00	00 00 00	0	0			
A	HD139664	00 00 00	00 00 00	0	0			
A	HD53143	00 00 00	00 00 00	0	0			
A	AU Mic	00 00 00	00 00 00	0	0			
A	Beta Pictoris	00 00 00	00 00 00	0	0			
A	HD107146	00 00 00	00 00 00	0	0			
A	HD15115	00 00 00	00 00 00	0	0			
A	HD141569	00 00 00	00 00 00	0	0			
A	HD32297	00 00 00	00 00 00	0	0			
A	BD+31 643	00 00 00	00 00 00	0	0			
A	HD97048	00 00 00	00 00 00	0	0			
A	Orion silhouette disks	00 00 00	00 00 00	0	0			
A	HH 30	00 00 00	00 00 00	0	0			
A	GG Tau	00 00 00	00 00 00	0	0			
A	DG Tau	00 00 00	00 00 00	0	0			
A	2MASS 1628−2430 (Ophiuchus)	00 00 00	00 00 00	0	0			
A	TW Hya	00 00 00	00 00 00	0	0			
B	HH 30	00 00 00	00 00 00	0	0			

Target Notes: The disks given are representative and all have resolved images at optical, near-, mid-, far-IR, and/or sub-mm wavelengths. See `http://astro.berkeley.edu/~kalas/disksite/pages/gallery.htm` for a gallery of images of debris disks, for example, and Caer-Eve McCabe's catalogue of resolved circumstellar disks (young and debris) at `www.circumstellardisks.org`, which presently number 99. Many other disks are known to exist via unresolved mid-IR and sub-mm excess measurements and these too would be prime targets for the E-ELT.

12b. ESO Archive - Are the data requested by this proposal in the ESO Archive (http://archive.eso.org)? If yes, explain why the need for new data.

No

13. Scheduling requirements

14. Instrument configuration

Period	Instrument	Run ID	Parameter	Value or list
80	NACO	A	IMG 13 mas/px IR-WFS	J, H, K_s
80	VISIR	B	IMG	10, 20 microns

European Organisation for Astronomical Research in the Southern Hemisphere

Organisation Européenne pour des Recherches Astronomiques dans l'Hémisphère Austral
Europäische Organisation für astronomische Forschung in der südlichen Hemisphäre

VISITING ASTRONOMERS DEPARTMENT • Karl-Schwarzschild-Straße 2 • D-85748 Garching bei München • e-mail: visas@eso.org • Tel. : +49-89-32 00 64 73

Important Notice:

By submitting this proposal, the PI takes full responsibility for the content of the proposal, in particular with regard to the names of CoIs and the agreement to act according to the ESO policy and regulations, should observing time be granted

1. Title — Category: **C–4**

Characterizing the lowest mass freely floating objects in star forming regions

2. Abstract

This proposal aims at the spectroscopic characterization of the lowest mass freely-floating objects selected in wide-area imaging surveys of young star forming regions, carried out with either survey telescopes or large format infrared imagers at 8-m telescopes. It is likely that the physical properties of such objects, as little massive as one or a few Jupiters, can be derived only from the comparison of spectra covering diagnostic atomic and molecular lines to the next generation of ultracool atmospheres. Therefore, the determination of the shape of the lowest-mass end of the substellar initial mass function may be only possible by using relatively high signal-to-noise spectroscopy that can only be provided by ELTs. Besides, spectroscopic monitoring will yield information on the meteorology of these objects, which are expected to display complex weather patterns. The goal of this project is to obtain near-simultaneous spectroscopy of the lowest-mass objects in star forming regions, in the wavelength interval ranging from the red (0.6 μm) spectral region dominated by the opacity in alkali lines, to the near-infrared (4 μm) dominated by molecular opacities (H_2, H_2O, CH_4, NH_3.)

3.

Run	Instrument	Time	Month	Moon	Seeing	Sky Trans.	Obs.Mode
A	Single-IFU	100h	any	n	n	THN	s
B	MIR-I/S	100h	any	n	n	THN	s

4. Principal Investigator: F. Comerón (ESO, ESO, fcomeron@eso.org)

CoI(s): H. Zinnecker (Potsdam, D)

5. Is this proposal linked to a PhD thesis preparation? State role of PhD student in this project

6. Description of the proposed programme

A) Scientific Rationale: Infrared survey telescopes have the potential of producing a complete census of the entire stellar and substellar contents of all nearby star forming regions down to at least the hypothetical mass where core fragmentation stops, set by the opacity limit at a few Jupiter masses (Rees 1976, MNRAS, 176, 483; Bate 2005, MNRAS, 363, 363, and references therein). The actual limit is uncertain, as some studies have shown that it depends on the initial conditions and may be lower in turbulent cores (Boyd & Whitworth 2005, A&A, 430, 1059). On the other hand, freely floating objects of truly planetary masses well below the opacity limit may exist as a consequence of the dynamical ejection from planetary systems in the earliest evolutionary stages (de la Fuente Marcos & de la Fuente Marcos 1999, New Astron., 4, 21), particularly if planets can form in the circumstellar disks around the individual components of binary systems where ejection is easier (Zinnecker 2001, ASP Conf. Ser. 239, 223). By the time the E-ELT enters operations, it is likely that complete initial mass functions down to giant-planet masses will be available for at least some of the nearest star forming regions thanks to the next generation of infrared wide-field surveys to be carried out by VISTA (e.g. Alcalá et al. 2004, Mem. Soc. Astron. Italiana Suppl., 5, 89), by pencil-beam deep observations with instruments like HAWK-I, and by other surveys of specific regions already ongoing (Caballero et al. 2007, A&A, 470, 903). Those imaging surveys may be able to show the existence of a cutoff to the lower end of the initial mass function of freely-floating objects and check theoretical predictions against actual statistics, but their quantitative results will be limited by the lack of spectroscopic information allowing accurate the determinations of their physical properties.

The faint end of the substellar mass range represents an as yet uncharted region of the parameter space populated by objects with low temperatures ($T < 1000$ K) and low surface gravities ($\log g < 3$); Baraffe et al. (2003), A&A, 402, 701). However, objects with similar temperatures, classified as late T dwarfs, are already known (e.g. Geballe et al. 2001, ApJ 556, 373; Saumon et al. 2007, ApJ, 656, 1136, and references therein) and have been theoretically modeled (Burrows et al. 2001, Rev. Mod. Phys., 73, 719; Marley et al. 2002, ApJ, 568, 335; Burrows et al. 2003, ApJ, 596, 587; 2006, ApJ, 640, 1063). These are typically evolved brown dwarfs found in the field, or in orbit around higher mass companions. Their masses are a few tens of a Jupiter mass, and they have had time during their evolution to cool down to the low temperatures observed now. The atmospheres of such objects are rich in chemistry, mineralogy, and meteorology, truly bridging the gap between the domain of stellar and planetary atmospheres. The formation of molecules and grains at such low temperatures is coupled to large-scale convection extending well into the atmosphere (Helling et al. 2008, A&A, in press), sedimentation of grains below the photosphere, raindown of chemical species, and the formation of clouds and of large-scale weather patterns (Burgasser et al. 2002, ApJ, 571, L151). Observationally, this means that the spectral features of brown dwarf atmospheres must display time variability. Photometric monitoring of T dwarfs has indeed revealed so far that aperiodic brightness variations with amplitudes of up to 0.3 mag and characteristic timescales ranging from several hours to a few days are frequent (Enoch et al. 2003, AJ, 126, 1006; Artigau et al. 2003, in "Brown Dwarfs", IAU Symp. 211, ASP Conf. Ser.). This variability has been attributed to the formation and evolution of cloud decks.

The main difference between the already well known evolved T dwarfs and objects with similar temperatures in a star forming region concerns the mass and the surface gravity. Assuming that such objects do form at all, Baraffe et al.'s (2003) models predict that a $T = 1000$ K object in a 5 Myr young star forming region should have a mass $M = 3$ M_{Jup} and a surface gravity $\log g = 3.6$, to be compared to $M = 55$ M_{Jup} and $\log g = 5.3$ for a 5 Gyr-old T dwarf of the same temperature. Observations of a handful of T dwarf-like objects in star forming regions (e.g. Zapatero Osorio et al. 2002, ApJ, 578, 536) indicate colors and spectral features that are similar, but not identical, to those of field T dwarfs, whereas recent atmosphere models clearly identify gravity-dependent spectral features (Burrows et al. 2006). On the other hand, an intriguing possibility currently under theoretical investigation is that the spectrum may contain as well telltale signatures of the formation mechanism. Such signatures would be the consequence of the differences in composition and internal structure between those objects formed in isolation by collapse and fragmentation, and those formed in an accretion disk around a star and subsequently ejected into interstellar space.

While they will be within the detection limits of imaging surveys, spectroscopy of these objects will require EELT capabilities due to their faintness even in the nearest star forming regions at a distance of ~ 150 pc. Expected magnitudes for a 5 Myr object of 1 M_{Jup} ($T_{\rm eff} \simeq 640$ K) are $I = 26.2$, $J = 22.7$, $H = 22.5$, $K = 22.9$, $L = 19.4$ (Baraffe et al. 2003). Interestingly, these same models predict that, for objects in the range of exotic colors of interest, a color-absolute magnitude relationship exists that is virtually independent of age over the entire range from 1 Myr up to 10 Gyr. The existence of such *pseudo-main sequence* provides a simple way to estimate the contamination of this region of the color-magnitude diagram of young aggregates due to non-members: objects with the same colors but brighter apparent magnitudes provide a direct measurement of the volume density of evolved objects of any age in the same color range between us and the aggregate. One should note however that luminosity estimates at such young ages are at present extremely uncertain and may fall considerably below those indicative values, as recently shown by Marley et al. (2007, ApJ, 655, 541).

Spectroscopy at moderate resolution is needed in order to analyze the composition of the dust clouds that form in the atmospheres of these objects. The main spectral regions of interest are the J, H, K, and L bands,

where the dominant sources of opacity (H_2O, NH_3, CH_4) appear, and also the far red (0.6-1.0 μm) where lines from neutral alkali elements (Li, Na, K, Cs, Rb) dominate the flux depression shortwards of 1 μm (Burrows et al. 2000, ApJ, 531, 438) and together with refractory elements (Al, Ca, Cr...) provide important diagnostics on temperature and pressure (Lodders 1999, ApJ, 519, 793) that can be extremely useful to constrain the evolutionary tracks where the objects lie. Distances to nearby star forming regions, which currently are the dominant source of uncertainty in the observational determination of luminosities, will be known to an accuracy better than ~ 1 % by the time the E-ELT enters operations thanks to the results provided by GAIA.
Time monitoring of ultracool objects will be important, as it will enable the observation of global weather pattern evolution and the development and disappearance of local atmospheric features, thus providing information on the atmospheric circulation, the timescale for the formation of local instabilities, or the role of rotation. The changing cloud coverage in these objects, the complex chemistry in their atmospheres, and the rotational modulation probably makes their spectrum evolve continuously, thus requiring integration times limited to a few hours in order to obtain meaningful snapshots. We expect rotation to be the main factor in limiting the duration of exposures, based on recent results by Zapatero Osorio et al. (2006, ApJ, 647, 1405) that show that field T dwarfs have rotation periods of 12.5 h or less and that the spin-down timescale at such low masses is very slow. Furthermore, with the advances in the modeling of the atmospheres and interiors of these objects expected in the coming decade, it may be foreseen that the spectroscopic characterization will be an even more valuable tool than it is today to relate the properties of ultracool young objects to their masses, thus becoming essential to determine the shape of the IMF at its lowest end.

B) Immediate Objective: This proposal relies on the expectation that theoretical modeling of cool brown dwarfs and freely-floating giant planet-mass objects in the next decade will have reached a stage in which spectroscopic diagnostics will be able to reliably constrain temperatures and surface gravities, and through them ages and masses. Therefore, spectroscopic classification will be the tool to use in order to derive the shape of the initial mass function, its possible truncation at the lowest masses, and perhaps even the origin of the lowest mass isolated objects.
We propose this project as one to be carried out under suboptimal conditions, when advanced AO capabilities of the EELT cannot be exploited. Therefore, we suggest to carry out the simulations assuming a modest level of AO performance.
This project should be rather straightforward from a simulation point of view. The red/infrared part should be carried out using the visible/infrared IFU spectrograph for the region shortwards of 2.5 μm, and the MIDIR infrared IFU at longer wavelengths. Theoretical spectra and fluxes of objects similar to those that we intend to observe are available, and the feasibility study should thus be reduced to the determination of the signal-to-noise ratio of the spectra of such objects after integrations of a (small) number of hours. The main questions that the simulation should address are:

* Given an exposure time limited to two hours by the object's rotation period, what are the limiting masses that can be reached at the nearest star forming regions (Chamaeleon, Lupus, Ophiuchus, R CrA...)?

* Given expected E-ELT performances in the infrared and the colors of the targets, what is the best spectral region for the spectroscopic study of these objects?

We will use Baraffe et al.'s (2003) predicted magnitudes at 5 Myr to set the fiducial cases to be considered. The second item above obviously involves a tradeoff between the maximization of the S/N ratio for a given integration time, and the scientific interest of the spectral features accessible in a given wavelength range. The table below summarizes the range of properties to be considered, for a distance modulus of 6.0 mag that is roughly applicable to all the regions listed in Section 12:

$M = 0.5\ M_{\rm Jup}$: $T = 455$ K, $\log g = 2.790$, $I = 28.25$, $J = 25.37$, $H = 24.91$, $K = 26.96$, $L = 21.03$
$M = 1\ M_{\rm Jup}$: $T = 644$ K, $\log g = 3.141$, $I = 26.34$, $J = 22.84$, $H = 22.57$, $K = 22.97$, $L = 19.56$
$M = 3\ M_{\rm Jup}$: $T = 1098$ K, $\log g = 3.576$, $I = 23.54$, $J = 19.89$, $H = 19.56$, $K = 16.21$, $L = 17.21$
$M = 10\ M_{\rm Jup}$: $T = 1965$ K, $\log g = 3.921$, $I = 19.61$, $J = 16.80$, $H = 16.44$, $K = 16.00$, $L = 15.06$

C) Telescope Justification: These are really faint objects, well beyond the reach of the VLT. While NIRSpec on the JWST will be highly competitive at near-infrared wavelengths longwards of 1 μm, the short-wavelength interval (< 1 μm) can be only covered by the E-ELT at the appropriate sensitivity level. The E-ELT also offers the advantage of making the monitoring aspect of this program easier to schedule. Finally, it may be noted again that this is a program with weak constraints on seeing, transparency, or adaptive optics performance, which would allow the E-ELT to keep doing unique science under worse-than-average atmosphere conditions.
The IFU capabilities offered by the instruments of choice are not directly relevant to this proposal, which would also be feasible with a long-slit spectrograph.

D) Strategy for Data Reduction and Analysis: Data reduction should be straightforward, as this is low-resolution spectroscopy of a single point source.

6. Attachments (Figures)

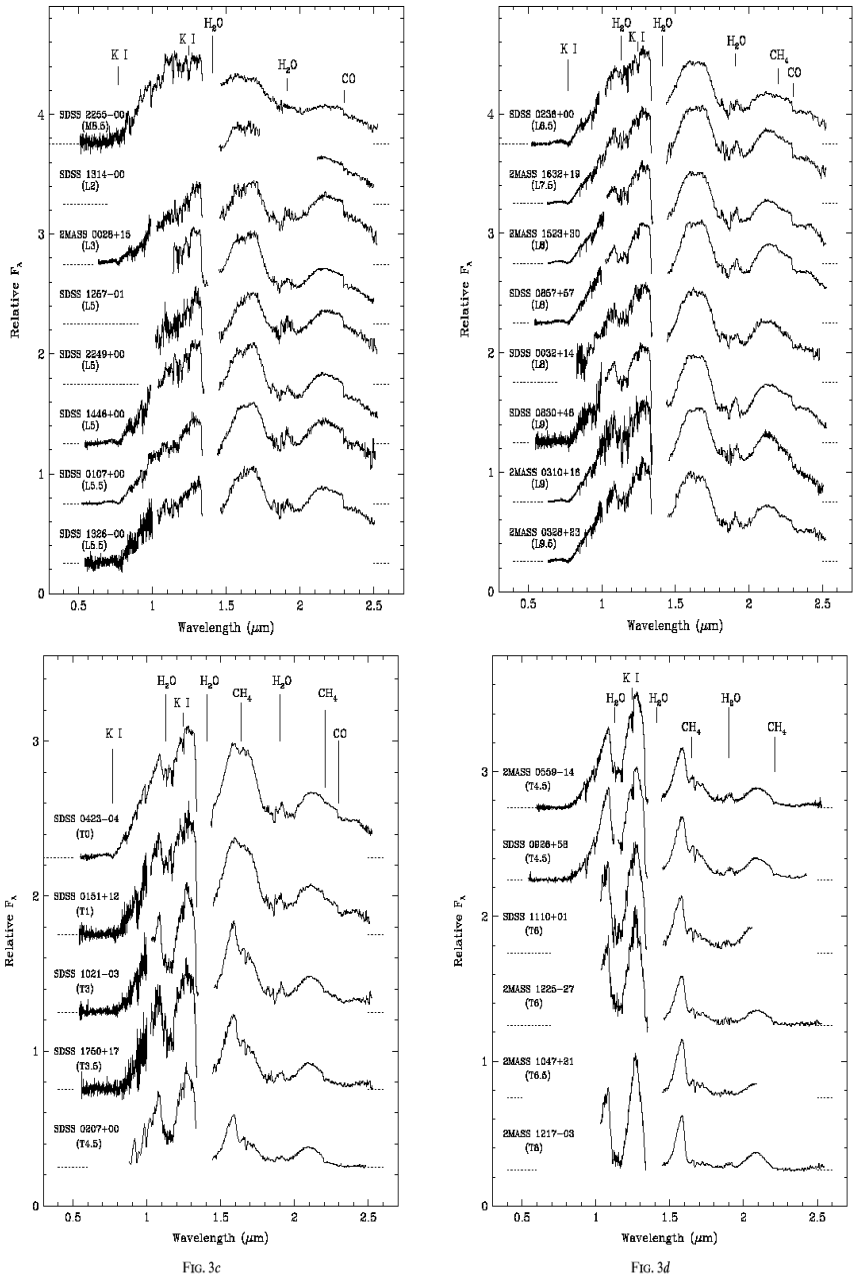

Fig. 1: Examples of field (evolved) cool spectra ranging from late M-type ($T \simeq 2300$ K) to late T-Type ($T < 1000$ K). Note the strong water vapor bands and the onset of strong CH_4 absorption at early T Types. From Geballe et al. (2002), ApJ, 564, 466.

7. Justification of requested observing time and lunar phase

Lunar Phase Justification: This is low-resolution red/infrared spectroscopy limited in time by the object spectral variability. The infrared part can be done under bright conditions, whereas grey conditions would be better for the red part.

Time Justification: (including seeing overhead) Monitoring should take place on several timescales. Every time that the program is executed the observations should extend continuously for one or possibly two full nights to sample the rotational modulation (thus sampling regional variations as the object presents different hemispheres to us). Then, they should be repeated once per week and once per month to sample the evolution of the cloud coverage and weather patterns. The estimate of 100h of observing time per instrument is provided to account for these different monitoring timescales on approximately 10 objects in a given star forming region.

Calibration Request: Standard Calibration

8. Instrument requirements

AO-fed single field IFU spectrograph (Single-IFU):
Wavelength range required: from 0.6 μm to 2.4 μm.
Spectral resolution: 3,000
Simultaneous wavelength coverage over a range as broad as possible is desirable. Together with depth it takes preference over spectral resolution.

Mid-IR imager, low- and high-resolution spectrograph (MIR-S):
Wavelength range required: L' band.
Spectral resolution: 3,000
Monitoring observations should be taken as closely as possible to those obtained with the Single-IFU instrument to avoid the effects of variability.

9. List of targets proposed in this programme

Run	Target/Field	α(J2000)	δ(J2000)	ToT	Mag.	Diam.	Additional info	Reference star
AB	Rho Oph	16 30 00	-24 00 00	2				
AB	R CrA	19 00 00	-37 00 00	2				
AB	Lupus 3	16 10 00	-39 00 00	2				
AB	Chamaeleon I	11 10 00	-77 00 00	2				

Target Notes: These are rough positions of the nearby star forming complexes targeted by this project.

EUROPEAN SOUTHERN OBSERVATORY

Organisation Européenne pour des Recherches Astronomiques dans l'Hémisphère Austral
Europäische Organisation für astronomische Forschung in der südlichen Hemisphäre

VISITING ASTRONOMERS SECTION • Karl-Schwarzschild-Straße 2 • D-85748 Garching bei München • e-mail: visas@eso.org • Tel.: +49-89-32 00 64 73

APPLICATION FOR OBSERVING TIME PERIOD: **78A**

Important Notice:

By submitting this proposal, the PI takes full responsibility for the content of the proposal, in particular with regard to the names of COIs and the agreement to act according to the ESO policy and regulations, should observing time be granted

1. Title Category: **B–2**

The Centers of Massive Dense Young Clusters: deep ELT infrared imaging and 3D spectroscopy

2. Abstract

We propose to use the 42 m ELT at 2 – 5 microns (broad band and narrow-band filters) to probe the the number density and brigthness of deeply embedded massive stars and protostars just formed in dense Galactic protocluster clouds (ultracompact HII regions, hot cores, outflow and maser sources), penetrating as much as 200 mag of visual extinction. The combination of astrometric, proper motion (1 mas/yr) and spectroscopic, radial velocity ($R \sim 10^4$) data are crucial to study dynamical processes associated with cluster formation, such as tight binary formation and gravitational interactions followed by stellar ejections. Integral field spectroscopy is needed for these dense and severely crowded clusters (up to 1000 objects per square arcsec at $K = 25 - 30$).

3.

Run	Period	Instrument	Time	Month	Moon	Seeing	Sky Trans.	Obs.Mode
A	79	ISAAC	20h	any	d	$\leq 0.4''$	PHO	v
B	79	SINFONI	66h	any	d	$\leq 0.4''$	PHO	v

4. Number of nights/hours Telescope(s) Amount of time

a) already awarded to this project:

b) still required to complete this project:

5. Special remarks:

In order to establish feasibility we need detailed simulations to determine the dynamic range of deep infrared LTAO imaging in a crowded field. Similarly, we need to determine the accuracy of sub-mas proper motion AO observations. The simulations should take into account the extra noise contribution due to infrared "cloudshine".

6. Principal Investigator: H. Zinnecker (AIP Potsdam, D, hzinnecker@aip.de)

CoI(s): F. Comeron (ESO, D), M. McCaughrean (Exeter, UK)

7. Is this proposal linked to a PhD thesis preparation? State role of PhD student in this project

8. Description of the proposed programme

A) Scientific Rationale: Massive stars ($M > 20\,M_\odot$) represent powerful engines of change/evolution within their galactic and extragalactic environment. When they die and explode as supernovae, they dramatically deposit lots of kinetic energy and chemically enriched material in the interstellar medium. Yet despite their enormous overall importance for galactic dynamical and chemical evolution, their origins and birth processes remain poorly understood (Bally & Zinnecker 2005; Zinnecker & Yorke 2007).

In this DRM proposal, we outline how future ELTs (> 30 m diameter) help to understand the formation and early dynamical evolution of massive stars embedded in highly obscured very compact HII regions. The wavelength range of 2 – 5 microns (K-,L-, and M-bands) is required to penetrate the visual extinction (up to $A_v = 200$ mag) in the dense molecular clumps ($N_{H2} = 3{\cdot}10^{23}\,cm^{-2}$) where massive stars are born. Note that J-band (1.2 micron) and H-band (1.6 micron) observations do not peer into such dense dusty birthplaces while K-band observation do ($A_J = 0.28\,A_V$, $A_H = 0.18\,A_V$, $A_K = 0.11\,A_V$; $A_V = 200$). The L- and M-band ($A_L = 0.06\,A_V$, $A_M = 0.02\,A_V$) data are needed to probe the hot and warm circumstellar matter (disks/envelopes) by means of the detection of infrared excess (K-L, L-M); the Pfund gamma (3.76 micron) and Brackett alpha (4.05 micron) recombination lines also fall into these L and M bands, respectively, and allow us (in conjunction with their flux ratio to Brackett gamma at 2.17 micron and via Menzel Case B recombination line theory) to determine the extinction towards resolved ionizing stellar sources, i.e. massive stars, in ultra-compact (size 0.1 pc) and hyper-compact (size 0.01 pc) radio HII regions, which indicate the deeply embedded sites of high-mass star formation. If we know the extinction, we can correct the K-band and V-band magnitudes and can determine the luminosities and masses of the bright central cluster stars. These massive stars are born on the Main-Sequence and are not dominated by accretion luminosity anymore (Zinnecker & Yorke 2007). They typically originate in dense, crowded stellar clusters or Trapezium systems, with linear star-to-star separations of the order of 1000 – 5000 AU (inferred from aperture synthesis radio observations of ultra- and hyper-compact HII regions). At the typical distances to massive star forming regions of 2 to 10 kpc, such linear separations translate into angular separations of 0.1 to 0.5 arcsec. Considering that there up to to a dozen embedded massive stars in such dense protocluster clumps implies a central surface density of 20 to 1000 objects per square arcsec! To convincingly resolve such crowded stellar systems, we need 10 mas angular resolution observations, the same as ALMA in the submm regime. Thus, these observations need 5 times better angular resolution at 2 microns than JWST will provide, while WFC3 on HST does not cover the required 2 – 5 micron domain but stops at 1.6 micron. HST/WFC3 may thus detect and resolve near-infrared sources in less extincted ($A_V = 50 – 100$ mag) less crowded regions (at a later stage of evolution), while JWST can go to higher extinction but only in more nearby protoclusters at 500 – 1000 pc, such as the Orion-KL region, not at 5 – 10 kpc, i.e. the Galactic Center region. The number of massive star formation sites is very limited in the solar neigborhood and grows with distance squared. Thus, ELT (together with ALMA) will be the telescope of choice, when it comes to study the variety of high-mass birthplaces in dense incipient clusters in the inner Galaxy. In the more nearby massive star formation regions (1 – 2 kpc) the angular resolution of the ELT may help to detect short-period binary and multiple systems by way of precise astrometric signatures in the reference frame of bright cluster stars and study their mass ratios. Close binary evolution of massive stars depends to a great deal on the initial binary mass ratio and is relevant as input for understanding the origin of low-mass and high-mass X-ray binaries. Finally, gravitational interactions will occur in such dense protoclusters and some heavyweights may be ejected at speeds in excess of 50 km/s (Blaauw's runaways). 50 km/s is 10 mas/yr at 1 kpc or 1 mas/yr at 10 kpc; this again shows we need 10 mas spatial resolution to catch these dynamical ejection phenomena in the distant clusters (e.g. near the Galactic Center).

To summarize: We plan to exploit the ELT's near- and mid-IR enhanced sensitivity and angular resolution to peer through huge amounts of dust extinction, and take direct nearly diffraction limited infrared images, in order to determine the stellar density in the centers of deeply embedded, massive clusters (e.g. the progenitor of NGC3603), before the gas has been expelled and the cluster hatches. This will allow us to infer if stellar collisions are likely in the formation of the most massive stars and perhaps, in combination with ALMA submm observations, if such clusters form as a result of global collapse of molecular clouds.

As an example, we derive the scientific requirements for an ELT to resolve a cluster of four massive stars (Trapezium system) embedded in a dense hypercompact HII region with radius 1000 AU at 2 kpc distance (diameter 1") behind 200 mag of visual extinction (22.5 mag extinction in the K-band at around 2 μm and 4.5 mag in the M-band at around 5 μm; see the relations given above based on the interstellar extinction law of Rieke & Lebofsky 1985). High proper motion (several mas/yr), including dynamical ejections at velocities in excess of 30 km/s due to close encounters (ca. 20 AU), in such multiple systems can be monitored over a 1 – 2 year timeline. We can also work out that we need to reach a limiting apparent magnitude of K = 28 (S/N = 5; 1 hr), to see deeply embedded early O-stars assuming an absolute O-star magnitude of M_K = -6 and a distance modulus of 11.5 (2 kpc). These requirements can be met by a giant telescope of 42 m diameter, if a natural near-IR bright guide star can be found, within 10 – 30", perhaps in the adjacent compact HII region, on which the wavefront sensing can be performed. Alternatively, a laser guide star must be used (LTAO).

Follow-up integral field spectroscopy of the incipient cluster will also be required. The first goal here would be to derive the extinction to the individual massive stars by measuring the flux ratio of the Brackett gamma

8. Description of the proposed programme (continued)

(2.17 μm) to Brackett alpha (4.05 μm) line, which changes with extinction from its intrinsic Menzel Case B value (1/3). Secondly, we expect to detect collimated H_2 jets (2.12 μm) if the stars are in a phase of active disk accretion. Also, CO bandhead emission and absorption (both the fundamental at 4.6 μm and the first overtone at 2.3 μm) must be searched for to determine the temperature of the warm gas near the stellar photosphere and its kinematics (rotation, infall). A velocity resolution of 30 km/s is required, i.e. R = 10^4. This would also allow us to detect radial velocity variations and orbital motion of massive close binaries, making use of near-IR photospheric absorption lines. Many visible massive OB stars are double-lined short-period spectroscopic binaries, but the question is whether they are born as close binaries or if they form by disk-assisted grazing collisions (Moeckel & Bally 2007).
The requirements on the integral field spectroscopy are as follows: we would hope to use 4 k × 4 k IR-arrays, which at 5 mas/pixel provide a FOV = 20" for imaging, thus catching the wider cluster core (0.1 pc). For IFU spectroscopy, we can concentrate the unit on the central or other 2" × 2" fields, requiring 400 × 400 = 160,000 fibers. Given the desired spectral resolution of 10,000, we need 2000 pixels to disperse the whole K-band. If we use only 1/2 of the K-band (2.10 – 2.18 μm for H_2 and Brg, and 2.29 – 2.31 μm for 2-0 CO bandhand), we can fit 4 spectra in one spectral row. As we have 4 k spectral rows, we can fit 16 k spectra or spexels on one IR array detector, hence we need 160,000/16,000 = 10 IR chips to match our scientific requirements. Similar for the M-band. Not impossible, by 2015.
Finally we mention that ALMA will have a similar resolution of 10 mas at submm wavelengths as the E-ELT in the near- and thermal infrared. Hence, while ELT will be particularly good at detecting OB stars at the end of their main accretion phase, ALMA should see OB protostars, i.e. massive stellar objects in their early or main mass assembly phase. Together ELT and ALMA are a powerful combination to reveal one of the most hidden but most important secrets of stellar astrophysics: the origin of massive stars.

B) Immediate Objective: Our ELT measurements will be directed to regions with signs of active massive star formation, such as the Orion BN/KL hidden source(s), Spitzer infrared dark clouds with UCHII regions as as well as hot molecular cores with methanol (or H_2O/OH) masers, all known to pinpoint massive stars in the process of formation. These measurements have the following immediate objectives: (1) see the photospheres of deeply embedded stars and protostars and determine the stellar number density in these dense molecular cloud clumps. This is important to estimate the likelihood of stellar collisions, a process under discussion for the growth of very massive stars (Bonnell, Bate & Zinnecker 1998); (2) check if the cluster center is mass segregated at birth, i.e. whether all bright massive stars are actually located near the cluster center, a profound prediction of the theory of massive star formation by competitive accretions including the merging of subclusters (Bonnell, Bate & Vine 2003). (3) from proper motion astrometric measurements at two epochs (1 year apart) find evidence for dynamical interactions, such as gravitational slingshot effects and stellar ejections at speeds characteristic of OB runway stars, i.e. some 40 – 100 km/s corresponding to (1 – 2 mas/yr) at 4 – 8 kpc distances of typical massive star forming regions; (4) integral field spectroscopic follow-up, as described above, to complement the proper motion data with radial velocity information, also relevant to catch the close spectroscopic massive binary population at birth. It is the interactions with these "hard" binaries that give rise to the ejection of single massive stars from the cluster (e.g. de Wit et al. 2005 and "runaway" O-stars references therein). (5) try to observe the subfragmentation of a centrally-peaked, dense molecular core destined to form a massive star in its center by identifying column density fluctuations on very small sub-arcsec scales, using the NICE method of Alves, Lombardi & Lada (2007) which measures the near-infrared color excess of background stars through the cloud core in question (here K_{long} vs. K_{short}, or even K-L). Such data would settle an ongoing debate between two schools of massive star formation: Dobbs & Bonnell (2005) vs. Krumholz & McKee (2005). The former authors predict fragmentation occurs, the latter disagree and predict a monolithic collapse without fragmentation.

C) Telescope Justification: The ELT is needed because of its sheer penetrating power into dusty protocluster clouds and because of its very high angular resolution. The latter is also instrumental for measuring the expected small proper motions (sub-mas astrometry).

D) Observing Mode Justification (visitor or service): Commissioning time, science verification is needed for DRM proposals. (visitor or service).

E) Strategy for Data Reduction and Analysis: Data reduction will be non-standard, cloudshine (Foster & Goodman 2006) maybe a noise-source compromising the photometric and astrometric precision.

8. Attachments (Figures)

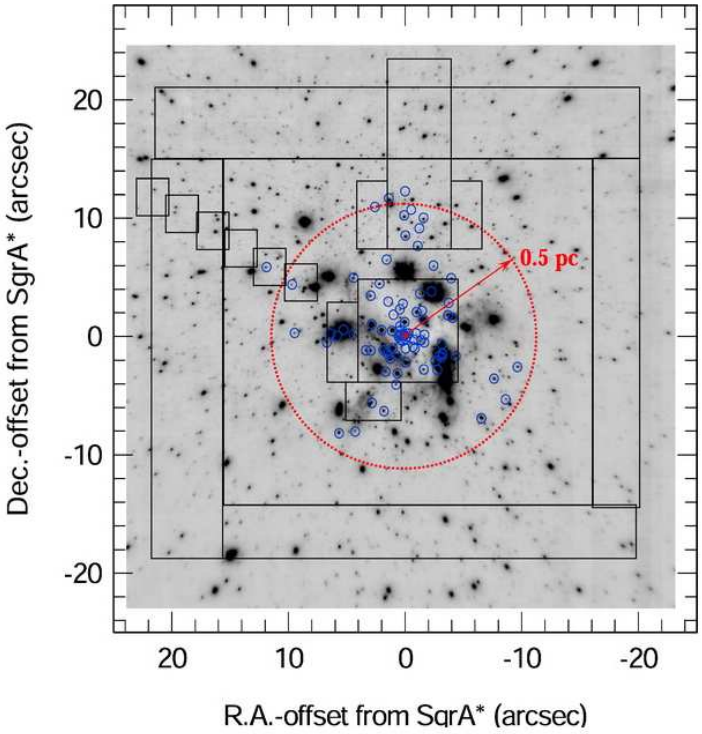

Fig. 1: Outline of the various 2003 – 2005 SPIFFI/SINFONI H+K- and K-band cubes, superposed on a $\simeq 100$ mas resolution, L-band NACO image (logarithmic scale). Small circles denote the 90 quality 1 and 2 early-type stars (OB I – V, Ofpe/WN9, W-R stars. A dotted circle denotes a 0.5 pc (20 arcsec) radius zone centered on Sgr A*, within which essentially all OB stars we have found appear to lie (from Paumard et al. 2006).

9. Justification of requested observing time and lunar phase

Lunar Phase Justification: Our infrared observations don't necessarily need dark time.

Time Justification: (including seeing overhead) Using the ELT- Experimental ETC (Version 2.4WG) assuming that Laser tomography Adaptive Optics is available, and 5mas pixels for an average star (e.g., A0), we can obtain point sources images down to a magnitude K = 28 in approximately 1 hour of integration (S/N = 5). We await detailed simulations to see how crowding and cloudshine will affect the limiting K-magnitude. L and M band limiting magnitudes are even more unclear, no information yet from the ETC at this point.
For the 6 – 8 targets selected (3 – 4 in the Galactic Center region) we estimate a total of 200 hr for the wide- and narrow-band infrared adaptive optics observations (24 hours or 2 nights per star forming region).

Calibration Request: Special Calibration -

10. Report on the use of ESO facilities during the last 2 years

Test observations with VLT/NACO

11. Applicant's publications related to the subject of this application during the last 2 years

Alves, Lombardi, & Lada 2007, A&A 462, L17; Bally & Zinnecker 2003, AJ 128, 2381; Bonnell, Bate, & Zinnecker 1998, MNRAS 298, 93; Bonnell, Bate, & Vine 2003, MNRAS 343, 413; de Witt et al. 2005, ASPC 347, 389; Dobbs & Bonnell 2005, PPV, 8170; Foster & Goodman 2006, ApJ 636, L105; Krumholz & McKee 2005, ApJ 630, 250; Megeath, Wilson, & Corbin 2005, ApJ 622, L141; Moeckel, N. & Bally 2007, ApJ 656, 275 Paumard et al. 2006, ApJ 643, 1011; Rieke & Lebofsky 1985, ApJ 288, 618; Zhang & Ho 1997, ApJ 488, 241; Zinnecker 2006, IAU-S232 Cape Town; Zinnecker & Yorke, Annual Review A&A 45, 481 – 563 (and 480 references therein)

12. List of targets proposed in this programme

Run	Target/Field	α(J2000)	δ(J2000)	ToT	Mag.	Diam.	Additional info	Reference star
A	name	RA	DEC	time(hrs)	mag	DM	ang diam(")	note
A	BN/KL	06 00	-05 00	12	M = 4	8.5	10"	Orion-IRc2 protostar
A	SgrA*	17 59	-29 00	24	10–25	14	40"	Galactic Center OB cluster
A	W51-IRS2	19 24	+14 30	8	10–25	14	10"	dense embedded cluster
A	G10.6-0.4	18 10	-19 56	8	10–25	14	10"	dense embedded HII region
B	BN/KL	06 00	-05 00	24	10–25	8.5	10"	Orion-IRc2 protostar
B	SgrA*	17 59	-29 00	24	10–25	14	40"	Galactic Center OB cluster
B	W51-IRS2	19 24	+14 30	8	10–25	14	10"	dense embedded cluster
B	G10.6-0.4	18 10	-19 56	8	10–25	14	10"	dense embedded HII region

Target Notes: Other targets near the Galactic Center include: Arches and Quintuplet clusters, a Spitzer protocluster in SgrB2; heavily obscured, deeply embedded protoclusters in infrared dark clouds (IRDC) including IRAS, ISO, MSX, and Spitzer selected regions (TBD; from Beuther/Garay) weakly obscured dense young clusters: NGC 3603, Tr14 and Tr16 in the Carina regions, and the dense center of 30 Dor (R136) in the Large Magellanic Cloud.
The targets in run A are for LTAO imaging, the targets in run B are for IFU imaging/spectroscopy.

12b. ESO Archive - Are the data requested by this proposal in the ESO Archive (http://archive.eso.org)? If yes, explain why the need for new data.

13. Scheduling requirements

14. Instrument configuration

Period	Instrument	Run ID	Parameter	Value or list
79	ISAAC	A	IMG	KLM
79	SINFONI	B	IMG	KLM

European Organisation for Astronomical Research in the Southern Hemisphere

Organisation Européenne pour des Recherches Astronomiques dans l'Hémisphère Austral
Europäische Organisation für astronomische Forschung in der südlichen Hemisphäre

VISITING ASTRONOMERS DEPARTMENT • Karl-Schwarzschild-Straße 2 • D-85748 Garching bei München • e-mail: visas@eso.org • Tel. : +49-89-32 00 64 73

Important Notice:

By submitting this proposal, the PI takes full responsibility for the content of the proposal, in particular with regard to the names of CoIs and the agreement to act according to the ESO policy and regulations, should observing time be granted

1. Title — Category: **C–4**

Giant-planet-mass objects in the Large Magellanic Cloud

2. Abstract

The goal of the proposal is to probe the complete substellar mass regime of a young star forming region in the LMC down to 5 M_{Jup}. This mass may be below the opacity limit setting the minimum mass of objects formed by fragmentation at the metallicity of the LMC. Therefore, the observations proposed here have the potential of revealing the opacity limit in a low-metallicity environment such as that of the early Milky Way, thus providing a data point that cannot be obtained from observations in our own Galaxy. The determination of the lowest-mass IMF, and eventually of the location of the opacity limit at low metallicity, will be helpful to constrain the volume density of evolved giant-planet-mass objects lurking in our own galactic disk, which have faded into invisibility since a long time ago. The challenge of these observations is set by the very small size of a typical star forming region, the crowdedness, and the coexistence of the main targets with much brighter stars that can be located a fraction of an arcsecond away.

3.

Run	Instrument	Time	Month	Moon	Seeing	Sky Trans.	Obs.Mode
A	DLI	25h	0.4	n	n	THN	s

4. Principal Investigator: F. Comerón (ESO, ESO, fcomeron@eso.org)
CoI(s): H. Zinnecker (Potsdam, D)

5. Is this proposal linked to a PhD thesis preparation? State role of PhD student in this project

6. Description of the proposed programme

A) Scientific Rationale: One of the main goals of current observational studies of the stellar and substellar content of star forming region is the detection of the lightest objects that can form in isolation (Lucas et al. 2000, MNRAS, 314, 858; 2006, MNRAS, 373, L60). The lower limit to the mass function of compact objects formed through molecular core fragmentation (stars, brown dwarfs, and isolated planetary mass object) is thought to be related to the opacity limit, i.e. the conditions under which a collapsing and fragmenting core becomes opaque to its own radiation, thus halting further fragmentation. Theoretical estimates for the opacity limit (Rees 1976, MNRAS, 176, 483; Bate 2005, MNRAS, 363, 363, and references therein) lie around 10 M_{Jup}, but simulations including realistic environment conditions of cores have shown that the opacity limit can be overcome to form objects of even smaller masses (Boyd & Whitworth 2005, A&A, 430, 1059). Indeed, observations of very young aggregates have revealed members with masses probably in the range of a few Jupiter masses only (e.g. Zapatero Osorio et al. 2002, ApJ, 578, 536; Luhman et al., 2005, ApJ 635, L93; Jayawardhana & Ivanov 2006, ApJ, 647, L167). As wide-field, deep infrared surveys of star forming regions are carried out in the next years, it may be expected that the existence and location of the lower limit to the substellar mass function will be firmly established.

Nevertheless, even if the lower limit to the initial mass function can be observationally derived soon, it will be so only for the young clusters of the solar neighborhood, all of which have virtually solar metallicity. Since the limit depends on the cooling curve of dense, cold molecular gas and the cloud opacity, metallicity is expected to play a major role in determining it. This has rather far-reaching consequences concerning the evolution of the mass function of galactic disks, since during most of its history the disk our Galaxy has been forming stars at subsolar metallicity. Therefore, knowing the shape of the lower end of the initial mass function at solar metallicity, and knowing the location of its lower mass cutoff, is of limited value for the purpose of estimating the numbers of extremely low mass objects that may exist nowadays in the galactic disk, since an extrapolation of their present-day formation rates would not be possible towards the low-metallicity ages of our Galaxy.

Can this problem be addressed observationally? It seems unlikely at present. Available models at solar metallicity (Baraffe et al. 2003, A&A, 382, 563) indicate that a 5 M_{Jup} object (similar to the lowest-mass ones detected in the σ Orionis cluster; Caballero et al. 2007, A&A, 470, 903) coeval with the Sun would have cooled down to 220 K of surface temperature having $M_J = 29.8$, $M_H = 28.8$, $M_K = 41$. The exotic colors, caused by ethane and ammonia absorption in the K band, would make the object easily identifiable as a K-band dropout (even if the sense of the dropout would be opposite to the common one, i.e., the object disappearing at the long wavelength). However, the faint absolute magnitudes mean that next-generation surveys reaching down to $J \sim 23$ could not detect them beyond 0.4 pc -or about 100,000 AU. Objects of the same mass formed at the beginning of the life of the Galaxy would have faded to $M_J = 34$, and would not be detected beyond 13,000 AU. Somewhat younger objects (ages less than ~ 3 Gyr) of the same mass could be detectable if present at a distance of a few parsecs, but their metallicities would be practically solar and thus not relevant to the determination of the evolution of the low-mass cutoff with metallicity.

Unless we are extremely lucky and one of those objects happens to be in the close vicinity of the Solar System, there is no possibility of detecting a real low-metallicity isolated planetary-mass objects in our Galaxy. Yet these objects might be extremely abundant if the early, low-metallicity Galaxy was at all able to form them.

Interestingly, the ELT will make barely possible the observation of low metallicity brown dwarfs and may answer the question of whether or not they formed, and in which amounts, in the early Galaxy. In principle old galactic planetary mass objects may be detectable with the ELT by carrying out deep surveys reaching up to a few parsecs from the Sun, but the chances to serendipitously detect one such object would be vanishingly small. However, the Magellanic Clouds are places where low-metallicity brown dwarfs are being formed *today*, and are thus many orders of magnitude brighter than their old counterparts of the same metallicity in the disk of our Galaxy. In a 1 Myr old star forming region of the LMC, our 5 M_{Jup} planetary-mass object would still have a temperature of 1900 K and magnitudes $M_J \simeq 10.6$, $M_H \simeq 10.2$, $M_K \simeq 9.7$, with a spectral type probably around mid L. At the distance modulus of the LMC ($DM = 18.5$), this implies $J \simeq 29.1$, $H \simeq 28.7$, $K \simeq 28.2$. The numbers are approximate only for several reasons. First, large uncertainties plague evolutionary models as such early ages (Baraffe et al. 2002, A&A, 382, 563), especially in the giant-planet-mass regime (Marley et al. 2007, ApJ, 655, 541). Secondly, these models are calculated for solar metallicity, whereas the atmosphere characteristics, interior opacities, and evolutionary tracks should be expected to vary sensibly for significantly subsolar metallicity.

B) Immediate Objective: The goal of this proposal is to constrain the low-mass luminosity function down to the giant planet-mass regime in some star forming regions of the Large Magellanic Cloud, with the hope of determining the as yet uncertain location of the opacity limit at the metallicity (40% solar; Hunter et al. 2007, A&A, 466, 277) of the LMC. Typical low-mass star forming regions of the solar neighbourhood like Lupus 3 or ρ Ophiuchi would subtend an angle $\sim 2''$ at the distance of the LMC, thus being appropriate for extreme AO observations which would provide diffraction-limited, high contrast, and deep imaging of the whole stellar and substellar content of the cluster. The instrument of reference is the XAO imager. The regions to be observed could be selected on the basis of their thermal infrared emission detected by other facilities, and would

normally contain stars of much higher masses, up to a few solar masses for a low-mass aggregate containing Herbig Ae/Be stars at the top end of its mass function, as is the case of well-known nearby star forming regions like Chamaeleon I, Lupus 3, or ρ Ophiuchi. The object density is unknown, as the initial mass function is unknown down to the masses of interest in this proposal, but a rough guess based on the extrapolation of the IMF of galactic star forming region (Chabrier 2003, PASP, 115, 763) yields about 20 members of the aggregate per square arcsecond.

The observations are challenging because of the high density of objects and the fact that they will cover a wide range in magnitudes, from $K \sim 18$ to $K \sim 29$. Simulations should determine whether meaningful photometry of the faintest members given the anticipated degree of crowding can be achieved, or whether the faintest members may be rendered completely undetectable because of the combined wings of the PSFs of (much) brighter members of the star forming region. Simulations can be made more realistic by assuming the presence of a variable background due to reflection nebulosity, especially in the J band, and variable extinction along the line of sight to each aggregate member. Photometric accuracy should be at the 0.1 mag level in order to confirm the low masses of the faintest detected objects on the basis of their colors. Another aspect that may be considered in the simulations is the fact that the LMC, at $\delta = -69°$, is observable at most only at high zenith distance ($> 45°$) from any of the potential ELT sites.

A possible result of the execution of this program could be the observational determination of an opacity limit lying at masses significantly above that of our fiducial 5 M_{Jup} object. Due to the steep mass-magnitude relationship at the lowest masses, an opacity limit located at 10 M_{Jup} would result in the non-detection of objects fainter than $J \sim 27.8$, $H \sim 27.4$, $K \sim 27.0$ (assuming no extinction), well above the limits that we require in this proposal.

The prescription for a realistic simulation should contain the following ingredients:

- A distribution of stellar masses drawn from the Chabrier (2003) initial mass function, whose most up-to-date analytical form (log-normal) can be found at Chabrier 2005, in "The initial mass function 50 years later", eds. Corbelli, Palla, and Zinnecker, Springer. Since we will consider only low-mass star forming regions analogous to the nearby ones, the mass function should be truncated in the high mass end to 2 $M_\odot$. The normalization factor will be chosen so that 100 objects are simulated.

- The mass is transformed to J, H, K absolute magnitudes using the evolutionary tracks of Baraffe et al. 2003 for an age of 5 Myr.

- A distance modulus of 18.50 is added to the absolute magnitudes, plus a random extinction in K, A_K, for each star drawn from a uniform probability distribution ranging from 0 to 1. The corresponding extinction in J and H are respectively taken as $A_J = 2.52A_K$, $A_H = 1.61A_K$.

- The simulated stars are scattered at random in a circle of $1''$ radius.

- Stellar light reflected by dust is introduced as an increase in the background of the region. Assuming reflection nebulosity with a total reprocessed emission in the K-band equivalent to 1/10 of the solar luminosity in that band, and typical emission nebulosity blue colors $(J-H) = -0.9$, $(H-K) = -0.6$, we obtain a background level of ~ 23.9, ~ 24.8, and ~ 25.4 mag per square arcsecond in the J, H, and K bands, respectively. It is proposed to simulate this contribution as 1- *a uniform background* level over the $1''$ radius of the star forming region, and 2- *a variable background* in which the emission per pixel takes a random value drawn from a uniform probability distribution going from zero background to a brightness that is twice the one corresponding to the values given above. This latter case, in which the background emission is taken as unresolved, is expected to be fundamental in determining limiting the detectability of the faintest sources.

C) Telescope Justification: This program may be marginally feasible only, even for a 42m ELT.

D) Strategy for Data Reduction and Analysis: Crowded field photometry will be needed to analyze the resulting images. Again, the simulations should determine the feasibility.

7. Justification of requested observing time and lunar phase

Lunar Phase Justification: This is broad-band infrared imaging, virtually unaffected by the presence of the Moon in the sky.

Time Justification: (including seeing overhead) To reach the limits $J = 29.1$, $H = 28.7$, $K = 28.2$ at $S/N = 10$, necessary exposure times according to the ETC are: 7h in J, 8h in H, and 10h in K (5 milliarcsecond sampling). The total time per region is thus 25 hours, plus overhead. Observation of several star forming regions is desirable, since their actual member surface densities will vary. The feasibility of this project is expected to be limited by a combination of factors that include the presence of too bright members, excessive crowding, excessive extinction, and variable background emission.

Calibration Request: Standard Calibration

8. Instrument requirements

Diffraction limit imager (DLI):
Wavelength range required: from 1.1 μm to 2.4 μm.

9. List of targets proposed in this programme

Run	Target/Field	α(J2000)	δ(J2000)	ToT	Mag.	Diam.	Additional info	Reference star
A	LMC	05 20 00	-69 00 00	25				

Target Notes: Accurate positions of low-mass star forming regions to be chosen on the basis of thermal infrared imaging available.

EUROPEAN SOUTHERN OBSERVATORY

Organisation Européenne pour des Recherches Astronomiques dans l'Hémisphère Austral
Europäische Organisation für astronomische Forschung in der südlichen Hemisphäre

VISITING ASTRONOMERS SECTION • Karl-Schwarzschild-Straße 2 • D-85748 Garching bei München • e-mail: visas@eso.org • Tel.: +49-89-32 00 64 73

APPLICATION FOR OBSERVING TIME PERIOD: **78A**

Important Notice:

By submitting this proposal, the PI takes full responsibility for the content of the proposal, in particular with regard to the names of COIs and the agreement to act according to the ESO policy and regulations, should observing time be granted

1. Title Category: **B–2**

The Resolved Stellar Populations of Elliptical Galaxies

2. Abstract

Elliptical galaxies represent the majority of luminous mass in the Universe and all the indirect observational indications, from the discovery of Elliptical galaxies in high red-shift surveys, to studies of integrated stellar populations, suggest that they are predominantly very old systems. However, the main theory of galaxy formation predicts that they assembled their mass relatively recently, and should therefore be dynamically young. It is important to accurately quantify this apparent contradiction. The only way to uniquely resolve this issue is to make CMDs of the resolved stellar populations in a sample of Elliptical galaxies, using the techniques developed for studies of Local Group galaxies. This means we need to reach the Virgo cluster, 17Mpc away. The detailed properties of Ellipticals will also be compared to the properties of a range of other large galaxies in the Local Group, and at distances out to and beyond Virgo to understand the effect of environment.

3.

Run	Period	Instrument	Time	Month	Moon	Seeing	Sky Trans.	Obs.Mode
A	79	FORS2	20h	any	d	$\leq 0.4''$	PHO	v
B	79	FORS2	66h	any	d	$\leq 0.4''$	PHO	v

4. Number of nights/hours Telescope(s) Amount of time

a) already awarded to this project:

b) still required to complete this project:

5. Special remarks:

Please give us time before we retire.

6. Principal Investigator: Eline Tolstoy (RUG, NL, etolstoy@astro.rug.nl)

CoI(s): The Rest (Elsewhere, ESO)

7. Is this proposal linked to a PhD thesis preparation? State role of PhD student in this project

8. Description of the proposed programme

A) Scientific Rationale: One of the main unsolved questions in astronomy is how galaxies form and evolve. In 1936 Edwin Hubble presented his tuning fork diagram classifying the different types of galaxy to be found in the Universe. This was the first attempt to find a pattern in the properties of different systems and thus search for a common evolutionary link. Elliptical galaxies appear to be shaped predominantly by a single component composed of old stars, while spiral galaxies have several components with a range of stellar ages, gas disks, dust and stellar bars. This 'Hubble tuning fork' diagram survives until today as the standard manner of presenting the morphologies of galaxies, with the addition of small dwarf spheroidal and irregular type galaxies. Fitting all these galaxy types into a common evolutionary scenario still remains to be achieved. According to the most widely accepted current structure formation scenario (the Cold Dark Matter, or CDM paradigm), all galaxies are built up from smaller pieces, the fundamental building blocks of galactic evolution, coalescing through time starting in the early universe (at high red-shift) to form the galaxies we see today. The present day systems we see around therefore provide a unique insight into the galactic assembly process and the chemo-dynamical evolution of the galaxy distribution.

The directly observable components of any galaxy are gas, dust and stars, and there is an intimate link between them. Stars form from gas, and synthesise elements in their interiors and the stars that explode at death disperse these elements into the gas from which dust and subsequent generations of stars are formed. All galaxies are thus the integrated products of all the star formation during their entire lifetimes, and the chemical elements in the stellar populations of different ages provide the most detailed evidence for this past star formation. Because low mass stars can have lifetimes comparable to the age of the Universe, the low mass tail of the ancient star formation that occurred at the formation epoch of a galaxy remains visible today and provides unique clues to the earliest physical process in the Universe. Stars of all ages provide an accurate and detailed probe of changing galaxy properties. By observing large numbers of individual stars we can measure how the rate of star formation and chemical composition of a galaxy has varied from its formation to the present and thus how galaxies were built up over time. To unravel this formative epoch detailed spatial, kinematic and chemical surveys of resolved stellar populations are required; providing a unified picture between local near-field cosmology, predictions from high red-shift surveys and theoretical simulations of galaxy formation and evolution.

Until now the sensitivity and resolution limitations have meant that detailed studies have only been possible within the Local Group and specifically around our own Galaxy. This means that only small dwarf type galaxies have had their ancient stellar populations accurately probed; massive galaxies still await this careful scrutiny. The Local Group contains only two massive galaxies (spiral systems M31 and the Milky Way) and around 40 smaller, mostly dwarf, galaxies. This is hardly representative of the range of galaxy types, and our Local Group is not necessarily representative of the high-density regions of the Universe where most galaxies live. Careful studies of dwarf galaxies have already shown inconsistencies between observations and the standard CDM picture and these need to be extended to larger galaxies to make an accurate comparison with the properties of small galaxies.

To make significant progress we need to study large numbers of resolved stars in a range of galaxy types and this requires us to look beyond the halo of the Milky Way.

B) Immediate Objective:

The Virgo cluster is the real prize for studying Elliptical galaxies. Virgo at an average distance of 17 Mpc, with over 2000 member galaxies of all morphological types, is the nearest large cluster of galaxies.

Thus, the objective of this proposal is to understand the formation and evolution of Elliptical galaxies in a range of environments using ancient stars as probes of the earliest history of star formation in the Universe. The first step is to resolve individual stars down to the oldest main sequence turnoffs, where possible (e.g., M32 a Local Group dwarf elliptical galaxy, and perhaps at the distance of Cen A). In M32 the problem to date in detecting the oldest main sequence turnoffs is partly due to sensitivity but most due to insufficient spatial resolution. For more distant galaxies, such as NGC3379, and the Virgo galaxies it is important to push for photometry below the Horizontal Branch (to detect the presence of stars > 10 Gyr old) and also to obtain main sequence turnoffs down to intermediate ages ($\sim 6-8$ Gyr). Two galaxies are selected for observing run A.

The goal is to make accurate Colour-Magnitude Diagrams (CMDs), with errors in magnitude < 0.05 mag at the faint limit, at least down to the Horizontal Branch (V,I,K ~ 32, in Virgo) in regions with a range of stellar density and stellar luminosity. These CMDs can then be used to determine accurate star formation histories and also to give targets for follow up spectroscopy to directly measure metallicity.

Elliptical galaxies require the highest spatial resolution and sensitivity. But, as with all galaxies - they have dense central region and a diffuse halo. This gives some degree of flexibility where we can make our CMDs, which is difficult to be certain of a priori - because the conversion from a surface brightness to a number of stars depends on the star formation history of the region. For example, for the same surface brightness a galaxy which is still forming stars will have fewer stars per square arcsec, because the individual stars are brighter, than a galaxy which stopped forming stars several Gyr ago.

Spiral galaxies are also interesting objects for the same type of study, but they have slightly different technical requirements, as there are quite a few examples closer than Virgo they have slightly lower demands on both

sensitivity and spatial resolution. However, partly because they are closer by they are typically larger on the sky (in our sample). They also contain (typically) a sparser stellar density in their outskirts - meaning that spatial resolution can be compromised for field of view. But these "wide" field images should always be complemented on deeper fields - with higher sensitivity and spatial resolution. This broader science case is represented in the galaxies listed in observing run B.

C) Telescope Justification: We need the ELT because of the spatial resolution AND the faintness of the point sources. We would ideally like to use images with a field of view of around 30 arcsec or greater but most important is to have images at the diffraction limit of the telescope.
I have included VIJHK filters, the minimum required is I and K. Ideally V and K.

D) Observing Mode Justification (visitor or service): visitor mode is preferred as we just love to sit in tin boxes and would like to see the Eiffeltower on its side.

E) Strategy for Data Reduction and Analysis: Here I outline the technical requirements for each galaxy in run A.

NGC4660 - A Virgo Elliptical

This is the most challenging case. At a distance modulus $(m-M)_0 \sim 31.2$. This means that the tip of the RGB is at I∼27.2, K∼25.; the HB at I∼31.8, K∼31.7; the oldest MSTO at I∼35.7, K∼35. The central surface brightness of this galaxy is given as I(0)= $\mu_0 = 16$ mag/arcsec^2 (in B, Gavazzi et 2005 A&A 430, 411), with a half-light radius $r_e = 8.6$arcsec and I(R_e) = 14.7 mag/arcsec^2 in H [H-K = 0.25mag, typically]; $r_e = 15.3$arcsec and I(R_e) = 19.6 mag/arcsec^2 in B; from Goldmine:
`http://goldmine.mib.infn.it/search-by-name.html`
(Gavazzi et al. 2004 A&A, 417, 499). The major axis of this galaxy is 1.9 arcmin; minor axis 1.24 arcmin. The total extent (∼4 disk scale lengths, 25mag/arcsec^2) is 140 arcsec ($2\times r_{ext}$). I estimate, at r_e we will have more than 5000 stars per arcsec^2 at the Horizontal branch magnitude (average star to star separation of 15mas), and in the centre, r_0, around 1 million stars per arcsec^2 (average star to star separation of 1mas), also at the Horizontal Branch. This means that the positioning of the image field depends strongly upon the achievable resolution, and the depth of the photometry.
See Tolstoy 2005 proceedings IAU 232 "Extremely Large Telescopes", eds Whitelock, Leibundgut and Dennefeld (astro-ph/0604065) for more details.
There are 2MASS images (in J,H,K), these can be obtained from web (via NED)
There are HST images (ACS), from Andres Jordan (g, z sloan filters).

NGC3379 (M105) - A slightly closer Giant Elliptical Galaxy, in Leo Group

At a distance modulus $(m-M)_0 \sim 30.3$. This means that the tip of the RGB is at I∼26.3, K∼24.1; the HB at I∼30.9, K∼30.8; the oldest MSTO at I∼34.8, K∼34.1.
The surface brightness at the half-light radius $r_e = 28.5$arcsec is I(R_e) = 15.5 mag/arcsec^2 in K (16.5 in J).
The major axis of this galaxy is 4.5 arcmin; minor axis 4.5arcmin (in R).
There are 2MASS images (in J,H,K), you can also find them via NED.
There are HST images (ACS + NICMOS, apparently quite deep).

I also send a list of stars in a typical elliptical galaxy, which stopped forming stars 5Gyr ago - so you have a representative stellar population to begin the modelling. It is just a list of stars with no fancy assumptions - no incompleteness - no noise etc. this should ideally come from the conversion into an image.
It is assuming a surface brightness= 24 mag/arcsec^2 and a limiting absolute magnitude= +6 (this is just to limit the number of stars - it doesn't make much difference after a certain point). List is no of stars, Mv, B-V, V-I, V-J, V-H, V-K, V-L, V-M.

9. Justification of requested observing time and lunar phase

Lunar Phase Justification: We require the maximum sensitivity and photometric accuracy possible which means dark time.

Time Justification: (including seeing overhead) Using the ELT- Experimental ETC (Version 2.4WG) assuming that Laser tomography Adaptive Optics is available, 5mas pixels for an average star (e.g., A0) we can obtain images with point sources down to a magnitude of V=28, I=31 and K=28 in approximately one hour of integration per filter. This means that in I, K filters we can detect the oldest MSTOS in a CMD for galaxies closer than (m-M)<25.; the Horizontal Branch for galaxies (m-M)< 27.5 - 30.5 (depending upon the HB properties); the tip of the RGB for galaxies closer than (m-M)<35.
In the case of the two galaxies in observing run A - we would like to detect the Horizontal Branch - which means V, I and K$\sim$ 31. I have not used the ETC to determine accurate exposures times instead I await the results of more detailed determinations of feasibility- and for each galaxy I list (run A and B) the required depth for detecting the TRGB (basic requirement); HB (intermediate requirement); Main Sequence Turnoffs (dream requirement).

Calibration Request: Special Calibration - Regular observations of standard star fields (!)

10. Report on the use of ESO facilities during the last 2 years

Report on the use of the ESO facilities during the last 2 years (4 observing periods). Describe the status of the data obtained and the scientific output generated.

11. Applicant's publications related to the subject of this application during the last 2 years

12. List of targets proposed in this programme

Run	Target/Field	α(J2000)	δ(J2000)	ToT	Mag.	Diam.	Additional info	Reference star
A	name	RA	DEC	time(hrs)	mag	DM	ang diam(')	note
A	NGC3379 (M105)	10 47 50	+12 34 54	5 fields	30.9, 30.8	30.3	5	second closest giant early type galaxy
A	NGC4660	12 44 32	+11 11 26	5 fields	31.8, 31.7	31.2	2	Virgo Elliptical
B	M87 (NGC4486)	12 30 49	+12 23 28	10x2hrs	31.8, 31.7	31.2	11	central elliptical in Virgo
B	M59 (NGC4621)	12 42 02	+11 38 49	5x2hrs	31.8, 31.7	31.2	8	Typical Virgo Giant Elliptical
B	M 32 (NGC221)	00 42 42	+40 51 55	5 fields	29, 28.3	24.5	9	Local Group dwarf elliptical
B	CenA (NGC5128)	13 25 28	-43 01 08.8	10x2hrs	30.5, 29.8	26	30	Closest giant early type galaxy
B	NGC4486B	12 30 32	+12 29 26	2hrs	31.8, 31.7	31.2	0.5	Virgo compact dwarf elliptical
B	Sombrero (M104)	12 39 59	-11 37 23	5x2hrs	30.3, 30.2	29.7	9	Nearest Sa galaxy
B	M83	13 37 00	-29 51 57	10x2hrs	28.9, 28.8	28.3	12	Nearest face on MW-type spiral
B	NGC300	00 54 54	-37 41 04	10x2hrs	30.3, 30.1	26.3	22	Sculptor Spiral galaxy
B	M100 (NGC4321)	12 22 55	+15 49 21	5x2hrs	31.6, 31.5	31	8	Virgo Spiral Galaxy
B	NGC 891	02 22 33	+42 20 57	5x2hrs	30.6, 30.5	30	15	Edge-on Milky Way twin
B	IC5052	20 52 01	-69 11 36	3x2hrs	29.4, 29.3	28.8	6	Edge-on galaxy
B	M 31	00 42 44	+41 16 09	20x2hrs	28.9, 28.2	24.4	200	Local Group Spiral Galaxy
B	M 33	01 33 51	+30 39 36	10x2hrs	29.1, 28.4	24.6	70	Local Group small spiral galaxy
B	NGC 55	00 14 54	-39 11 48	10x2hrs	30.2, 29.5	25.8	32	Nearby Edge-on irregular/spiral galaxy

Target Notes: The targets in run A and the main ones for this exercise; the targets in run B give a more general impression of the types of (large) galaxy we might like to observe with an ELT with the same requirements as run A.
Number of points per field listed with time, I, K mags listed. Angular diameter is the total size of the entire galaxy not always necessary to observe the entire galaxy.

12b. ESO Archive - Are the data requested by this proposal in the ESO Archive (http://archive.eso.org)? If yes, explain why the need for new data.

13. Scheduling requirements

14. Instrument configuration

Period	Instrument	Run ID	Parameter	Value or list
79	FORS2	A	IMG	VIJHK
79	FORS2	B	IMG	VIJHK

European Organisation for Astronomical Research in the Southern Hemisphere

Organisation Européenne pour des Recherches Astronomiques dans l'Hémisphère Austral
Europäische Organisation für astronomische Forschung in der südlichen Hemisphäre

VISITING ASTRONOMERS DEPARTMENT • Karl-Schwarzschild-Straße 2 • D-85748 Garching bei München • e-mail: visas@eso.org • Tel. : +49-89-32 00 64 73

APPLICATION FOR OBSERVING TIME PERIOD: **80A**

Important Notice:

By submitting this proposal, the PI takes full responsibility for the content of the proposal, in particular with regard to the names of CoIs and the agreement to act according to the ESO policy and regulations, should observing time be granted

1. Title Category: **B–2**

The Chemo-Dynamical Structure of Galaxies

2. Abstract

We propose a multi-object spectrosopic survey of the resolved stellar populations in a range of nearby galaxies out to the distance of Virgo, sampling the entire galaxy, from inner to outer regions. This will allow us to probe the total mass of all galaxies very accuratly and also to accurately determine the chemical evolution history of all components of nearby galaxies (ie., disk, bulge & halo), including a kinematic deconvolution of all these components and a comparison of their detailed properties over a range of galaxy type and environment. We will obtain a more accurate picture of the dark matter properties of a range of different galaxy types, the effects of tidal perturbation and the ubiquity of metallicity distribution functions.

3.

Run	Period	Instrument	Time	Month	Moon	Seeing	Sky Trans.	Obs.Mode
A	79	FORS2	20h	any	d	$\leq 0.4''$	PHO	v
B	79	FORS2	66h	any	d	$\leq 0.4''$	PHO	v

4. Number of nights/hours Telescope(s) Amount of time

a) already awarded to this project:

b) still required to complete this project:

5. Special remarks:

6. Principal Investigator: Eline Tolstoy (RUG, NL, etolstoy@astro.rug.nl)

CoI(s): Giuseppina Battaglia (ESO, ESO)

7. Is this proposal linked to a PhD thesis preparation? State role of PhD student in this project

8. Description of the proposed programme

A) Scientific Rationale: One of the main unsolved questions in astronomy is how galaxies form and evolve. In 1936 Edwin Hubble presented his tuning fork diagram classifying the different types of galaxy to be found in the Universe. This was the first attempt to find a pattern in the properties of different systems and thus search for a common evolutionary link. Elliptical galaxies appear to be shaped predominantly by a single component composed of old stars, while spiral galaxies have several components with a range of stellar ages, gas disks, dust and stellar bars. This 'Hubble tuning fork' diagram survives until today as the standard manner of presenting the morphologies of galaxies, with the addition of small dwarf spheroidal and irregular type galaxies. Fitting all these galaxy types into a common evolutionary scenario still remains to be achieved. According to the most widely accepted current structure formation scenario (the Cold Dark Matter, or CDM paradigm), all galaxies are built up from smaller pieces, the fundamental building blocks of galactic evolution, coalescing through time starting in the early universe (at high red-shift) to form the galaxies we see today. The present day systems we see around therefore provide a unique insight into the galactic assembly process and the chemo-dynamical evolution of the galaxy distribution.

The directly observable components of any galaxy are gas, dust and stars, and there is an intimate link between them. Stars form from gas, and synthesise elements in their interiors and the stars that explode at death disperse these elements into the gas from which dust and subsequent generations of stars are formed. All galaxies are thus the integrated products of all the star formation during their entire lifetimes, and the chemical elements in the stellar populations of different ages provide the most detailed evidence for this past star formation. Because low mass stars can have lifetimes comparable to the age of the Universe, the low mass tail of the ancient star formation that occurred at the formation epoch of a galaxy remains visible today and provides unique clues to the earliest physical process in the Universe. Stars of all ages provide an accurate and detailed probe of changing galaxy properties. By observing large numbers of individual stars we can measure how the rate of star formation and chemical composition of a galaxy has varied from its formation to the present and thus how galaxies were built up over time. To unravel this formative epoch detailed spatial, kinematic and chemical surveys of resolved stellar populations are required; providing a unified picture between local near-field cosmology, predictions from high red-shift surveys and theoretical simulations of galaxy formation and evolution.

Until now the sensitivity and resolution limitations have meant that detailed studies have only been possible within the Local Group and specifically around our own Galaxy. This means that only small dwarf type galaxies have had their ancient stellar populations accurately probed; massive galaxies still await this careful scrutiny. The Local Group contains only two massive galaxies (spiral systems M31 and the Milky Way) and around 40 smaller, mostly dwarf, galaxies. This is hardly representative of the range of galaxy types, and our Local Group is not necessarily representative of the high-density regions of the Universe where most galaxies live. Careful studies of dwarf galaxies have already shown inconsistencies between observations and the standard CDM picture and these need to be extended to larger galaxies to make an accurate comparison with the properties of small galaxies.

To make significant progress we need to study large numbers of resolved stars in a range of galaxy types and this requires us to look beyond the halo of the Milky Way.

B) Immediate Objective:

The Virgo cluster is the real prize for studying Elliptical galaxies. Virgo at an average distance of 17 Mpc, with over 2000 member galaxies of all morphological types, is the nearest large cluster of galaxies. But the main aim of this proposal is to a study a wide range of galaxy types, including spiral galaxies (e.g., NGC253, NGC300, M81) in nearby galaxy groups, such as Sculptor (in the south) or M81 (in the north) both around 2 Mpc distance. We also propose to look in more detail at galaxies in and on the edge of the Local Group, such as M31, M33 and and more distant dwarf irregular galaxies. We also include starburst galaxies NGC1569 and NGC1705 (at 2.2Mpc) and faint low surface brightness galaxies such as DD0154 (at $\sim$ 4 Mpc). This survey would also include the closest Elliptical galaxies (Centaurus A, 3.5Mpc away and NGC3379, 10Mpc) if at all possible and also dwarf elliptical galaxies, such as M32 and the large samples found in nearby galaxy groups (e.g., sample of De Rijcke et al. 2005, A&A, 438, 491).

Thus, the objective of this proposal is to understand the formation and evolution of galaxies in a range of environments using direct measurements of the chemo-dynamical properties of resolved stars as probes of the evolutionary history of these galaxies and their different components as well as their current dynamical state and thus their dark matter masses and distributions.

The most basic requirement is spectra in the CaT spectral region (8500-9000Å) of at least 1000 stars per galaxy at R$\sim$6000. Ideally each separate component of the galaxies (e.g., halo, disk, bulge) should have similarly large sample. It is also desireable to have complementary observations at high resolution (R$\sim$20000) in each system of at least 100 stars per galaxy, or component of the larger systems. This would ideally cover at least the wavelength range from 480-680nm. This would thus mimic the capabilities of the current FLAMES/GIRAFFE instrumental setup on the VLT.

8. Description of the proposed programme (continued)

Elliptical galaxies require the highest spatial resolution and sensitivity. But, as with all galaxies - they have dense central region and a diffuse halo. This gives some degree of flexibility where we can make our spectroscopic selections.
Spiral galaxies are the most easily available targets for this type of study, but they have slightly different technical requirements, as there are quite a few examples closer than Virgo they have slightly lower demands on both sensitivity and spatial resolution. However, partly because they are closer by they are typically larger on the sky (in our sample). They also contain (typically) a sparser stellar density in their outskirts - meaning that spatial resolution can be compromised for increased field of view.

C) Telescope Justification: We need the ELT because of the spatial resolution AND the faintness of the spectroscopic targets. We would ideally like to take spectra over a field of view of around 2 arcmin, or greater. It is unlikely that diffraction limited performace is required, although in the most distant targets, in Virgo, crowding of the targets is likely to be a problem if the PSF is too large.

D) Observing Mode Justification (visitor or service): visitor mode is preferred as we just love to sit in tin boxes and would like to see the Eiffeltower on its side.

E) Strategy for Data Reduction and Analysis: We still need to work out the detailed technical requirements.
Basically need imaging of the targets (with accurate colours and magnitudes), enabling a selection of RGB stars to be made.

NGC4660 - A Virgo Elliptical

This is the most challenging case. At a distance modulus $(m-M)_0 \sim 31.2$. This means that the tip of the RGB is at I∼27.2, K∼25. The central surface brightness of this galaxy is given as I(0)= $\mu_0 = 16$ mag/arcsec^2 (in B, Gavazzi et 2005 A&A 430, 411), with a half-light radius $r_e = 8.6$arcsec and I(R_e) = 14.7 mag/arcsec^2 in H [H-K = 0.25mag, typically]; $r_e = 15.3$arcsec and I(R_e) = 19.6 mag/arcsec^2 in B; from Goldmine:
`http://goldmine.mib.infn.it/search-by-name.html`
(Gavazzi et al. 2004 A&A, 417, 499). The major axis of this galaxy is 1.9 arcmin; minor axis 1.24 arcmin. The total extent (∼4 disk scale lengths, 25mag/arcsec^2) is 140 arcsec ($2\times r_{ext}$). I estimate, at r_e we will have more than 5000 stars per arcsec^2 at the Horizontal branch magnitude (average star to star separation of 15mas), and in the centre, r_0, around 1 million stars per arcsec^2 (average star to star separation of 1mas), also at the Horizontal Branch. This means that the positioning of the field depends strongly upon the achievable resolution, and the depth of the photometry.
There are 2MASS images (in J,H,K), these can be obtained from web (via NED)
There are HST images (ACS), from Andres Jordan (g, z sloan filters).

M83 - the nearest face-on spiral galaxy

At a distance modulus $(m-M)_0 \sim 28.3$. This means that the tip of the RGB is at I∼24.3 K∼22; the HB at I∼28, K∼28; the oldest MSTO at I∼32, K∼32(?).
The surface brightness is variable, depending where observations are made. The major axis of this galaxy is 13 arcmin; minor axis 11 arcmin.
There are 2MASS images (in J,H,K), you can also find them via NED.
There are HST images (ACS + NICMOS, apparently quite deep).

9. Justification of requested observing time and lunar phase

Lunar Phase Justification: We can probably live with grey time, but sensitivity is likely to be an issue, so more distant sources need dark.

Time Justification: (including seeing overhead) Using the ELT- Experimental ETC (Version 2.5WG) assuming that Laser tomography Adaptive Optics is available, 10mas pixels for an average star (e.g., K0) we can obtain spectra of point sources down to a magnitude of I=27 (S/N=10) in approximately 40 hours of integration at R=5000 in dark time, and assuming LTAO (because this gives the shortest exposure time). This means that we can obtain CaT spectroscopy for galaxies out to a distance of Virgo. For M83 this limit will be about 3 hours for stars on the tip of the RGB, although ideally one want to look at stars down to a magnitude or two below TRGB, which leads to integration times of around 20hours. with the currently available ETC numbers.
It is a bit difficult to explore too much with the ETC - because I can't play around with the resolution & pixel size too much - so the exposure times are extremely uncertain - there is flexibility in the resolution & pixel scale.
The high resolution observations are likely to be much more expensive in observing time - the ETC gives unrealistically bad predictions from what I can see - as usual in the optical.
I have selected LTAO only because it gives the shortest exposure times. the most photons on the detector is the prefered option.

Calibration Request: Special Calibration - Observations of globular cluster calibrators

10. Report on the use of ESO facilities during the last 2 years

Report on the use of the ESO facilities during the last 2 years (4 observing periods). Describe the status of the data obtained and the scientific output generated.

11. Applicant's publications related to the subject of this application during the last 2 years

12. List of targets proposed in this programme

Run	Target/Field	α(J2000)	δ(J2000)	ToT	Mag.	Diam.	Additional info	Reference star
A	name	RA	DEC	time(hrs)	mag	DM	ang diam(')	note
A	*M83	13 37 00	-29 51 57	10x2hrs	28.9, 28.8	28.3	12	Nearest face on MW-type spiral
B	*NGC3379 (M105)	10 47 50	+12 34 54	5 fields	30.9, 30.8	30.3	5	second closest giant early type galaxy
A	NGC4660	12 44 32	+11 11 26	5 fields	31.8, 31.7	31.2	2	Virgo Elliptical
B	M87 (NGC4486)	12 30 49	+12 23 28	10x2hrs	31.8, 31.7	31.2	11	central elliptical in Virgo
B	M59 (NGC4621)	12 42 02	+11 38 49	5x2hrs	31.8, 31.7	31.2	8	Typical Virgo Giant Elliptical
B	M 32 (NGC221)	00 42 42	+40 51 55	5 fields	29, 28.3	24.5	9	Local Group dwarf elliptical
B	*CenA (NGC5128)	13 25 28	-43 01 08.8	10x2hrs	30.5, 29.8	26	30	Closest giant early type galaxy
B	NGC4486B	12 30 32	+12 29 26	2hrs	31.8, 31.7	31.2	0.5	Virgo compact dwarf elliptical
B	Sombrero (M104)	12 39 59	-11 37 23	5x2hrs	30.3, 30.2	29.7	9	Nearest Sa galaxy
B	*NGC300	00 54 54	-37 41 04	10x2hrs	30.3, 30.1	26.3	22	Sculptor Spiral galaxy
B	M100 (NGC4321)	12 22 55	+15 49 21	5x2hrs	31.6, 31.5	31	8	Virgo Spiral Galaxy
B	*NGC 891	02 22 33	+42 20 57	5x2hrs	30.6, 30.5	30	15	Edge-on Milky Way twin
B	IC5052	20 52 01	-69 11 36	3x2hrs	29.4, 29.3	28.8	6	Edge-on galaxy
B	*M 31	00 42 44	+41 16 09	20x2hrs	28.9, 28.2	24.4	200	Local Group Spiral Galaxy
B	M 33	01 33 51	+30 39 36	10x2hrs	29.1, 28.4	24.6	70	Local Group small spiral galaxy
B	*ngc1569	04 30 49	+64 50 53	?				Nearby BCD galaxy
B	*ngc1705	04 54 13.5	-53 21 40	?				Nearby BCD galaxy
B	ddo154	12 54 05	+27 08 59	?				Nearby LSB/BCD galaxy
B	ngc253	00 47 33	-25 17 18	?				Nearby (Scl) spiral galaxy
B	M81	09 55 33	+69 03 55	?				Nearby spiral galaxy
B	*Sculptor Group	00 23 38	-38 00 01	?				Nearby group of galaxies

Target Notes: The targets in run A and the main ones for this exercise; the targets in run B give a more general impression of the types of (large) galaxy we might like to observe with an ELT with the same requirements as run A.
Number of points per field listed with time, I, K mags listed. Angular diameter is the total size of the entire galaxy not always necessary to observe the entire galaxy.

12b. ESO Archive - Are the data requested by this proposal in the ESO Archive (http://archive.eso.org)? If yes, explain why the need for new data.

13. Scheduling requirements

14. Instrument configuration

Period	Instrument	Run ID	Parameter	Value or list
79	FORS2	A	IMG	VIJHK
79	FORS2	B	IMG	VIJHK

EUROPEAN SOUTHERN OBSERVATORY

Organisation Européenne pour des Recherches Astronomiques dans l'Hémisphère Austral
Europäische Organisation für astronomische Forschung in der südlichen Hemisphäre

VISITING ASTRONOMERS SECTION • Karl-Schwarzschild-Straße 2 • D-85748 Garching bei München • e-mail: visas@eso.org • Tel.: +49-89-32 00 64 73

APPLICATION FOR OBSERVING TIME PERIOD: **78A**

Important Notice:

By submitting this proposal, the PI takes full responsibility for the content of the proposal, in particular with regard to the names of COIs and the agreement to act according to the ESO policy and regulations, should observing time be granted

1. Title Category: **B–3**

First Stars relics in the Milky-Way and satellites

2. Abstract

There are two basic ways to probe the very first stellar formation events: look for pop III stars at high redshift, or observe, locally, the nucleosynthetic imprints of these first stars. Searches for the most metal-poor stellar content of the galactic halo are advancing fast and in the coming decade, we expect to have large samples of extremely metal-poor stars identified throughout the Milky-Way halo and the nearby galaxies (Magellanic Clouds and closest dwarf spheroidal galaxies). These stars still display in their atmospheres the imprints of metal-enrichement by pop III stars, and allow to gain insight on the nature and nucleosynthesis of the earliest chemical enrichement processes, provided that the detailed abundance patterns of these stars can be derived (eg Cayrel et al. 2004). This DRM reviews what major steps could be taken by a 42m ELT, examining the wavelength domains needed for various applications.

3.

Run	Period	Instrument	Time	Month	Moon	Seeing	Sky Trans.	Obs.Mode
A	79	UVES	20h	any	d	$\leq 0.8''$	PHO	v

4. Number of nights/hours Telescope(s) Amount of time

a) already awarded to this project:

b) still required to complete this project:

5. Special remarks:

6. Principal Investigator: Vanessa Hill (Observatoire de Paris, F, Vanessa.Hill@obspm.fr)

CoI(s):

7. Is this proposal linked to a PhD thesis preparation? State role of PhD student in this project

8. Description of the proposed programme

A) Scientific Rationale: There are two basic ways to probe the very first stellar formation events: look for pop III stars at high redshift, or observe, locally, the nucleosynthetic imprints of these first stars. This latter approach, often referred to as stellar archeology, has been so far persued in the Milky-Way halo.
Searches for the most metal-poor stellar content of the galactic halo have yielded significant samples of low-mass stars, that have lived close to a Hubble time, and still display in their atmospheres the imprints of metal-enrichement by pop III stars. They allow to gain insight on the nature (masses in particular) and nucleosynthesis of the earliest chemical enrichement processes, provided that the detailed abundance patterns of these stars can be derived (eg Cayrel et al. 2004).
So far, both the extremely metal-poor stars (EMPS) searches, and even more strikingly the high resolution spectroscopic follow up needed to derive accurate elemental abundances, have foccussed on a small volume around the solar neighborhood. Typically, main-sequence stars are confined within ∼XX kpc of the sun while giants can typically reach out to ∼XX kpc. In the coming decade, these sample will be enlarged and deepened (to V∼20), finally reaching to the full galactic halo (out to ∼80-100kpc). These will also includes the most metal-poor stars outside of the Milky-Way galaxy, in nearby satellite galaxies such as the Magellanic Clouds and the most nearby dwarf spheroidal galaxies. With this major step forward, one can hope to adress numerus questions that are today hampered by the lack of statistics and volume confined in the inner halo:

- what is the true lowest-Z boundery for low-mass stars ?
- and its corollary: how does star-formation proceeds close to zero metallicity ?
- what is the role of extremely massive popIII stars in the metal build-up ?
- is the outer halo preferentially populated by EMP stars (as has been suggested by the finding of the first [Fe/H]=-5 star in the HES sample, a category of stars that was absent from more nearby survey)
- does the population of most ancient stars (lowest-metallicity) vary upon the host-galaxy (or DM halo) they reside in (as has been suggested by Helmi et al. 2006) ?
- can the Lithium abundances (^{6}Li and ^{7}Li) observed in dwarf EMP stars be reconciled with Big Bang Nucleosynthesis under the concordance cosmological model ?
- what is the age of the most metal-poor stars (radioactive age-dating using Th and U, eg Cayrel et al. 2001, Hill et al. 2002, Frebel et al. 2007). What is the astrophysical site for the r-process neutron captures ?
- what are the role of cosmic ray spallation in the early galaxy (production of Be and ^{6}Li)

Below is an outline of the near future searches for EMPS that will be in need of high spectral resolution follow-up in the coming decade, together with the number of stars below [Fe/H]< -3 and < -5 expected to be found from these searches.

Survey	Hemisph.	Start	Eff. sky coverage	Eff. mag limit	$N < -3.0$ (EMP)	$N < -5.0$ (HMP)
HES	South	1989	6400 deg^2	$B < 17.0$	400	4
SEGUE	North	2005	1000 deg^2	$B < 19.0$	1000	10
LAMOST	North	2007	10,000 deg^2	$B < 19.0$	10,000	100
SSS	South	2007	20,000 deg^2	$B < 18.0$	5000	50

The estimated numbers of stars expected to be discovered in ongoing and near-future surveys for metal-poor stars, extrapolating the current results from the HES. [Courtesy N. Christlieb.]
By the time of the operation of an ELT, searches for EMPS will have been further extended to yet larger volumes and depth, reaching B=20-21. This is a very significant step, since it means that they will have *reached significant samples outside of the Milky-Way*, in the closest small galaxies: LMC, SMC and a few dSph galaxies. This will provide with the first strong test on the similarities / differences of the metallicity distribution function of the oldest stars in large (Milky-Way) and small (MC and dSph) galactic systems, a *direct constraint to galaxy formation and assembly scenarios at the earliest times.*

B) Immediate Objective:
To perform these detailed analyses of the genesis of the most ancient populations in the Milky-Way and nearby systems, one needs to perform high-resolution spectroscopy to measure accurately the elemental abundance patterns in stars of magnitudes down to B=19 to 20.
Because EMPS are very metal-poor, many atomic and molecular transitions in the V, R and IR bands are exeedingly weak, and the UV becomes a very important wavelength domain where instrisically strong lines are still detectable. Below is a short list of unique oportinities to be gained, at a resolution R 40000-70000 (single object, no requirement on spatial resol) for the study of EMPS with metallicities < -3 and down to < -5.
1) in the blue (370nm-520nm):
Most of the elements in very metal-poor stars can only be measured in the blue. This includes:

- numerous lines of Fe (two ionisation stages to constrain the stellar gravity -or luminosity and hence distance, aswell as microturbulent velocity)
- α elements (Mg, Ca, Ti) iron peak elements (Mn, Co, Ni), Zn, that allow to constrain the nucleosynthetic origin of the metal-enrichment and hence the nature and metal-production of pop III stars.

8. Description of the proposed programme (continued)

• heavy neutron capture elements (r-process elements). The radioactive elements Th and U that can be used for age-dating the stars fall in that category, whith theit most proeminent lines at 401.9nm and 386nm respectively (each extremely weak, Thλ401.9 can reach some 30mÅwhile Uλ385.96 reaches at most a few mÅ). This make radioactive age-dating will become possible for almost all the very metal-poor stars with r-process enhancements that will be detected by LAMOST for example (several hundred candidates expected) provided B=18 can be reached.

2) in the UV (310nm-370nm):

• Berylium (310nm) probes cosmic-ray spallation in the earliest phases of galaxy evolution. Reaching B=15-15.5 will turn the current handfull of stars ([Fe/H]<-2.5) where Be is measurable into a whole population (some 100 turnoff stars are currently known in the galactic halo with [Fe/H]<-3 at that magnitude limit)

• the r-process is still at loss of a production site, and, because of its rare nature, a very good probe of mixing of SNe products in the early galactic gaz. Many r-process elements can only be detected bluer than 3700A, most notably among the heaviest and lightest (Nb, Mo, Ru, Ag and Ir, Pt, Bi)

• Oxygen has other transitions (in V, R and H bands) but all of them become undetectable at the lowest metallicities, leaving the OH bands in the UV (310nm) the *only* probe of oxygen abundance in these stars. Oxygen is the most abundant element after He, and yeilds very important clues on the nature of popIII stars since it's production is mass-dependant

• similarly, Nitrogen is only detectable at the lowest metallicities from the NH bands around 336nm. Nitrogen is one of the current major puzzles for early nucleosynthesis (secondary element that should therefore be underabundant in metal-poor stars but instead shows a primary behaviour, and is even enhanced in the reccord-holders -5 metallicity stars).

3) The case of Li (670.7nm)

Lithium in old stars can give constraints on Big-Bang Nucleosynthesis and $\Omega_{\rm baryon}$. So far, its abundance has only been measured in the Milky Way. The ELT will give for the first time the opportunity to measure it in another galaxy (other DM halo). It is detected by its 670.7nm doublet, and will be measurable in the Sagittarius dwarf galaxy dwarf (turnoff) stars. (LMC will still be too faint, even for a 42m ELT).

C) Telescope Justification: We need the ELT because of the photon collecting power. Similar observations could be done with EXPRESSO in its 4-UT combined mode, **provided it goes all the way to UV** which is technically very difficult given the light-path to the common focus. Even in the (unlikely) event that the combined 4-UT mode reaches a 16m equivalent telescope, the limiting magnitudes that would be reachable there would be too shallow to reach the outer galactic halo stars and the nearest satellite galaxies (LMC, SMC, Sgr, Scl, Draco..).

D) Observing Mode Justification (visitor or service): visitor mode is preferred as we just love to sit in tin boxes and would like to see the Eiffeltower on its side.

E) Strategy for Data Reduction and Analysis: In parallel to the increasing accuracy of the observations of spectra, the corresponding stellar atmosphere modelling is also expected to progress significantly in the coming decade. Most notably, radiative transfer in 3D hydrodynamical stellar atmospheres is now starting to emerge, and will become the norm in the coming decade(s). Similarly, NLTE modelling is becoming feasible for numerou atoms. This has a strong impact especially in metal-poor stars where the UV flux is high and collisions are few (and LTE more easily violated).

8. Attachments (Figures)

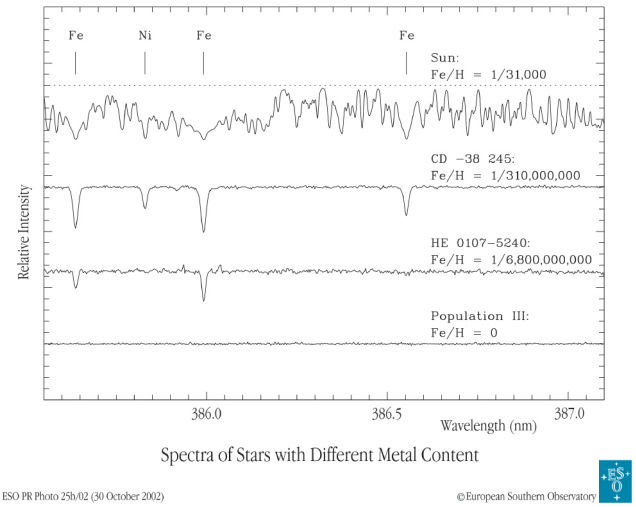

Fig. 1: Metal-lines in stars of various metallicities, in the blue wavelength. The stars targetted here are of [Fe/H]<-3 to -4. In the red wavelength domain, only very few and exceedingly weak lines persist at these metallicities.

9. Justification of requested observing time and lunar phase

Lunar Phase Justification: -

Time Justification: (including seeing overhead) The requirement for such a science case is a visual (and UV) single-object spectrograph with resolution R= 20000 - 70000 (around 40000-50000 being the best compromise), with no particular requirement on the spatial resolution (seeing limited is fine).
Using the ELT- Experimental ETC (Version 2.7WG) (integrated over 0.8" in a seeing limited observation with seeing=0.8"). These numbers wer checked and refined (at various wavelength) scaling up the current performances of VLT+UVES to a 42m telescope (1247m^2 vs 45.24m^2 collecting area):
1) Blue and UV
• Detailed abundance analysis of extremely metal-poor giants down to B=20 (i.e. reaching down to the level of the horizontal branch in the LMC): S/N=80-100 per resolution element is achieved in 2h in the B band at R=50000
• Uranium detection (& r-process elements) S/N=200 @386nm at R=70000 in 9h for a B=18 magnitude [Fe/H]=-3 giant.
2) Extreme UV • Nitrogen detection S/N=40 @336nm at R=50000 (or S/N=70 336nm at R=20000, since the bandhead is wide) in 3h for a B=19 [Fe/H]=-4 lower-RGB giant or turnoff star.
• Oxygen detection R=50000 S/N=50 @310nm in 10h for a B=17.5 [Fe/H]=-4 giant.
• Berylium detection R=50000, S/N=150 @310nm in 10h for a B=15.5 turnoff star at [Fe/H]<-3.
3) Red • Lithium detection S/N=60 @670.7nm in 7h for a V=21.5 turnoff star in the Sgr Galaxy.

Calibration Request: Standard Calibration

10. Report on the use of ESO facilities during the last 2 years

11. Applicant's publications related to the subject of this application during the last 2 years

12. List of targets proposed in this programme

Run	Target/Field	α(J2000)	δ(J2000)	ToT	Mag.	Diam.	Additional info	Reference star
A	name	RA	DEC	time(hrs)	mag	DM	ang diam(')	note

12b. ESO Archive - Are the data requested by this proposal in the ESO Archive (http://archive.eso.org)? If yes, explain why the need for new data.

13. Scheduling requirements

14. Instrument configuration

Period	Instrument	Run ID	Parameter	Value or list
79	UVES	A		

EUROPEAN SOUTHERN OBSERVATORY

Organisation Européenne pour des Recherches Astronomiques dans l'Hémisphère Austral
Europäische Organisation für astronomische Forschung in der südlichen Hemisphäre

VISITING ASTRONOMERS SECTION • Karl-Schwarzschild-Straße 2 • D-85748 Garching bei München • e-mail: visas@eso.org • Tel.: +49-89-32 00 64 73

APPLICATION FOR OBSERVING TIME PERIOD: **78A**

Important Notice:

By submitting this proposal, the PI takes full responsibility for the content of the proposal, in particular with regard to the names of COIs and the agreement to act according to the ESO policy and regulations, should observing time be granted

1. Title Category: **B–2**

A Survey of Black Holes in Different Environments

2. Abstract

We propose to carry out a spatially resolved spectroscopic survey of the centers of elliptical and early type galaxies with very low and very high central velocity dispersions. The goal of the survey is to resolve the Sphere of Influence (SoI) of the suspected black holes (BHs) in those galaxies and thereby investigate the low and hi-mass end of the M_{BH}-σ relation. Our sample includes galaxies in a broad range of clusters as well as field galaxies. The centers of all galaxies in our sample have previously been imaged by HST and/or JWST, and expected BH masses have been estimated based on the cusp brightness and velocity dispersion. The survey will for the first time resolve the SoI of BH with masses M_{BH} of around 10^6 outside the local Universe, and at the same time yield spectroscopic data of a significant sample of the most massive BHs currently known. The survey also might discover supermassive BHs with masses of up to about 10^{10} $M_\odot$.

3.

Run	Period	Instrument	Time	Month	Moon	Seeing	Sky Trans.	Obs.Mode
A	79	FORS2	200h	any	d	$\leq 0.4''$	CLR	v
B	79	FORS2	3h	any	d	$\leq 0.4''$	CLR	v

4. Number of nights/hours Telescope(s) Amount of time

a) already awarded to this project:

b) still required to complete this project:

5. Special remarks:

In order to establish feasibility we need detailed simulations which included the expected s/n.

6. Principal Investigator: Wolfram Freudling (ESO, D, wfreudli@eso.org)

CoI(s): Eric Ensellem (U. Lyon, F), Alessandro Marconi (Arcetri, I), The Rest (Elsewhere, ESO)

7. Is this proposal linked to a PhD thesis preparation? State role of PhD student in this project

8. Description of the proposed programme

A) Scientific Rationale:

The relationship between host galaxy bulge mass and black hole (BH) mass is well established for both active and currently inactive galaxies (McLure & Dunlop, 2002, MNRAS, 311, 795). Wyithe (2006, MNRAS 365, 1082) and Greene & Ho (2006, ApJ 641, L21) have recently argued that the linear M_{BH}-σ relation steepens at high black hole masses and flattens at low black hole masses. Lauer et al. (2006, ApJ, astro-ph/0606739) argue that cusp brightness might in fact be a better estimator of the black hole masses than σ for the most massive BHs with $M_{BH} > 10^9 M_\odot$. If confirmed, this might explain why it has been so difficult to find the supermassive black holes in the local Universe, which are expected to exist as the counter parts of the high-z QSOs. It is currently not known whether the M_{BH}-σ relation depends on environment or if and how it evolves.

To study the M_{BH}-σ relation at its extreme ends, it is necessary to directly determine the BH masses for statistical significant samples. To confidently detect and measure the mass of a nuclear BH, we need to probe the volume within which the BH dominated the galactic dynamics. Called the 'Sphere of Influence', this region has a radius defined as:

$$r_i = GM_{BH}/\sigma^2 = 4.3pc(M_{BH}/10^7 M_\odot)/(\sigma/100km/s)^2 \quad (1)$$

where σ is the stellar velocity dispersion. A typical scale for BH masses of $\approx 10^7 M_\odot$is about 7 pc.

Unfortunately, it is quite difficult to probe the SoI in galaxies at the extreme ends of the M_{BH}-σ relation. There are currently only two cases where this region has been probed directly to show that a massive BH is the only physical possibility: our Galaxy and NGC4258. The projected diameter of the SoI for a BH with masses of $\approx 10^7 M_\odot$ is significantly smaller than 100mas at the distance of the Virgo cluster. Because of the low volume density of high-mass BHs, studying the high-mass end of the M_{BH}-σ relation requires samples at moderately high redshifts out to $z \approx 0.4$. At such distance, the SoI is again smaller than 100mas. In addition, the onset of significant surface brightness dimming at such redshifts makes it impossible to obtain spatially resolved spectroscopy of the cores of such galaxies with 8m class telescopes. Therefore, progress in this field will be moderate before the arrival of the ELT.

However, the high angular resolution and sensitivity of an ELT will allow:

1. to resolve nuclear sub-structures down to a few pc at distances of tens of Mpc (depending on aperture and PSF). This will allow mass determination of BHs with masses similar to the one in the center of the Milky Way out to the distance of Virgo.
2. resolve the sphere of influence for the most massive BHs with masses of greater than $10^9 M_\odot$ at cosmological redshifts. Mass determination of black holes will be limited by the available light only. Mass determination for $10^9 M_\odot$ BHs will be possible out to a redshift of about 0.4, allowing the collection of statistical samples of such objects.

Such measurements are fundamental to the understanding of the relationship between the evolution of the BH and the host galaxy, including the possible connection between AGN and starburst activity.

A rough estimate of the maximum redshift at which BHs can be spectroscopically resolved is shown in Fig. 1. For each BH mass, the size of the SoI was computed. Subsequently, the maximum distance at which the projected diameter of the SoI is more than 10 mas was determined. An additional complication is that the most massive BHs tend to be in galaxies with relatively low surface brightness in the center. This effect was taken into account by assuming that the central surface brightness is related to the BH mass as given in Equ. 10 of Lauer et al. (2006). It can be seen that accurate mass determination will be limited to redshifts less than $z \approx 0.6$ even for the most massive BHs.

B) Immediate Objective:

We propose to carry out 2 distinct spatially resolved spectroscopic surveys of the nucleus of early type galaxies. The first survey targets galaxies with estimated BH masses of more $10^{9.5} M_\odot$. Candidate galaxies have been extracted from the SDSS, and are at redshifts up to $z \approx 0.4$.

The second survey targets BH with masses less than $10^{6.5} M_\odot$ in a variety of environments. This includes galaxies in the Virgo Cluster and galaxies at similar redshift.

C) Telescope Justification:

The program needs the ELT both for the spatial resolution and the light collection power. The most efficient way to map the full velocity field around the nucleus is to use an IFU. If an IFU is not available, several long slit positions could be used.

In order to determine the feasibility to detect BHs with spatially resolved spectroscopy, we have simulated the shape of spectral lines based on a kinematic model of NGC 3377.

We first fit a multi-Gaussian expansion to the density and velocity fields of that galaxy (see Emsellem et al., 1994, AA 285, 739). We then projected the density and velocity field to a given redshift, convolved the fields with a I-band LTAO PSF (provided by Jochen Liske), and finally dispersed the light. To illustrate the results, we present in Fig. 2 a single slice through such a position-velocity cube perpendicular to the rotation axis. For real observations, the optimal position angle of such a long-slit spectrum is not know a priory, but can be determined from IFU observations or alternatively require several slit positions.

8. Description of the proposed programme (continued)

The simulated spectrum in Fig. 2 shows that the proposed project requires very high s/n and Nyquist sampling of the PSF. A more rigorous simulation should include a realistic galaxy spectrum and noise estimate. Such a simulated spectrum could then be used with existing programs to estimate the BH mass.

D) Observing Mode Justification (visitor or service): Would like to visit the ELT.

E) Strategy for Data Reduction and Analysis:

8. Attachments (Figures)

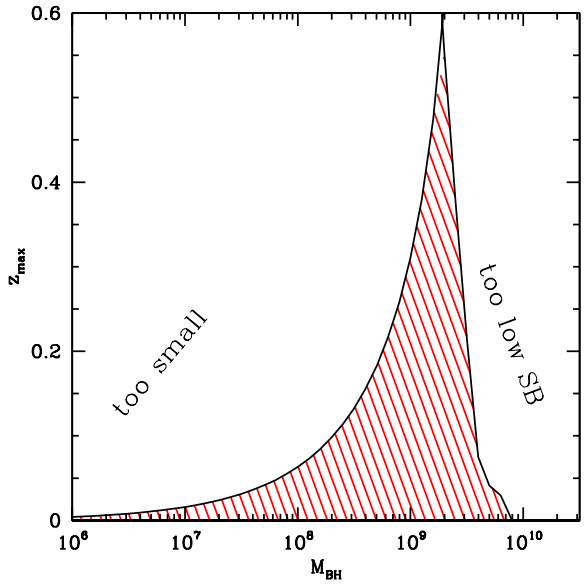

Fig. 1: Maximum observable distance of black holes.

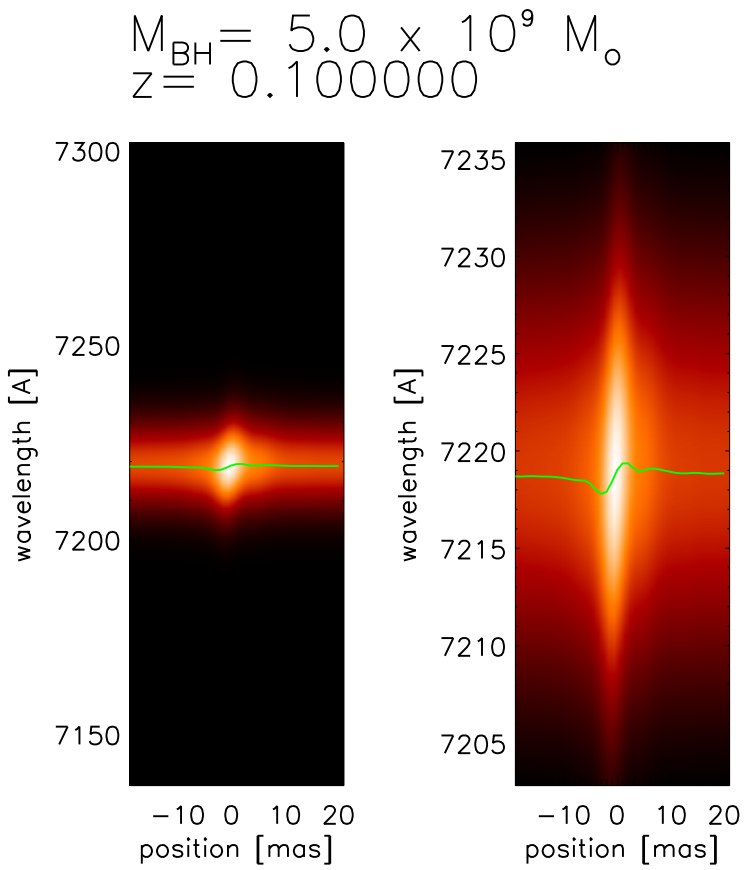

Fig. 2: Simulation of the imprint of a BH on one spectral line at z=0.1. Both panels show the same spectral line at a different wavelength scale. The green line is the center of mass of the spectral line determined for different position along a slit.

9. Justification of requested observing time and lunar phase

Lunar Phase Justification: Grey time would probably be acceptable.

Time Justification: (including seeing overhead) We estimated the exposure times using the ELT ETC, assuming a necessary S/N of 30 **per 5 mas pixel** and a spectral resolution of 5000. We believe such a high s/n will be necessary to accurately determine the center of the broad spectral line as a function of position (see Fig. 2.). In box 15, we list for each target the exposure time for a single spectrum. This corresponds to the time requirement with an IFU. The total exposure time for the hi-z sample is about 200h. For the low-z sample, the actual exposure time for each target is small, the total for the full sample is about 3h. If a longslit spectrograph is used instead the required observing time would triple assuming three position angles on each target.

Calibration Request: Standard Calibration

10. Report on the use of ESO facilities during the last 2 years

11. Applicant's publications related to the subject of this application during the last 2 years

12. List of targets proposed in this programme

Run	Target/Field	α(J2000)	δ(J2000)	ToT	Mag.	Diam.	Additional info	Reference star
A	J03081586-234128	030815	-234129	1.0		76	z=0.07	
A	J09481931+00471	094819	+004717	4.4		49	z=0.21	
A	J112604.67-00280	112604	-002809	4.6		35	z=0.34	
A	J122927.38+00584	122927	+005849	5.9		41	z=0.26	
A	J154647.81-00321	154647	-003219	5.7		42	z=0.25	
A	J01415759-01062	014157	-010626	2.5		61	z=0.16	
A	J02114830+00164	021148	+001640	4.1		36	z=0.30	
A	J01035416+14481	010354	+144814	5.1		46	z=0.23	
A	J040316.96-05162	040316	-051622	9.8		37	z=0.30	
A	J091254.53-00092	091254	-000926	5.2		37	z=0.30	
A	J10383663+01174	103836	+011749	2.3		73	z=0.13	
A	J11463450+02214	114634	+022147	4.4		50	z=0.19	
A	J121430.79+01221	121430	+012212	4.2		52	z=0.19	
A	J13315353+03175	133153	+031750	2.9		56	z=0.18	
A	J08533336+02433	085333	+024334	4.7		41	z=0.27	
A	J10494034+05030	104940	+050306	4.8		34	z=0.31	
A	J15322891+02391	153228	+023916	3.1		73	z=0.13	
A	J23233141-10255	232331	-102551	4.6		37	z=0.29	
A	J002750.60-10052	002750	-100524	7.1		30	z=0.40	
A	J002627.47-09260	002627	-092602	7.2		36	z=0.32	
A	J00413973-10544	004139	-105449	5.4		36	z=0.32	
A	J004450.76-09415	004450	-094158	6.1		30	z=0.39	
A	J210656.09+09373	210656	+093734	5.3		35	z=0.34	
A	J13012184+05201	130121	+052016	5.6		46	z=0.23	
A	J13281513+05521	132815	+055210	8.7		35	z=0.32	
A	J134613.21-02160	134613	-021600	5.7		34	z=0.33	
A	J21000912-00324	210009	-003249	5.7		42	z=0.25	
A	J09183893+05280	091838	+052803	2.4		55	z=0.18	
A	J100548.18+06424	100548	+064241	3.2		45	z=0.23	
A	J10421721+07153	104217	+071539	4.8		36	z=0.30	
A	J12022628+10334	120226	+103344	4.2		46	z=0.23	
A	J095327.04+10493	095327	+104933	6.4		37	z=0.31	
A	J10000065+09204	100000	+092044	5.7		35	z=0.32	
A	J112712.60+11563	112712	+115637	7.4		40	z=0.26	
A	J131746.33+12453	131746	+124536	9.1		31	z=0.40	

Following targets moved at the end of the document ...

Target Notes: 1. Run A is for the high-mass BHs, run B for the low-mass BHs. 2. The given diameter the expected one for the sphere of influence in mas.

12b. ESO Archive - Are the data requested by this proposal in the ESO Archive (http://archive.eso.org)? If yes, explain why the need for new data.

13. Scheduling requirements

14. Instrument configuration

Period	Instrument	Run ID	Parameter	Value or list
79	FORS2	A	SPEC	R

15. List of targets proposed in this programme

Run	Target/Field	α(J2000)	δ(J2000)	ToT	Mag.	Diam.	Additional info	Reference star
	...continuing from box 3							
A	J13313548+12503	133135	+125032	4.0		43	z=0.25	
A	J155209.36+07475	155209	+074751	4.9		35	z=0.32	
A	J093904.29+11130	093904	+111306	6.1		32	z=0.38	
A	J093653.30+10395	093653	+103958	9.6		33	z=0.36	
A	J09343336+10431	093433	+104312	0.7		77	z=0.12	
A	J100446.38+12200	100446	+122003	3.6		38	z=0.27	
A	J08083058+07201	080830	+072011	0.6		91	z=0.10	
A	J121859.87+14251	121859	+142516	4.6		41	z=0.27	
B	NGC0516	01 24 08.2	+09 33 05	0.1		12	z= 2456 km/s	
B	ESO545G040	02 38 11.3	-20 10 01	0.1		48	z= 1474 km/s	
B	IC1919	03 26 02.0	-32 53 45	0.1		19	z= 1327 km/s	
B	ESO358G006	03 27 17.6	-34 31 37	0.1		23	z= 1254 km/s	
B	ESO548G033	03 32 28.6	-18 56 53	0.1		43	z= 1652 km/s	
B	NGC1373	03 34 59.1	-35 10 15	0.1		35	z= 1383 km/s	
B	NGC1369	03 36 45.0	-36 15 22	0.1		31	z= 1414 km/s	
B	NGC1390	03 37 52.0	-19 00 32	0.1		20	z= 1211 km/s	
B	MCG-06-09-023	03 42 45.5	-33 55 13	0.1		25	z= 1268 km/s	
B	ESO358G059	03 45 03.5	-35 58 22	0.1		32	z= 1042 km/s	
B	ESO250G005	04 04 35.1	-46 02 35	0.1		34	z= 1230 km/s	
B	NGC2328	07 02 35.8	-42 04 07	0.1		25	z= 1187 km/s	
B	UGC05467	10 08 12.8	+18 42 25	0.1		24	z= 2883 km/s	
B	NGC3457	10 54 48.5	+17 37 13	0.1		59	z= 1161 km/s	
B	ESO440G038	12 01 42.5	-31 42 12	0.1		21	z= 2320 km/s	
B	NGC4415	12 26 40.5	+08 26 09	0.1		34	z= 933 km/s	
B	NGC4467	12 29 30.4	+07 59 38	0.1		33	z= 1423 km/s	
B	NGC4587	12 38 35.5	+02 39 25	0.1		49	z= 913 km/s	
B	NGC4612	12 41 32.6	+07 18 52	0.1		28	z= 1781 km/s	
B	NGC4733	12 51 06.9	+10 54 45	0.1		50	z= 929 km/s	
B	NGC5206	13 33 43.9	-48 09 08	0.1		42	z= 555 km/s	
B	NGC5666	14 33 09.3	+10 30 38	0.1		32	z= 2224 km/s	
B	IC4653	17 27 07.1	-60 52 50	0.1		23	z= 1551 km/s	
B	ESO286G050	21 06 41.0	-42 33 26	0.1		20	z= 2672 km/s	
B	NGC7077	21 29 59.6	+02 24 50	0.1		39	z= 1166 km/s	
B	ESO466G046	22 02 44.1	-31 59 26	0.1		22	z= 2326 km/s	
B	NGC7351	22 41 26.8	-04 26 40	0.1		45	z= 890 km/s	
B	IC5267B	22 56 57.1	-43 45 35	0.1		15	z= 1758 km/s	

European Organisation for Astronomical Research in the Southern Hemisphere

Organisation Européenne pour des Recherches Astronomiques dans l'Hémisphère Austral
Europäische Organisation für astronomische Forschung in der südlichen Hemisphäre

VISITING ASTRONOMERS DEPARTMENT • Karl-Schwarzschild-Straße 2 • D-85748 Garching bei München • e-mail: visas@eso.org • Tel. : +49-89-32 00 64 73

Important Notice:

By submitting this proposal, the PI takes full responsibility for the content of the proposal, in particular with regard to the names of CoIs and the agreement to act according to the ESO policy and regulations, should observing time be granted

1. Title — Category: **X–0**

The Physics and Mass Assembly of Galaxies out to $z \sim 6$

2. Abstract

We propose to obtain ELT spatially resolved spectroscopy of a sample of a thousand massive galaxies at $2 < z < 6$, selected from future large area optical-nearIR surveys. These observations will yield direct kinematics of stars and gas in the first generation of massive galaxies (in the range $0.1 < M < 5 \times 10^{11} M_{\odot}$), as well as their stellar population properties. One will be able to derive dynamical masses, ages, metallicities, star-formation rates, dust exinction maps, to investigate the presence of disk and spheroidal components and the importance of dynamical processes (e.g. merging, in/outflows) which govern galaxy evolution. These data will also allow one to study the onset of well known scaling relations at low redshifts, and to witness the gradual shift of star formation from the most massive galaxies in the highest density regions to less massive galaxies in the field. The whole program is designed to provide the ultimate test of galaxy formation theories.

3.

Run	Instrument	Time	Month	Moon	Seeing	Sky Trans.	Obs.Mode
A	Multi-IFU	30n	jun	n	$\leq 0.8''$	CLR	v

4. Principal Investigator: Piero Rosati (ESO, ESO, prosati@eso.org)

CoI(s): Mathieu Puech (ESO, ESO), Andrea Cimatti (Bologna, I), Sune Toft (ESO, ESO), Marijn Franx (Leiden, NL), Jacqueline Bergeron (IAP, F)

5. Is this proposal linked to a PhD thesis preparation? State role of PhD student in this project

Yes / N/A / mid-course

6. Description of the proposed programme

A) Scientific Rationale: Over the last decade, the synergy of 8-10m class telescopes with HST has strongly reinvigorated the field of galaxy formation and evolution by unveiling very distant galaxies ($z \sim 6$), by allowing the first determination of the global star formation history since redshift $z \sim 6$ and by providing the first insights on the stellar mass assembly history out to $z \sim 3$. Despite this recent progress, the outstanding question remains on *how and when* galaxies assembled their *baryonic* mass across cosmic time. The ΛCDM standard model has provided a satisfactory scenario describing the hierarchical assembly of dark matter halos, in a bottom-up sequence which is now well-established over the whole mass structure spectrum. In contrast, little progress has been made in the physical understanding of the formation and evolution of the baryonic component because the conversion of baryons into stars is a complex, poorly understood process. As a result, all intellectual advances in galaxy formation and evolution over the last decade have been essentially empirical, often based on phenomenological (or "semi-analytical") models which heavily rely on observations to describe, with simplistic rules, such processes as star formation efficiency, energy feedback from star formation and AGN, chemical evolution, angular momentum transfer in merging, etc. Cornerstones observations in this empirical framework are the (total and stellar) mass of galaxies and their physical properties, including age and metallicities of their underlying stellar populations, dust extinction, SF rate, structural/morphological parameters. The study of well-established scaling relations involving a number of these physical parameters (e.g. mass-metallicity, fundamental plane, color-magnitude, morphology-density) are essential for understanding the physical processes driving galaxy evolution. However, with the current generation of 10m-class telescopes, we have been able to construct for example the fundamental plane of early-type galaxies, or to measure the Tully-Fisher relation of late types over a wide range of masses only at low redshifts, whereas only the brightest or most massive galaxies have been accessible at $z > 1$, and a direct measuremnt of masses has been almost completely out of reach at $z > 2$. Thus, our ability to explore the evolution and origin of the aforementioned scaling relations has rapidly reached the limit of 10m-class telescopes. As a result, most of the outstanding questions arisen from recent observational galaxy evolution studies which have pushed the VLT to its limit call for an ELT, specifically to extend the spectroscopic limit by at least two magnitudes with near-diffraction limit angular resolution. HST/Nicmos imaging studies have shown that massive galaxies with disk-like morphologies exist at least out to $z \sim 2.5$ (see Fig.3). IFU spectrographs on 9-10m class telescopes (e.g., VLT/SINFONI and Keck/Osiris) are currently able to distinguish merging from rotating disks in star-forming (SFR$= 100 - 200$ M$_\odot$/yr) massive galaxies ($v \gtrsim 200$ km/s, or $\sim 10^{11} M_\odot$) out to $z \approx 2$ (Genzel et al. 2006; Förster Schreiber et al. 2006; Wright et al. 2007, Shapiro et al. 2008). Spatially resolved metallicity maps have been attempted only for a few very bright late-type galaxies at $z \sim 2$ (Förster Schreiber et al. 2008, in prep.). In addition, observations of bright Ly-break galaxies at $z \sim 3$ and their sorrounding IGM have revealed very large and ubiquitous outflows which play a crucial role for the energy feedback and metal enrichment of the IGM (Steidel et al. 2007). To date, these studies have been confined only to the brightest and most massive galaxies at $z \lesssim 2.5$, thus providing a limited and biased view on galaxy formation.
An ELT, equipped with an optical-nearIR AO-assisted MOS, will remove these limitations, by directly probing physical properties of galaxies as a function of their masses and environment over 90% of the age of the Universe. Specifically, we will extend current studies of massive galaxies to $z \sim 5$ and extend the accessible mass range by more than a factor of 10 at $z \sim 2$. This will provide crucial information on the baryon physics in galaxy evolution models and will drive the transition from current phenomenological models to a physical understanding of galaxy formation and evolution.

B) Immediate Objective: The proposed experiment consists of spatially resolved ELT spectroscopic observations of a sample of ~ 1000 galaxies over the redshift range $2 < z < 6$. Specific requirements on the telescope/instrument system to reach the scientific objective discussed here are discussed in the following section. Specific science goals and measurements for this project include:
– High spatial resolution spectroscopy will provide direct kinematics of stars and gas in the first generation of massive galaxies in the range ($0.1 < M_{star} < 5 \times 10^{11} M_\odot$). The resolution will be of the order of FWHM$= 0.01$ arcsec in the H band, corresponding to 80pc at $z = 3$. This will allow one to assess the relative importance of different dynamical processes (e.g. merging, in/out-flows) in governing galaxy evolution at different cosmic times. Thus, it will be possible to quantify the role, physics and rate of merging in modulating star formation, mass assembly and morphology evolution.
– The evolution of the galaxy mass function will be traced by measuring dynamical masses (accurately derived from kinematic maps) of the first massive galaxies, i.e. 10 times more massive of the Milky Way ($M \sim 10 M_{MW}$), at virtually any redshift and as well as lower mass galaxies ($\sim 0.1 M_{MW}$) over 90% of the age of the Universe. This will require moderate-to-high resoution ($R > 1000$) integral field or multi-object spectroscopy to measure rotation curves from emission lines in disk-like systems and velocity dispersions from absorption-lines in early-type galaxies.
– The cosmic epochs over which the mass assembly of spheroids and disks occured will be determined. Current observations suggest that the build-up of spheroids was largely over by $z \sim 1$, just when the major build-up of disks was ramping up to dominate later on.

– Dust extinction is currently very difficult to estimate because the Balmer lines are redshifted in the near-IR, where the current sensitivity is often not enough to detect lines such as Hβ. With ELT spectroscopy, it will be possible to derive dust extinction maps from Balmer line ratios. Such a knowledge is essential to reliably derive physical parameters and to break the complex degeneracies due to dust extinction.
– The star formation rate (SFR) will be derived from the extinction-corrected emission line luminosity (e.g. Hα to $z \sim 2.5$, [OII] to $z \sim 5$), and compared with other indicators coming from multi-wavelength observations (e.g. ALMA, Herschel, JWST). The knowledge of the dynamical masses will constrain the specific star formation rates (SSFR=SFR/Mass) and the mass growth history in galaxies.
– The detailed properties of the stellar continuum, absorption and emission lines will be used to derive the age, metallicity and star formation history of the stellar population and to break the degeneracies between these quantities that affect the interpretation of the current results.
– The relative velocity of emission and absorption lines of different species will place important constraints of the ISM physical state and dynamics. This will also allow to identify the sources of feedback processes (energy injection into ISM, superwinds, outflows, AGN) and the origin of the SF quenching in massive halos and its possible link with the onset of the massive black hole - bulge relation and galaxy/AGN co-evolution.
– Current determination of the stellar mass function and its evolution is based on the model dependent method of "photometic masses" from multi-wavelength surveys, whose reliability remains largely untested. The ELT will provide crucial calibration of this method by allowing masses to be determined, via dynamical methods or strong gravitational lensing, for galaxies over the widest range of redshifts and morphological type.
– The transition of massive star formation from high-density environments to the field will be observed (current evidence tentatively places the epoch of such transition at $z > \ 4 - 5$).
– The migration of star formation rate from high to low masses as galaxies evolve ("downsizing effect") will be studied, by measuring M/L ratios of distant galaxies over a wide range of masses (only the most massive galaxies can be probed today out to $z \sim 1$).
– The evolution of velocity distribution function of satellites of nearby galaxies down to ~ 30 km/s (Loeb & Peebles 2003) will be observed as a direct probe of DM halo properties and as a test of CDM predictions.

Sample Selection and Synergies with other facilities:
Most of the issues to be addressed with the ELT in the field of galaxy evolution will benefit greatly from the synergy with JWST, which will provide rest-frame optical imaging of galaxies with superp spatial resolution over the widest redshift range, as well as low resolution integrated spectroscopy. These space capabilities are likely to remain unmatched even in the era of ELT with LTAO. The galaxy samples needed for this science case will likely be drawn from the next-generation large field surveys in the optical (e.g. LSST and Pan-STARRS), near-IR (VISTA), and or sub-mm surveys with ALMA. Several mass selected samples of galaxies out to $z \sim 6$ will be available over the next decade, however the ELT observations proposed here will likely result into a flux-limited sample.

References:
Genzel R. et al., 2006, Nature, 442, 786
Grazian A. et al., 2006, A&A, 449, 951
Förster Schreiber N. M. et al., 2006, ApJ, 645, 1062
Loeb A. & P.J.E. Peebles 2003, Apj, 589, 29
Shapiro K.L et al. 2008, Ap.J, in press
Steidel C.C. et al. 2007, in proc. of "Cooling, Star Formation, and Feedback in High Redshift Galaxies", Gaching bei München
Toft S. et al., 2007, ApJ, 671, 285
Wright S.A. et al., 2007, ApJ, 658, 78

C) Telescope Justification:

Technical requirements for the telescope/instrument systems are derived as follows.
A sample of at least 1000 galaxies at $z < 2 < 5$ is required to properly sample galaxy diversity over the mass, redshift and "morphological type" (mergers, disks, spheroids). If we take a redshift range of $z = 1.5 - 5.6$, current observations suggest a K-band magnitude for L^* galaxies in this redshift range of $K^*_{AB} = 22 - 25$, with corresponding surface densities of $\Sigma(\mathrm{gal/arcmin}^2) \approx 10, 1, \sim 0.1$, for $L > L^*$, at $z = 2, 4, 5$. Across the probed redshift range, the minimum mass accessible will vary from 0.1 to a few times M^*.
The sample size and surface densities imply a field of view of 20-100 arcmin^2 and a multiplex factor of at least 10, possibly up to 100. Spectral resolutions $R \sim 3 - 5000$ are required for adequate OH removal and detection of kinematic gradients. As a result, either a deployable IFU system with MOAO or a large monolithic IFU (LTAO) should be adequate for this project.
Required spatial sampling is of ~ 50 mas to resolve galaxies in at least 10 spatial bins given their effective radii of $R_E = [0.1 - 0.3]$ arcsec. This implies that AO performance are critical for this experiment, requiring an LTAO system, eventually MOAO.

6. Description of the proposed programme (continued)

Mapping physical properties (SFR, metallicity, dust) and dynamics requires mostly near-IR coverage, however to cover standard diagnostic lines, [OII]3727, [OIII]5007, Hβ, Hα, over the entire redshift range [1-5], one would need a wavelength coverage of $0.8 - 2.5\mu$, with trade-offs due to varying AO performance across this range.
Typical exposure times (see Sect.7 below) for M^* galaxies out to $z \sim 5$ are in the range 10-50 hours, making the whole program feasible in roughly 20-50 nights.
A first analysis of these requirements based on scaling arguments shows that no break points are apparent in achieving science goals, when considering a telescope diameter in the range of 30-42 meters. There is however a clear advantage for large apertures if one wants to access the highest redshift of galaxy assembly and an adequate range of galaxy masses (i.e. $0.1\,M^* < M <$ a few M^*) over $2 < z < 5$.
Requirements and main parameters of targets are summarised in Table 1. below.

D) Strategy for Data Reduction and Analysis: The Puech et al. reduction pipeline used for the simulations described below (see Fig.2) will be well suited to reduce these data.

6. Attachments (Figures)

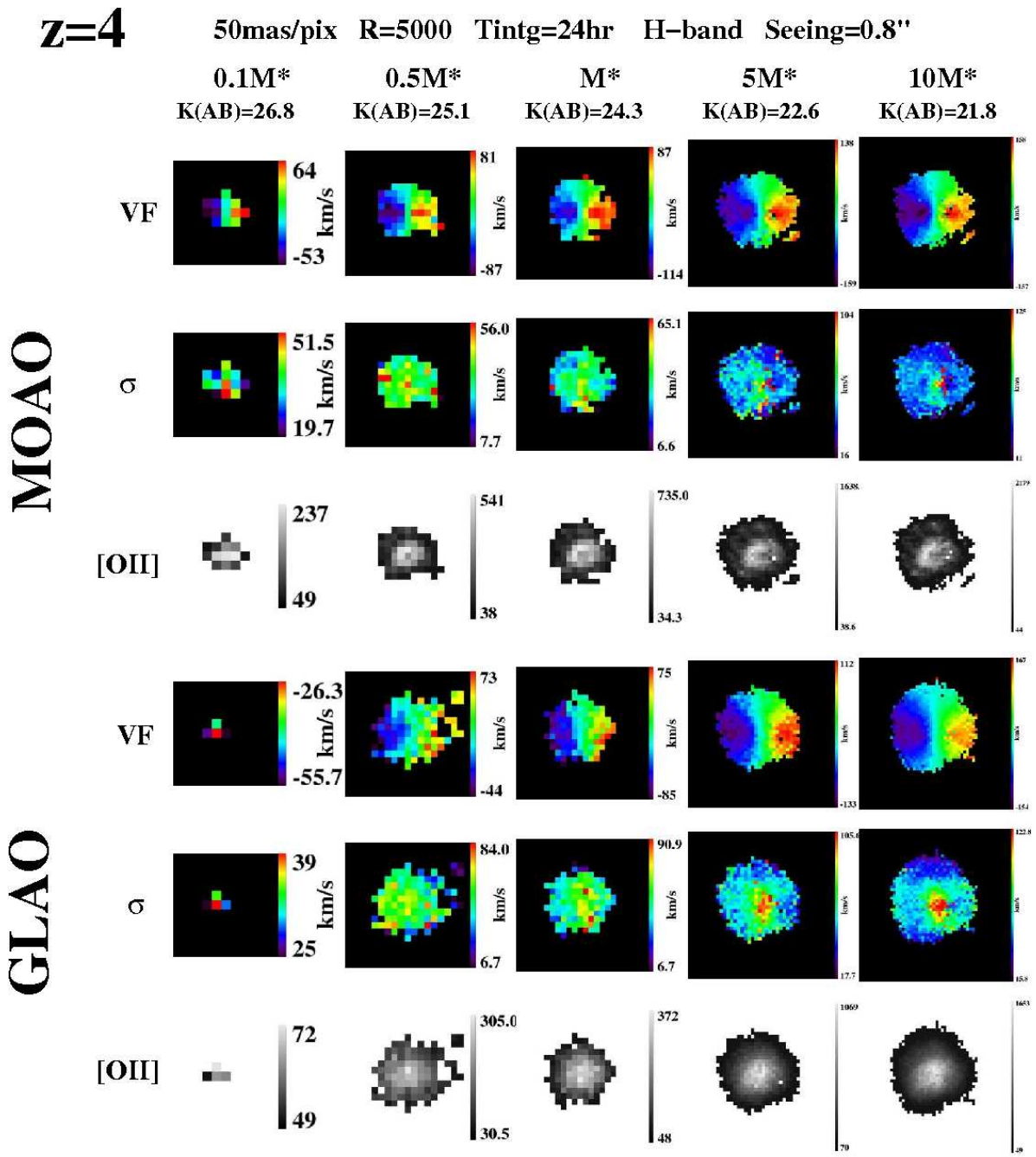

Fig. 1: Example of E-ELT simulation of spatially-resolved spectroscopic observations. An E-ELT equipped with a 3D spectrograph will provide kinematic maps of distant galaxies down to M* and up to $z \sim 4.5$. The M* galaxy case (central column) has $K(AB) = 24.4$ K and a line flux of 9×10^{-18} erg cm^2 s^{-1}. This figure illustrates the $z = 4$ case showing a rotating disk rescaled at different stellar masses (see Puech et al., in prep., for details). MOAO allows one to reach a better spatial resolution than GLAO (see the more curved isovelocities and the sub-structures in the [OII] emission map in the MOAO case), which is better suited for recovering the rotation curve of distant galaxies. However, GLAO already provides adequate spatially resolved kinematical information, which can be used to infer the dynamical nature of this source, e.g. a rotating disk. The latest AO PSF models (Feb 2008) have been used here.

6. Attachments (Figures)

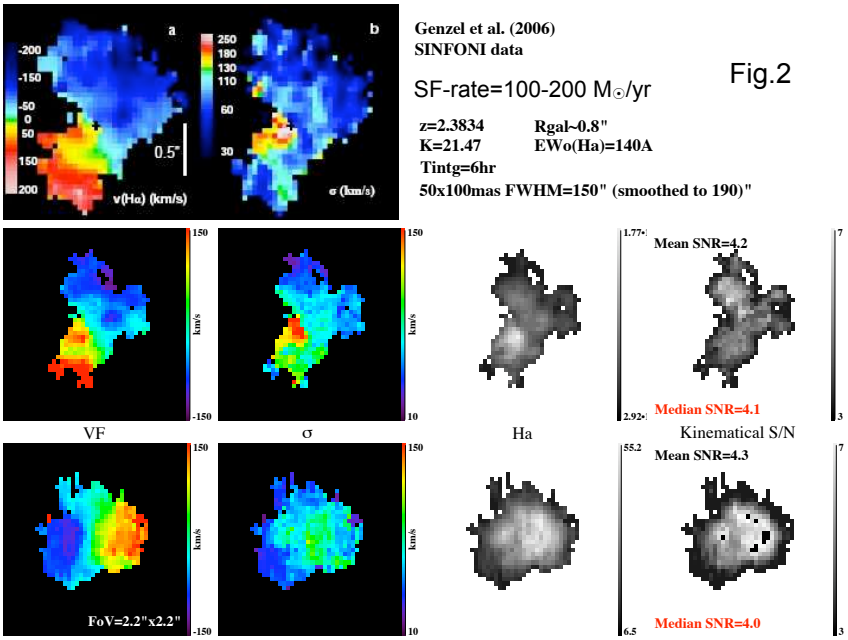

Fig.2

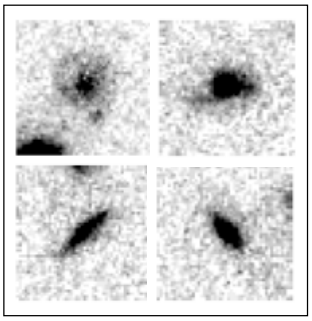

Fig.3

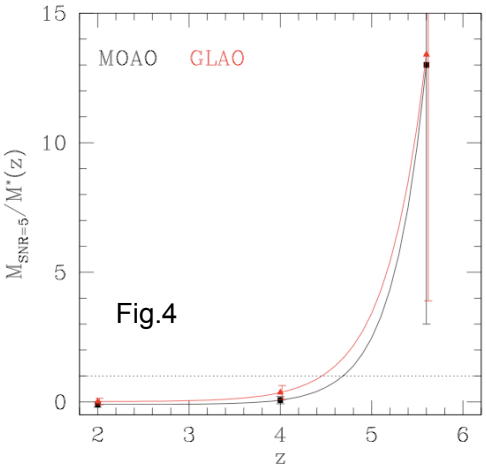

Fig.4

Fig. 2: Simulations of Genzel et al. observations of a massive star-forming galaxy at $z = 2.3$ with SINFONI on the VLT. *First line*: original Genzel et al. results; *second line*: re-analysis of the Genzel data cube with Puech's reduction pipeline (from left to right: velocity field, velocity dispersion, and Hα SNR; *third line*: simulation of VLT/SINFONI observations of a local disk galaxy rescaled to match physical parameters (mass, size, line EW, K-magnitude and kinematics) of Genzel's galaxy. Note how these simulations reproduce very well the observed signal-to-noise (rightmost images) in the Hα line.

Fig. 3: Nicmos-3 images of 4 massive ($\sim 10^{11}\,M_\odot$) star-forming (with a typical star-formation rate of 100 $M_\odot$/yr) galaxies at $z = 2 - 3$. Images are 5″across. (From Toft et al. 2007).

Fig. 4: Galaxy (stellar) mass corresponding to the minimum S/N needed for reliable kinematic measurements (see text) as function of redshift. The vertical bars cover a wide range of dynamical and morphological cases (Puech et al. 2008, in prep.). This plot shows that one can obtain reliable kinematically maps of galaxies down to M^*, out to $z \simeq 4.5$.

7. Justification of requested observing time and lunar phase

Lunar Phase Justification: Provide here a careful justification of the requested lunar phase.

Time Justification: (including seeing overhead) As a reference case, we consider a $z = 4$ star-forming galaxy with $H_{AB} = 24.3$ (or 22.9 Vega) which is a good estimate for an M^* galaxy from current stellar mass function studies. We assume an half-light radius of $R_H = 0.2$ arcsec. A robust assessment of the feasibility for this science case requires detailed 3D spectroscopy simulations, which were carried out using the Puech et al. methodology (Puech et al. MNRAS, submitted). Simulations were performed for a reference disk galaxy with a circular velocity of 230 km/s, size $= 4R_H = 0.8''(5.6\,kpc)$, observed in the [OII] line with $EW_0 = 30\,\mathring{A}$. We adopt a redshift dependence of the size as reported by Ferguson et al. 2004, and rescaled galaxy sizes as a function of mass following Courteau et al. 2007 (i.e. $R_{1/2} \propto M_{stellar}^{0.35}$). The S/N map peaks at $\sim$25 in the center, with a mean over the entire galaxy of $S/N = 12$. The sky model spectrum for Mauna Kea from the Gemini web pages was used (this corresponds to a mean sky brightness of 16.4 Vega-mag/arcsec^2). An MOAO correction was assumed giving 45% encircled energy within 100 mas). With this S/N one can recover the bidimension kinematics, in particular classify kinematically the galaxy (e.g. distinguish a disk from a merger) and reliably measure the rotation curve over the entire galaxy. A full grid of simulations (Puech et al. in prep., see example ein Fig.1) for different set up parameters (R, sky brightness, AO case, pixel scale), as well as physical parameters (R_H, mass, redshifts) have been run.
Results from these simulations suggest that to ensure robust kinematic measurements, we need a signa-to-noise over the galaxy size of at least 5, with the following scaling according to exposure time T, telescope diameter D, equivalent width EW, spectral resolution R, and pixel scale Δpix:

$$< S/N >_{min} = 5 \left(\frac{T}{24h}\right)^{0.5} \left(\frac{D}{42\,m}\right) \left(\frac{EW}{30\text{Å}}\right) \left(\frac{R}{5000}\right)^{-0.5} \left(\frac{\Delta pix}{50\,mas}\right)$$

The stellar mass that corresponds to such a minimal SNR increases with redshift z, according to

$$M_{lim}(z)/M^* = 0.1 + 3.3 \times 10^{-6} \exp(z/0.37)$$

For the characteristic stellar mass, M^*, we assumed $\log_{10} M^*(z) \approx 11.6 + 0.17z - 0.07z^2$, an empirical relation which describes the evolution of the observed mass function (Grazian et al. 2006).

Calibration Request: Special Calibration - None

8. Instrument requirements

Table 1. – **Main Technical Requirements and Target Parameters**

Requirements	Range of values (minimum/ideal)
Field of View	25–100 arcmin^2
Wavelength range	0.8–2.5 μ
Multiplexing	10–100
Spatial sampling	50–100 mas
Spectral resolution R	3–5000
50% enclosed energy diameter	$<$ 250 mas (MOAO)
Target parameters	
Typical magnitude	22–25 (K_{AB})
Redshift range	2–6
Typical mass	$> 10^{10}$ M$_\odot$
Target density	0.1–10 arcmin^{-2}
Object size	0.1–0.5 arcsec
Typical Exptimes	10-50 hrs

9. List of targets proposed in this programme

Run	Target/Field	α(J2000)	δ(J2000)	ToT	Mag.	Diam.	Additional info	Reference star
A	UDF	3:32:39.0	-27:47:00	10-100	25	DM	ang diam(')	note

Target Notes: N/A

EUROPEAN SOUTHERN OBSERVATORY

Organisation Européenne pour des Recherches Astronomiques dans l'Hémisphère Austral
Europäische Organisation für astronomische Forschung in der südlichen Hemisphäre

VISITING ASTRONOMERS SECTION • Karl-Schwarzschild-Straße 2 • D-85748 Garching bei München • e-mail: visas@eso.org • Tel.: +49-89-32 00 64 73

APPLICATION FOR OBSERVING TIME PERIOD: **78A**

Important Notice:

By submitting this proposal, the PI takes full responsibility for the content of the proposal, in particular with regard to the names of COIs and the agreement to act according to the ESO policy and regulations, should observing time be granted

1. Title Category: **A–1**

High resolution imaging of high redshift galaxies

2. Abstract

We propose to take very high resolution imaging of high redshift galaxies to understand their physical conditions. These observations will be taken at the diffraction limit of the telescope (6,8,10 milli-arcsec in J, H, K for a 42m telescope). The resolution will be approximately 60 pc in physical length, and will be of comparable quality 1 arcsec imaging on VIRGO galaxies. They will give typically > 100 resolution elements over the galaxies, and produce detailed information about the morphology, dynamical state, and variations in physical parameters across the galaxy. These observations will image galaxies beyond the Balmer break up to $z = 4.5$, enabling the separation of old and young stars. High contrast observations of QSOs will produce host morphologies, colors, and physical properties.

3.

Run	Period	Instrument	Time	Month	Moon	Seeing	Sky Trans.	Obs.Mode
A	79	FORS2	300h	any	d	$\leq 0.4''$	PHO	v
B	79	FORS2	300h	any	d	$\leq 0.4''$	PHO	v
C	79	FORS2	300h	any	d	$\leq 0.4''$	PHO	v

4. Number of nights/hours Telescope(s) Amount of time

a) already awarded to this project:

b) still required to complete this project:

5. Special remarks:

In order to establish feasibility simulations are needed to verify integration time.

6. Principal Investigator: Marijn Franx (Leiden University, NL, franx@strw.leidenuniv.nl)

CoI(s): The Rest (Elsewhere, ESO)

7. Is this proposal linked to a PhD thesis preparation? State role of PhD student in this project

8. Description of the proposed programme

A) Scientific Rationale: The physics of high redshift galaxies is still an enigma. Most galaxies look irregular in current imaging (either from space or from the ground); their dynamical state is not well defined, and it is not well understood what processes drive their evolution.
High resolution imaging of these galaxies will be crucial for a proper understanding. Their typical sizes are on the order of 0.2-0.3 arcsec at redshifts 2-4, decreasing at higher redshift (roughly like 1/(1+z) at fixed restframe luminosity), and decreasing with smaller luminosity (roughly like $L^{0.5}$). HST imaging generally gives very few resolution elements, at a level of 10 within an effective radius. This is equivalent to imaging of Virgo galaxies with a seeing of 10 arcsec !
It is clear that high resolution imaging will provide direct insight into the dynamical state of the galaxies, and allows us insight into the processes that drive these galaxies. As an example, recent work based on HST imaging has indicated that a fair fraction of z=2-3 galaxies may be mergers. However, the resolution is truely poor to establish this well. High resolution imaging should be able to provide details about the occurence of double nuclei, tidal effects in the main body of the galaxies, and the occurance of tidal arms.

B) Immediate Objective:
We will image high redshift galaxies in well studied "deep fields", including the HDF-South, UDF, Goods, and other areas for which multi-wavelength studies have been made. We will want to image about 1e4 L* galaxies at redshifts between 2 and 5. Given the number density of a few/sq arcmin, we'll need an area of 2000 sq arcmin in total.
The resolution should be close to the diffraction limit, and constant over the field. An MCAO system with a wide-field imager would be appropriate for these observations.
We will also image QSO hosts. Some of these will fall automatically on the fields; but it may be better to actually image these systems with LTAO. This requires careful simulations. LTAO will provide better psf, but it may be hard to know it. MCAO will give worse psf, but better stability - and hence knowledge.

C) Telescope Justification: The ELT is needed for both the spatial resolution and the faintness of the galaxies.

D) Observing Mode Justification (visitor or service):

E) Strategy for Data Reduction and Analysis:

9. Justification of requested observing time and lunar phase

Lunar Phase Justification: We require the maximum sensitivity and photometric accuracy possible which means dark time.

Time Justification: (including seeing overhead) We used the ETC (May 2007) to calculate the required integration time. We assumed that we want to observe a galaxy with Hab=23 and a size of 0.2 arcsec. We used the "surface brightness" option of the ETC. The surface brightness of these galaxies is about 20.2 in AB, or 18.8 in Johnson. According to the ETC, this should give a S/N per 5mas pixel of about 25 in 10 hours . To reach this, I had took S/N(1x1arcsec)/sqrt(pixels in1x1 arcsec) = 5000/200= 25. It is not clear whether this is correct.
The number of L* galaxies per arcmin is at a level of 5-10. Hence, to image 1000 L* galaxies, we like to image at least 100 sq arc min. This requires 10hours*(100/instantaneous FOV). For an MCAO system, this might lead to acceptable total integration time if the coverage is > 1 arcmin. We'll assume 2arcmin FOV, leading to about 500 hours integration time/band. We'll need at least 2 filters, hence 1000 hours total.
Notice that this will image many more galaxies than a 1000, but many are far below L*. For example, UDF has 2300 galaxies to AB=29 with z¿2 in a field of 6 sq arcmin - hence this will give 4e4 galaxies.
Careful simulations are needed of the QSO hosts. The ETC is cannot be used without such simulations.

Calibration Request: Special Calibration - Regular observations of standard star fields (!)

10. Report on the use of ESO facilities during the last 2 years

Report on the use of the ESO facilities during the last 2 years (4 observing periods). Describe the status of the data obtained and the scientific output generated.

11. Applicant's publications related to the subject of this application during the last 2 years

12. List of targets proposed in this programme

Run	Target/Field	α(J2000)	δ(J2000)	ToT	Mag.	Diam.	Additional info	Reference star
A	name	RA	DEC	time(hrs)	mag	DM	ang diam(')	note
A	COSMOS	10	+2:12	300	26-28	45	90	COSMOS FIELD
B	CDFS	03 32 28	-27 48 30	300	26-28	45	90	cdf-south
C	HDFS	22 32 56	-60 33 02	300	26-28	45	90	hdf-south

Target Notes: The targets in run A and the main ones for this exercise; the targets in run B give a more general impression of the types of (large) galaxy we might like to observe with an ELT with the same requirements as run A.
Number of points per field listed with time, I, K mags listed. Angular diameter is the total size of the entire galaxy not always necessary to observe the entire galaxy.

12b. ESO Archive - Are the data requested by this proposal in the ESO Archive (http://archive.eso.org)? If yes, explain why the need for new data.

13. Scheduling requirements

14. Instrument configuration

Period	Instrument	Run ID	Parameter	Value or list
79	FORS2	A	IMG	VIJHK
79	FORS2	B	IMG	VIJHK

European Organisation for Astronomical Research in the Southern Hemisphere

Organisation Européenne pour des Recherches Astronomiques dans l'Hémisphère Austral
Europäische Organisation für astronomische Forschung in der südlichen Hemisphäre

VISITING ASTRONOMERS DEPARTMENT • Karl-Schwarzschild-Straße 2 • D-85748 Garching bei München • e-mail: visas@eso.org • Tel. : +49-89-32 00 64 73

Important Notice:

By submitting this proposal, the PI takes full responsibility for the content of the proposal, in particular with regard to the names of CoIs and the agreement to act according to the ESO policy and regulations, should observing time be granted

1. Title Category: **A–1**

"First Light" – the highest redshift galaxies

2. Abstract

We propose to use integrated JHK spectroscopy on the ELT (in GLAO mode) to identify the first UV emitting sources in the Universe responsible for reionization at $z > 6$. The galaxy candidates will be identified via the Lyman-break technique with upcoming deep space-based and groundbased near-IR surveys (with HST/WFC3, Spitzer, VISTA, and JWST). The source densities at $z \sim 8$ are about 1 arcmin^{-2} at $H_{AB} = 28.5$, for which an ELT spectrograph equipped with a GLAO or LTAO/MOAO system will be necessary to obtain redshifts and to understand their physical properties. These observations will probe an era of a few hundreds million years right at the end of the "dark ages", will identify the sources of reionization and yield their luminosity function. Such a program will give the first insight of the crucial early stages of galaxy formation process.

3.

Run	Instrument	Time	Month	Moon	Seeing	Sky Trans.	Obs.Mode
A	79Multi-IFU	300h	any	d	$\leq 0.4''$	PHO	v
B	79Multi-IFU	250h	any	d	$\leq 0.4''$	PHO	v
C	79Multi-IFU	250h	any	d	$\leq 0.4''$	PHO	v

4. Principal Investigator: Matt Lehnert (GEPI, Observatoire de Paris, F, matthew.lehnert@obspm.fr)
CoI(s): Marijn Franx (Leiden University, NL), Bram Venemans (ESO, ESO)

5. Is this proposal linked to a PhD thesis preparation? State role of PhD student in this project

6. Description of the proposed programme

A) Scientific Rationale: **I. The First Galaxies** When and where did the first galaxies form? From very general considerations, the critical mass scale at which the "first galaxies" likely formed is in $\sim 5 \times 10^7$ $M_\odot$ halos assembled at z$\sim$10. At this scale, the virial temperature of the halo exceeds 10^4 K and atomic cooling becomes the dominant cooling mechanism. Halos above this mass can retain photoionized gas and therefore likely maintained a self-regulated multi-phase ISM sustained by the energy due to star-formation (e.g., Mac Low & Ferrara 1999; Madau et al., 2001; Ricotti et al., 2002). These are the "first galaxies".

Many other questions surround the formation of the first galaxies. What was the shape of the IMF, if such a designation even makes sense for the first galaxies? The IGM is likely to have been heavily enriched (Becker et al. 2009), the radiation fields may have been intense and the turbulence high during this epoch. The balance between metalicity, turbulence generated by collapsing gas and mechanical energy injection from the stars, and the strength of the radiation is not at all academic as it determines whether the star-formation is regulated by cooling through molecular or atomic lines. These two cooling mechanisms could lead to substantially different IMFs (Bromm et al. 2003; Greif et al. 2007; Wise & Abel 2008; Greif & Bromm 2006). Moreover, the accretion of cold gas from filaments in combination with a softened equation of state drives strong shocks and turbulence (e.g., Keres et al. 2005) which might lead to vigorous fragmentation and the formation of the first star clusters. Such a situation would be an important early step towards conditions that are similar to star-formation in the local Universe and hence a "normal" IMF.

However, to know whether or not any of our modeling represents reality, and to solve this "tension" in the models of the first galaxies, observers must find the first galaxies and then confirm that they are indeed in the early Universe. To overcome the problem of discovering galaxies at high redshift, regardless of their properties, we should use selection methods which rely on extrinsic rather than intrinsic galaxy properties and which, if possible, can be applied over the entire history of the Universe. The dropout technique, which selects galaxies according to the strength of their Lyman break (so-called Lyman break galaxies) is such a selection method. LBGs have been the benchmark against which other samples of high z galaxies have been compared since their discovery by Steidel and collaborators and they allow us to study galaxies selected in a consistent manner over z=0-7 and now perhaps even up to z=8.5 (Bouwens et al. 2009). So, while LBGs can in principle select every type of galaxy, we inherently add a bias which is that they must be actively star-forming with little reddening to "clean" the samples of "interlopers". At higher and higher redshifts it becomes difficult to remove interlopers as data of sufficient depth at longer wavelengths (i.e., beyond 2.5μm) are difficult to obtain. However, despite this, it is the simplicity of selection of the LBG technique, its robustness, and wide redshift span probed that makes the LBG selection technique unrivaled for (comparative) galaxy evolution studies.

Luminosity Function and Correlation Lengths: One of the most important parameters that we would like to know is the spatial distribution and clustering properties of the galaxies as a function of magnitude. From such studies we can determine the halo mass and the likely evolutionary outcomes for this first galaxies. In order to accomplish this, we need to:

1. measure the redshifts, confirm that the galaxies are at high redshift; give the redshift distribution for the population as a function of magnitude. From this, we can determine their 3-D spatial correlation and their likely halo masses.

2. determine the luminosity function of the galaxies at a variety of wavelengths. From this combined with thier spatial distributions will allow us to begin guage their impact on the IGM and their integral mass to light ratios - useful for constraining the IMF.

The Violent ISM of Young Galaxies at z>6: All the of the most distant galaxies known have properties suggesting that they are in their formative phases and are sites of intense star-formation. The star-formation intensities are well above the threshold for galaxies driving winds in the local universe (0.05 $M_\odot$ yr^{-1} kpc^{-2}) and LBGs at lower redshifts have strong evidence for driving winds from rest-frame UV spectroscopy (e.g., Shapley et al. 2001; Vanzella et al. 2009). Studies of starburst-driven outflows in nearby galaxies have determined that these flows are a multiphase phenomenon where each phase reveals important details about the ionization state of the ISM and the distribution of mass, energy, and thermal and ram pressure. Without constraining all of these characteristics, it is difficult to know what the ultimate fate of the gas may be. How much is retained in the halo and how much is likely ejected? For example, in starbursts, the observed outflow velocities of the optical emission line gas are much lower than the expected terminal velocity of the wind (i.e., Lehnert & Heckman 1996). In contrast, line widths and offset velocities of the *absorption* line gas seem to trace the terminal velocities and correspond to the wind velocities estimated from the X-ray emission (Heckman et al. 2000). The absorption line gas also has the advantage that it probes the gas between us and the UV/optical continuum emission. Thus if we see the gas blueshifted relative to systemic it is outflowing. With emission line gas, there is not this sense of direction - all velocities can be seen. In addition, the absorption line gas probes

large ranges of temperature and ionization state (T$\sim 10^4$ to 10^6 K and many ionization states, for example CI to CIV).
By determining the properties of the rest-frame UV emission and absorption lines, we can:

1. Determine if the warm/hot ISM will escape the galaxy and halo potential.
2. Determine the metallicity of the galaxy and the rate at which metals are ejected into the halo and/or IGM by the outflowing gas. Are their sufficient baryons to account for the metallicity of the early IGM?
3. Determine the velocity and total energy in the outflowing gas. Is there sufficient energy and velocity (to estimate how far the outflow reaches) to explain the topology of the IGM as determined through absorption line studies from background QSOs (if any), GRBs, and supernovae. Is the energy per baryon sufficient to explain the phase of the early IGM?
4. What is the covering fraction of neutral and highly ionized gas? What is the topology of the ISM in these early galaxies? What is the escape fraction? Combined with the luminosity function, is it sufficiently high given the number of sources to account for reionization?
5. Combined with observations with infrared/submm/mm facilities such as ALMA and JWST/MIRI (CO and H_2 lines) how much of the mechanical energy of the star-formation dissipated in the warm/cold molecular gas or advected out of the system through winds? How much mass is in each phase of the ISM in these early galaxies? Is the ISM turbulent?
6. Observations with infrared/submm/mm facilities such as ALMA and JWST/MIRI (CO and H_2 lines) and direct detection of stellar photospheric lines with E-ELT we can ask: How efficient is the star-formation in these early galaxies?
7. Combined with rest-frame optical and near-infrared imaging and a direct detection of the UV stellar photospheric lines which form in the atmospheres of O and B stars, we can determine the age and number of O and B-stars. This will allow us to determine the shape of the IMF. Was it top heavy? These lines will also allow us to estimate the metallicity of the young stars. HeII emission might be seen which in combination with the UV absorption lines may tell us what fraction of Population III stars these galaxies may host.
8. Galaxies at high redshifts are often multi-modal (Douglas et al. 2009) similar to low-z LBGs (Overzier et al. 2009). What is the nature of these components, velocity differences, internal properties? Are these mergers or clumpy star forming complexes (see Overzier et al. 2009)?

B) Immediate Objective: Our immediate objective is a low spectral resolution study (R=1000-3000) to obtain redshifts followed by a detailed investigation of the nature of the stellar and interstellar absorption lines in the most distant galaxies known. All galaxies are taken from our large sample of galaxies drawn from the deep surveys with HST/WFC3, JWST/NIRCam, and ultra deep surveys with the VLT and other 10m class telescopes. Some fields already have a substantial number of galaxies which have been spectroscopically confirmed by JWST/NIRSpec.
Moderate resolution spectroscopy of the most distant galaxies with a high multiplex spectrometer will provide:
• A wide variety of absorption lines in the UV and blue optical. These include the important interstellar lines of: Lyα, NV$\lambda\lambda$1239, 1243, SiII$\lambda\lambda$1260,1265 OI/SiII$\lambda\lambda$1302,1305, SiII$\lambda\lambda$1304, 1309, CII$\lambda\lambda$1334,1336, SiIV$\lambda\lambda$1393, 1402, SiII$\lambda\lambda$1527,1533 CIV$\lambda\lambda$1548, 1550, MgII$\lambda\lambda$2796, 2803, FeII lines, etc. as well as several strong stellar absorption lines (e.g, SVλ1502, CII$\lambda\lambda$1426,1428; some of the lines listed have contributions from both). Such a line list will allow us to probe gas ranging from 10^4 to 10^6 K.
•A wide variety of strong emission lines including: Lyα,NV, CIV, HeII, etc.
Specifically, with these intermediate resolution spectra, we will:
(1) measure the relative velocity offsets of the various UV absorption lines as a function of ionization to Lyα HeIIλ1640,etc. This will uniquely and quantitatively show whether they are driving winds. Blueshifted absorption lines relative to recombination lines will be strong evidence for winds and allow to quantify their

terminal velocities. From ionization calculations, covering fraction estimates (see next point), and column density measurements, we can estimate the mass/energy outflow rates.
(2) estimate the covering fractions of low-ionization and high-ionization absorption line gas. The lines are resolved, which means that the covering fraction can be estimated from the residual core intensity of the line. Our spectra will contain both low and high ionization UV lines, hence we can estimate the covering fraction of various gas phases. This will give us clues on the distribution of the ISM in this galaxies and help constrain the outflow rates.
(3) estimate the escape fraction of ionizing photons. An HI column depth of only 1.6×10^{17} cm^{-2} produces $\tau \sim 1$ at the Lyman edge. Since the mean columns of galaxies are much greater ($\sim10^{21}$ cm^{-2}), the leakage of photons from any system must be determined by its topology. We expect to resolve the UV absorption lines, hence if we know the velocity widths, we can derive the line depth from absorption line modeling. CIIλ1335 is produced in the same regions as HI; thus the core residual intensity of CII will yield direct constraints on the areal coverage of optically thick, neutral gas.
(4) measure the photospheric lines. From the photospheric and stellar wind lines we will derive the main properties of the UV-emitting stellar population (composition, age, metallicity). If any of these galaxies have H_2 mass determinations, we can estimate the star-formation efficiency. The UV absorption lines in comparison with emission lines such as HeII will allow us to constrain the metallicity and ionization state of the gas. Can we find evidence for population III stars?

C) Telescope Justification: Only the large light grasp and high spectral and spatial resolution of the E-ELT can think about accomplishing this experiment.

D) Strategy for Data Reduction and Analysis: The strategy for this proposal depends on a number of factors. The main one is the competitiveness with JWST. At resolutions greater than about 3000, a 42m E-ELT is about a factor of 3 faster in integration time to each the same depths. For example, in 10^5 s, JWST/NIRSpec can reach K_{AB}=24 at 10σ which E-ELT can reach this depth in about 3.5×104s. The number of Y-J break galaxies is about 0.5 arcmin^2 down to H_{AB}=28.5 (Bouwens et al. 2009). In the field of view of an E-ELT, assumed to be about 30-40 arcmin^2. Thus, if a spectrograph with deployable IFUs has a multiplex of about 20, its effective multiplex is larger than JWST/NIRSpec because of the larger field of view. The E-ELT can actually obtain redshifts faster at these resolutions. At R=1000, the situation is reversed and JWST is much faster (about a factor of 5 in integration time). R=1000 is sufficient for redshifts and characterizing the absorption line EQWs but not anything more. Also, if JWST stays on schedule, they would have a headstart of a few years for obtaining redshifts, stacking analyses, clustering amplitude determinations, metallicity/age estimates, etc - albeit on relatively small numbers of galaxies (perhaps 10s). The niche of the E-ELT is then in the physics of the ISM and the topology of the IGM. If JWST does not fly, then even a R=3000-5000 multi-object spectrometer will be a capable redshift machine. It is not clear if lower resolutions will be able to get redshifts if there are not strong emission lines.

6. Attachments (Figures)

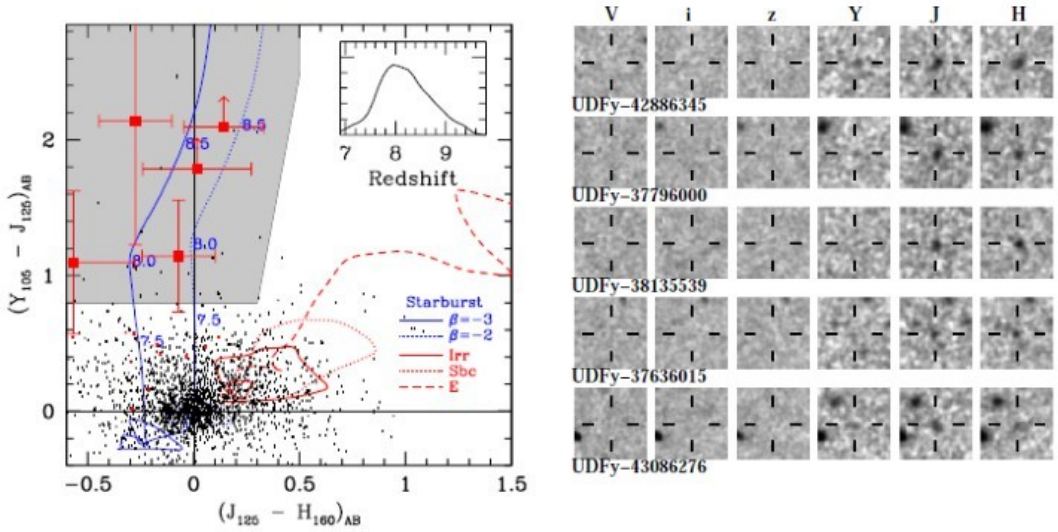

Fig. 1: *(left)* The source color selection, J_{125}-H_{160} versus Y_{105}-J_{125}, used by Bouwens et al 2009 to select very high redshift galaxies. The red squares represent the distant galaxies and the numbers are the best fit photometric redshift for each source. *(right)* Postage stamps of the images of the 5 candidate z∼8 galaxies. The surface density of z∼8 candidates is about 0.5 arcmin^2 at H_{AB}∼28.5.

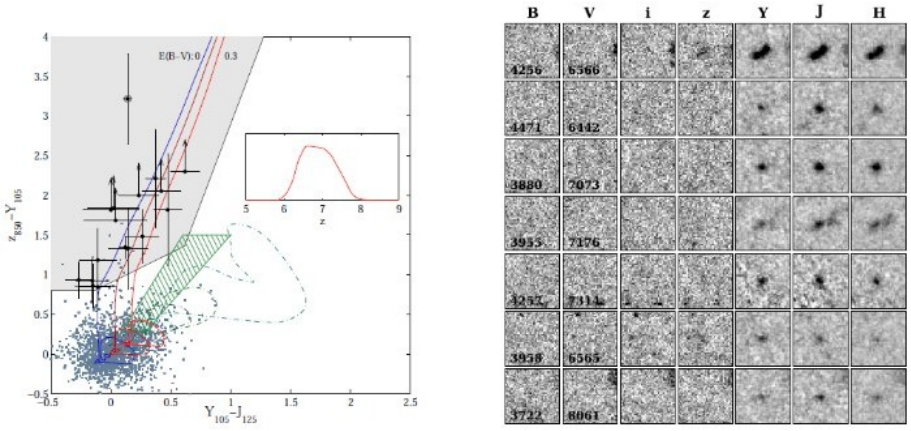

Fig. 2: *(left)* The source color selection, Y_{105}-J_{125} versus z_{850}-Y_{105}, used by Oesch et al. 2009 to select z∼7 galaxies. The black circles represent the distant galaxies with galaxies of various redshifts with different reddenings are shown by the lines. The shaded area in the color-color diagram is the selection region. *(right)* Postage stamps of the images of some of the candidate z∼7 galaxies (see Oesch et al. for the rest).

6. Attachments (Figures)

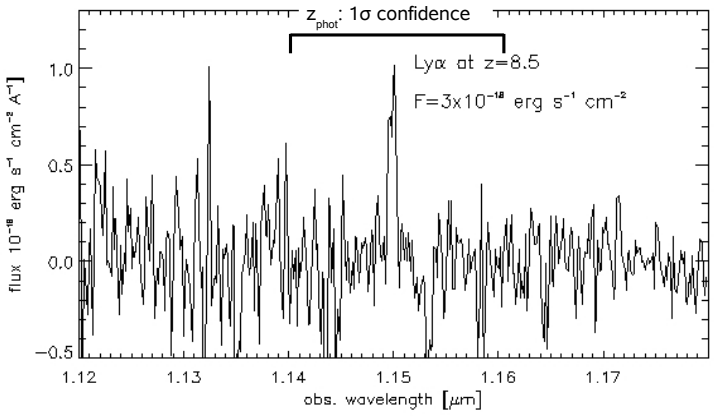

Fig. 3: Artificial (Gaussian) Lyα line with observed FWHM=200 km s^{-1} and flux 3×10^{-18} erg s^{-1} cm^{-2} at z$\sim$8.5 superimposed on a real SINFONI J-band spectrum. The cube is was scaled such that it represents 12 hrs of integration time. The spectrum was extracted from a box aperture of 5×5 pixels ($0.625'' \times 0.625''$) and smoothed by 3 pixels along the spectral axis. Note that there are several bright night sky lines near 1.15μm which we were able to subtract robustly with our routines (which is a strong factor in determining whether this experiment is even feasible in the first place). The black bar shows the 1σ confidence interval of the photometric redshift given by Bouwens et al. (2009). Such a line would imply equivalent widths of more than about 150Å in the rest-frame for H_{AB}=28.2. An equivalent integration with a E-ELT would push this a factor of 5 deeper or down to 30Å. Most distant lyman break galaxies have EQWs higher than this.

7. Justification of requested observing time and lunar phase

Lunar Phase Justification: The spectrograph will be used in the near-infrared and so we have no lunar phase requirement.

Time Justification: (including seeing overhead) The ETC indicates that H_{AB} 27 can be reached in 30 hours. Hence the ULTRA VISTA galaxies require integration time of that amount. This assumes that the galaxies convolved with the psf have a half-light radius of 200mas. This is unknown ofcourse, but hopefully reachable with GLAO. The surface density of these galaxies is approximately 0.38 / sq arcmin; and we wish to cover 2600 sq arcmin. Obviously, wide field access is critical. If the effective field is 10 arcmin, we can do this program in approximately 30 pointings, hence 300 hours. Deeper observations will be done on a bright subsample to give better physical information.
The galaxies with Hab=28 will be done in approximately 100 hours per integration. Their surface density is expected to be higher by about a factor of 6 - about 2 per sq arcmin. This number is of course quite uncertain - it should be measured. Working over a field of 10x10 arcmin, this would give 200 galaxies per 100 hours. In 500 hours, 1000 galaxies can be measured, yielding a meaning full sample.

Calibration Request: Special Calibration - Regular observations of standard star fields (!)

8. Instrument requirements

AO-fed multi-IFU spectrograph (Multi-IFU)
Wavelength range required: 10000Å to 14000Å.
Spectral resolution: 4 000 –10 000

9. List of targets proposed in this programme

Run	Target/Field	α(J2000)	δ(J2000)	ToT	Mag.	Diam.	Additional info	Reference star
A	COSMOS	10	+2:12	300	26-28	45	90	COSMOS FIELD
B	CDFS	03 32 28	-27 48 30	250	26-28	45	90	CDF-south
C	HDFS	22 32 56.22	-60 33 02.69	250	26-28	45	90	HDF-south

Target Notes: These are typical targets. Cosmos field will likely happen; other fields are more uncertain, they depend on what JWST will do.

European Organisation for Astronomical Research in the Southern Hemisphere

Organisation Européenne pour des Recherches Astronomiques dans l'Hémisphère Austral
Europäische Organisation für astronomische Forschung in der südlichen Hemisphäre

VISITING ASTRONOMERS DEPARTMENT • Karl-Schwarzschild-Straße 2 • D-85748 Garching bei München • e-mail: visas@eso.org • Tel. : +49-89-32 00 64 73

APPLICATION FOR OBSERVING TIME LARGE PROGRAMME PERIOD: **80A**

Important Notice:

By submitting this proposal, the PI takes full responsibility for the content of the proposal, in particular with regard to the names of CoIs and the agreement to act according to the ESO policy and regulations, should observing time be granted

1. Title Category: **A–8**

Monitoring the redshift-drift of the Lyman-alpha forest – a direct measurement of the dynamical evolution of the Universe

2. Abstract / Total Time Requested

Total Amount of Time: 3750h Total Number of Semesters: 16 (30)

We propose to monitor the redshift drift of the Lyα forest and associated metal lines towards a sample of 20 very bright QSOs in the redshift range $2 < z < 4.2$ with the ultra-stable high-resolution optical spectrograph on the E-ELT over a period of 15 yr (with observations spread over 8 yr). The redshift drift is sensitive to the difference of the expansion rate today and the expansion rate at the redshift of the absorbing structures and is directly related to the de- or acceleration of the Universe. With the proposed observations we can achieve an overall measurement accuracy of 3.5 cm/s. By monitoring the drift of the Lyα forest over a wide redshift range we will measure the instantaneous expansion rate of the Universe today and the expansion rate at high redshift. The measurement of the expansion rate and its evolution at high redshift will be an important test of General Relativity. The measurement of the instantaneous expansion rate will test whether the Universe expands today at the rate expected from other observations which, generally, measure the expansion rate averaged over a hundred Myr or more. The observations proposed can be used as a first epoch measurement for more accurate measurements by a future generation of astronomers and will thus represent a long-lasting legacy. The acquired spectra will also represent a unique resource for a wide range of other QSO absorption line studies.

3.

Run	Period	Instrument	Time	Month	Moon	Seeing	Sky Trans.	Obs.Mode
A	80	UVES	234h	any	d	$\leq 0.8''$	PHO	s

4. Principal Investigator: L. Pasquini (ESO, ESO, lpasquin@eso.org)

CoI(s): M. Haehnelt (IoA, UK), J. Liske (ESO, ESO), on behalf of the CODEX team (OTHER, OTHER), and the E-ELT SWG (OTHER, OTHER)

5. Description of the proposed programme

A) Scientific Rationale: The universal expansion was the first observational evidence that general relativity might be applicable to the Universe as a whole. Since Hubble's (1929) discovery much effort has been invested into completing the basic picture of relativistic cosmology. It is generally accepted that the observable Universe is homogeneous and isotropic and can be described by the Robertson Walker metric. Whether the evolution of its scale factor can, however, be accurately described by general relativity in four spacetime dimensions as we know it is still completely uncertain. The observational result of the mid-late 1990s that the expansion of the Universe has recently begun accelerating came as a big surprise. Within standard relativistic cosmology such an acceleration can be accommodated by modifying the stress-energy tensor of general relativity to include a new component with negative pressure. In its simplest incarnation this so-called dark energy is the cosmological constant already introduced by Einstein and all observational evidence so far is consistent with this interpretation. The unexpected deviation from the Einstein-de Sitter model favoured for a long time by many theoreticians may, however, be the smoking gun of the necessity for a more radical paradigm shift. A model-independent approach to mapping the expansion history as a function of redshift in detail appears therefore prudent. Using SNIa to measure the luminosity distance as a function of redshift has been the most successful attempt of this kind so far. The caveats are that distance is 'only' related to the expansion history through an integral over redshift and that one still requires a prior on the spatial curvature of the Universe.
We propose here to instead directly measure the expansion history of the Universe with a method first suggested by Sandage (1962). Sandage pointed out that the evolution of the expansion rate causes the redshifts of objects partaking in the Hubble flow to change slowly with time:

$$\dot{z} = (1+z)H_0 - H(z) \tag{1}$$

This implies that the expansion history can be determined, at least in principle, by means of a straightforward spectroscopic monitoring campaign. The signal is, however, very small. Fig. 1 shows the redshift drift as a function of redshift as expected in standard relativistic cosmology for three different sets of cosmological parameters. For a monitoring period of 10 yr the redshift drift at $z = 3$ lies in the range of 3-30 cm/s, depending on the choice of cosmological parameters, and is hence well within reach of what can be measured with the ultra-stable high-resolution optical spectrograph at the E-ELT in terms of accuracy and stability of the wavelength calibration.
As originally proposed by Loeb (1998) and discussed in detail by Liske et al. (2007), monitoring of the absorption spectra of bright QSOs with their large number of spectral features offers the best prospects for a measurement of the redshift drift. Note that the measurement of $\dot{z}$ for spectral features covering a range of redshift will allow us to measure the instantaneous present-day expansion rate as well as the expansion rate at high redshift averaged over about a Gyr. The measurement of the expansion rate and its evolution at high redshift will be an important test of general relativity and will – if general relativity is the correct theory of gravity on large scales – contribute significantly to characterizing the nature and evolution of dark energy. The measurement of the instantaneous expansion rate today will test whether the Universe expands today at the rate expected from other astronomical measurements of the present-day Hubble constant which generally constitute a measurement of the expansion rate averaged over hundred Myrs or more.
Finally, we point out that our data may also serve as a first epoch measurement to be supplemented with similar or superior data in the (distant) future in order to yield a much more accurate measurement of the redshift drift. Hence, these data will have a long-lasting legacy value.

B) Immediate Objective: We will obtain a series of extremely high S/N absorption spectra of a sample of about 20 very bright QSOs with the ultra-stable high-resolution optical spectrograph. In order to monitor the redshift drift the series of spectra will be taken over two periods of four years, separated by a further seven years, so that the total duration of the experiment is $\Delta t = 15$ yr.
Radial velocity measurements of the effect of planets with the ultra-stable high-resolution optical spectrograph have demonstrated that its design goal of a stable wavelength calibration with an accuracy of 1 cm/s is achievable over (arbitrarily) long periods of time. This superb accuracy of the wavelength calibration and wavelength stability means that the accuracy of our proposed measurement will be photon-noise limited (Pasquini et al. 2005; Liske et al. 2007). We have performed detailed simulations using mock absorption spectra to establish how well the redshift drift can be measured in the photon limited case using the method of Bouchy et al. (2001) originally developed for the detection of planets (Liske et al. 2007). In the photon noise limited case the accuracy with which the redshift drift can be measured depends only on the total number of photons collected and the number and sharpness of the spectral features. This accuracy can be conveniently parameterized as

$$\sigma_v = 1.4 \left(\frac{S/N}{2370}\right)^{-1} \left(\frac{N_{\rm QSO}}{30}\right)^{-0.5} \left(\frac{1+z_{\rm QSO}}{5}\right)^{-1.7}, \tag{2}$$

where $N_{\rm QSO}$ and $z_{\rm QSO}$ are the number and redshift of the QSOs. S/N refers to the S/N ratio per 0.0125 Å pixel in a single epoch measurement. The first two terms in equation (2) describe the dependence on the number of photons collected while the third term is an approximation based on our extensive simulations of mock absorption spectra that take into account the evolution of the flux distribution in the absorption spectra. Since the number of absorption features increases with increasing redshift the measurement error decreases.

Choosing the sample of target QSOs requires to find the best compromise between a sample that is as large as possible in order to optimize redshift and RA coverage and to check for systematic uncertainties, and a sample that contains only the very brightest QSOs in order to maximize the total number of photons collected. Here we fix the number of target QSOs at 20. Based on currently known QSOs accessible from the southern hemisphere (i.e. with DEC $< 35°$) we have chosen those 20 QSOs with the most favourable combination of magnitude and redshift, in the sense that this sample will give the best total accuracy (σ_v) that is possible for any sample of 20 QSOs and a fixed amount of observing time. The selected sample covers the redshift range $2 < z < 4.2$ and provides a wide coverage in RA to facilitate observations all year round.

At the time of our measurement the number of suitable bright QSOs should have increased and it should be possible to choose a larger sample of QSOs with a similar number of photons collected in the same amount of time. An additional benefit of a large sample of QSOs is the increased amount of additional science which can be expected to be done with these spectra.

Fig. 2 shows a Monte Carlo simulation of the expected outcome of the experiment using our sample of 20 QSOs, where the data have been split into four redshift bins. With an overall, total experiment accuracy of ~ 3.5 cm/s the redshift drift should be just about measurable.

The CODEX study has identified and investigated a number of possible systematic uncertainties, most notably peculiar accelerations of the absorbing structures and changes in the density, temperature and ionization state of the absorbers. The additional error components that are induced by these effects were all found to be at least two orders of magnitude smaller than the error due to photon noise. Hence, the uncertainties due to peculiar motions and the evolution of the physical state of the absorbers is not expected to limit our ability to measure the redshift drift.

Finally, we note that the unprecedented S/N of the spectra collected for the redshift drift measurement will also make them extremely valuable for a number of other QSO absorption line studies addressing fundamental questions. For example, the dataset should increase the accuracy with which one can test for the variability of physical constants by two orders of magnitude compared to present-day data.

C) Telescope Justification: The measurement is photon noise limited and requires the large aperture of the ELT and the accuracy and the stability of the wavelength calibration of the ultra-stable high resolution optical spectrograph. Proper correction of the earth motion requires that the photon barycenter of the time of the observation must be determined to better than 0.6 s. In order to minimize this effect and other possible sources of instability (e.g. airmass variations, instrument instabilities), the integration time of each exposure should be as short as possible. In order to reach the photon noise-limited regime in the total measurement each individual observation must be photon noise limited. This further adds to the necessity for the large photon collection area of the ELT. A S/N ratio of 30 per pixel is obtained in just 11 minutes for an object of magnitude 18.5. Hence there should be no problem with using reasonably short integration times (e.g. 20 minutes) with the ELT.

D) Observing Mode Justification (visitor or service): The observations are standard and are distributed over a large number of targets with a wide range of RA. Service mode is therefore the preferred option.

5. Attachments (Figures)

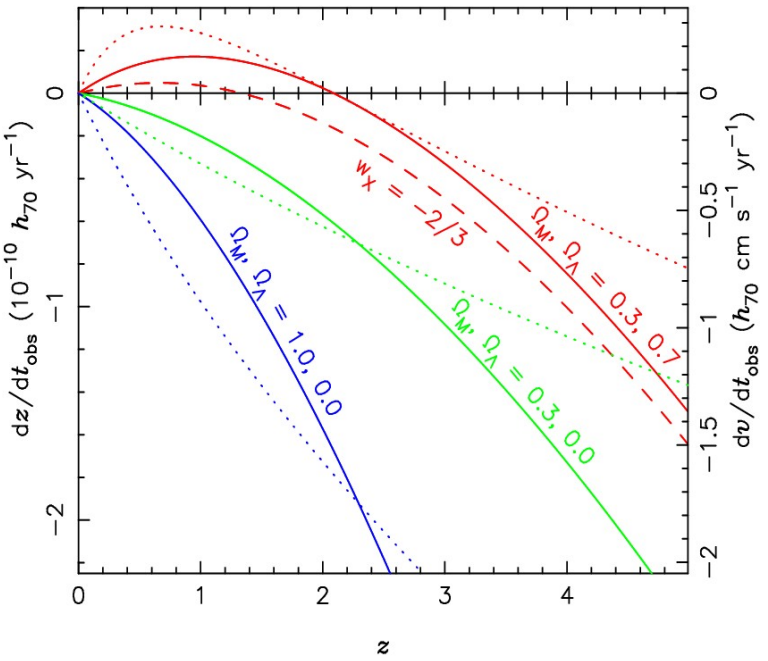

Fig. 1: The solid (dotted) curves and left (right) axis show the redshift drift $\dot{z}$ ($\dot{v} = c\,\dot{z}\,(1+z)^{-1}$) as a function of redshift for various combinations of $\Omega_{\rm M}$ and Ω_Λ as indicated. The dashed curve shows $\dot{z}$ for the case of dark energy with an equation of state parameter $w_X = -\frac{2}{3}$ (and $\Omega_{\rm M}, \Omega_X = 0.3, 0.7$).

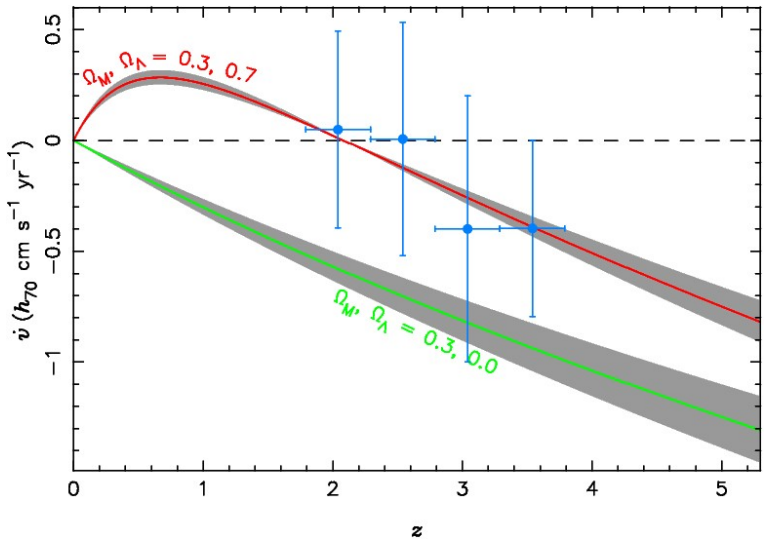

Fig. 2: Monte Carlo simulation of a $\dot{z}$ measurement using our sample of 20 bright QSOs in the redshift range $2 < z_{\rm QSO} < 4.2$. The time between the first and last observations was assumed to be $\Delta t = 15$ yr, with observations carried out during years 1–4 and 12–15. The blue points show the simulated data in bins of 0.5 and the expected error bars. The overall accuracy of this experiment is $\sigma_v = 3.5$ cm/s. The solid red curve shows the expectation for the input cosmological model of $H_0 = 70$ km/s/Mpc, $\Omega_{\rm M} = 0.3$ and $\Omega_\Lambda = 0.7$. For comparison the green curve shows the expectation for a model without a cosmological constant. The grey shaded areas result from varying H_0 by ± 8 km/s/Mpc. The discovery of any QSOs in the southern hemisphere that are brighter than any of the ones in our sample would increase the total accuracy of the experiment at a fixed total observing time.

6. Experience of the applicants with telescopes, instruments and data reduction

The expertise in the CODEX team covers all aspects of the project. The applicants have ample experience with using (and building) echelle spectrographs and team members have pioneered the precision measurement of radial velocities based on an accurate wavelength calibration. The team further has considerable experience with the reduction and analysis of QSO absorption spectra. The theoreticians on the team are familiar with all theoretical aspects of the project and will ensure an adequate interpretation of the data.

7. Resources available to the team, such as: computing facilities, research assistants, etc.

The proposing team has sufficient resources to perform the required data analysis. Following the positive experience with HARPS, the development of an optimized data reduction pipeline should have been part of the instrument development plan and we assume here the availability of such a pipeline. The final reduced data and intermediate data products will be made available to the community.

8. Special remarks:

The proposal assumes that the ultra-stable high resolution optical spectrograph reaches its design wavelength calibration accuracy of 1 cm/s over long periods of time and has adequate wavelength coverage. This puts the following requirements on the telescope

- availability of a gravity-stable Coudé focus
- plate scale stability: 0.02–0.05 arcsec centering
- wavelength coverage: 0.38–0.7 μm

Note that the instrument is seeing limited and that the anticipated aperture is 1 arcsec. The measurement would nevertheless benefit from GLAO in order to maximize the amount of energy in the fibre aperture. In this case it would have the following AO requirement: 80% EE in 1 arcsec.

9. Justification of requested observing time and lunar phase

Lunar Phase Justification: The reflected solar spectrum contains many spectral lines, which would produce a variable contaminating radial velocity signal in our QSO spectra. We have discussed this in detail in Pasquini et al. (2005: CODEX concept study, page 80). In order to limit the shift produced by an individual contaminating absorption line to 30 cm/s or less, its contrast with respect to the science target has to be larger than 10^4. This clearly calls for observations in dark time. How much moon light can be tolerated will be investigated in future using detailed simulations.

Time Justification: (including seeing overhead) We have used the ELT ETC for the high-resolution spectrograph (R=100000). The ETC results have been treated in the following way: First, the number of photons per pixel has been reduced to bring the ETC pixel size to that of our own detailed simulations (eq. 2, 0.0125 Å) which we consider an adequate pixel size for the ultra-stable high-resolution spectrograph. Secondly, we undid the ETC's application of an "aperture" loss factor. Depending on the selected spatial pixel scale and number of pixels the ETC uses simulated PSFs to calculate the fraction of the flux ensquared by the selected area. However, we expect the high-resolution spectrograph to be built with a 1 arcsec aperture so that aperture losses should be minimal. The current version of the ETC does not allow for this possibility, hence the correction.

The total exposure time was selected so as to achieve a total, overall accuracy of the experiment of 3.5 cm/s, resulting in an overall time request of 3750 hours. This time will be split among the targets in such a way as to achieve a similar overall S/N for each object.

Assuming that on average 1 night per week will be available for CODEX observations (or 26 nights per semester) it will take 4 years to complete each of two epochs. Separating the end of the first epoch and the start of the second epoch by 7 years results in a total, overall experiment duration of 15 years.

We reiterate that with the advent of future large surveys, a substantial addition to the list of bright QSOs is expected, especially in the southern hemisphere, enabling us to either reduce our time request or to obtain stronger results.

Calibration Request: Standard Calibration

10. Report on the use of ESO facilities during the last 2 years

N/A

11. Applicant's publications related to the subject of this application during the last 2 years

Pasquini L. and the CODEX Team, 2005, ESO publication OWL-CSR-ESO-00000-0160: OWL Instrument Concept Study: COsmic Dynamics EXperiment

Pasquini L. and the CODEX Team, 2005, The Messenger, 122, 10: CODEX: Measuring the Expansion of the Universe (and beyond)

Liske J. and the CODEX Team, 2007, MNRAS, submitted: Cosmic dynamics in the era of Extremely Large Telescopes

12. List of targets proposed in this programme

Run	Target/Field	α(J2000)	δ(J2000)	ToT	Mag.	Diam.	Additional info	Reference star
A	Q 0016-357	00 18 40.80	-35 29 11.9	11	r=15.21		z=3.1900	
A	PKS 0858-279	09 00 40.07	-28 08 23.6	4	g=14.56		z=2.1520	
A	B 1422+231	14 24 38.11	+22 56 00.3	26	i=15.99		z=3.6200	
A	PG 1247+268	12 50 05.72	+26 31 07.5	7	g=15.01		z=2.0420	
A	KP 1623.9+26.8	16 25 57.76	+26 44 44.2	14	r=15.54		z=2.6070	
A	B2 1225+31	12 28 24.96	+31 28 37.5	10	r=15.32		z=2.2190	
A	Q 0049-3936	00 52 09.28	-39 19 45.7	12	r=15.42		z=2.3000	
A	Q 0031-3939	00 33 32.66	-39 22 45.4	11	r=15.40		z=2.2000	
A	TON 1530	12 25 27.42	+22 35 13.2	10	g=15.43		z=2.0580	
A	Q 0158-3731	02 00 23.35	-37 17 23.9	13	r=15.54		z=2.2500	
A	CTS A33.02	05 54 45.80	-33 05 17.1	15	r=15.71		z=2.3600	
A	PSS J0926+3055	09 26 36.32	+30 55 04.9	66	i=16.96		z=4.1900	
A	HE 0940-1050	09 42 53.50	-11 04 26.2	31	r=16.36		z=3.0540	
A	1208+1011	12 10 57.04	+09 54 27.0	56	i=16.80		z=3.8030	
A	Q 0038-3416	00 41 15.10	-34 00 12.7	15	r=15.69		z=2.2000	
A	SDSS J120006.25+312630.8	12 00 06.25	+31 26 30.9	32	r=16.40		z=2.9889	
A	CTS C15.05	23 50 34.26	-43 25 59.7	29	r=16.33		z=2.8850	
A	Q 0000-26	00 03 22.95	-26 03 18.4	76	i=17.11		z=4.1110	
A	Q 1101-264	11 03 25.30	-26 45 15.8	15	g=15.87		z=2.1450	
A	Q 0025-4047	00 27 42.83	-40 31 00.8	16	g=15.94		z=2.1800	

Target Notes: All targets were chosen to lie south of +35°. The exposure time is in hours and refers to the time requested per year. To obtain the total exposure time multiply the given value by 8 (i.e. the plan is to perform identical observations for 4 years for each of two epochs). The total exposure time of 3750 hours was distributed among the targets so as to achieve similar S/N for all of them.

12b. ESO Archive - Are the data requested by this proposal in the ESO Archive (http://archive.eso.org)? If yes, explain why the need for new data.

No

13. Scheduling requirements

14. Instrument configuration

Period	Instrument	Run ID	Parameter	Value or list
80	UVES	A	BLUE	Standard setting: 346