SUPPLEMENTARY INFORMATION

doi:10.1038/nature12184

Here we give details of the multi-wavelength observations (§ 1), the photometry deblending procedures (§ 2), the weak lensing model (§ 3), and the estimation of the physical properties of HXMM01 (§ 4).

Throughout we adopt a Chabrier[31] initial mass function (IMF), and the concordance ΛCDM cosmology[32] with $\Omega_{\rm m} = 0.27$, $\Omega_\Lambda = 0.73$, and $H_0 = 70$ km s^{-1} Mpc^{-1}.

1 Observations and Data Reduction

HXMM01 was initially identified as a strong lensing candidate[5] in SPIRE[33] observations from the *Herschel** Space Observatory[34] as part of the HerMES† survey[6] of the XMM-LSS field. The source has an unusually high 500 μm flux density ($S_{500} = 132$ mJy), which suggests gravitational lensing of a fainter background galaxy[35].

Along with other lensing candidates[5], we observed HXMM01 with the Submillimeter Array[36] to refine its position and to examine if its morphology is consistent with gravitational lensing. The *Herschel* source was resolved into two similarly bright SMA sources separated by only $\sim 3''$ and close to two low-redshift galaxies, making it a high-priority target for further followup observations.

Guided by the far-IR photometric redshift, the spectroscopic redshift of HXMM01 was first determined from a blind search for CO $J = 3 \rightarrow 2$ emission with the Combined Array for Research in Millimeter-wave Astronomy (CARMA). Subsequently, the redshift was confirmed at IRAM/Plateau de Bure Interferometer (PdBI) with a measurement of CO $J = 4 \rightarrow 3$ emission. We also obtained single-dish and interferometric CO $J = 1 \rightarrow 0$ data with the National Radio Astronomy Observatory‡'s Robert C. Byrd Green Bank Telescope (GBT) and Karl G. Jansky Very Large Array (JVLA)[37]. Finally, with Keck, we obtained a near-IR spectrum covering the Hα emission of HXMM01 and optical spectra of the nearby foreground galaxies to determine their redshifts. The Keck $K_{\rm S}$-band adaptive optics image and the *HST*/WFC3 F110W image further revealed the clumpy rest-frame optical morphology of both SMA sources. In the following, we describe the observations and data reduction procedures in more details.

**Herschel* is an ESA space observatory with science instruments provided by European-led Principal Investigator consortia and with important participation from NASA.

†hermes.sussex.ac.uk

‡NRAO is operated by Associated Universities Inc., under a cooperative agreement with the National Science Foundation.

1.1 Herschel Far-IR Imaging

The original *Herschel* SPIRE imaging data for this source, as part of the HerMES XMM-LSS field observations[6], are now publicly available from the *Herschel* Science Archive. The observational identifications (Obs Ids) of those datasets are 1342189003 and 1342190312. As the parallel-mode Photodetector Array Camera and Spectrometer (PACS)[38] data at 100 and 160 μm is shallow, we then obtained deeper scans of HXMM01 with integration times of 840 s at these wavelengths using part of the remaining SPIRE team's guaranteed time (Obs Ids: 1342258806 and 1342258807). Finally, a 70 μm scan map of HXMM01 with an integration time of 276 s was obtained as part of an OT2 open-time program (Obs Ids: 1342261958 and 1342261957). The SPIRE data are reduced and analyzed following the standard procedures[39]. We used the high fidelity maps produced using an internal map-making package[40]. The PACS data are reduced with HIPE.

1.2 Submillimeter Imaging

We obtained high-resolution interferometric imaging of HXMM01 at 876.5 μm (342.02 GHz) with the SMA (PIs: Gurwell, Bussmann). HXMM01 was observed in the subcompact array configuration on 2010 Aug 14 with an on-source integration time ($t_{\rm int}$) of 58 min, in the extended configuration on 2010 Sep 26 with $t_{\rm int} = 141$ min, and in the very extended configuration on 2011 Jan 4 with $t_{\rm int} = 54$ min. Atmospheric opacity and phase noise for all observations was very low ($\tau_{225\ \rm GHz} = 0.07$, $\phi \approx 20$ deg). All observations used an intermediate frequency (IF) coverage of 48 GHz to provide a total of 8 GHz bandwidth (considering both sidebands). The quasars 0217+017 and 0238+116 were used for time-variable gain (amplitude and phase) calibration. 3C 454.3 served as the primary bandpass calibrator. We used a combination of Uranus and Callisto as the absolute flux calibrators.

We used the MIR software package in Interactive Data Language (IDL) to calibrate the visibility data. For imaging, we used the Multichannel Image Reconstruction, Image Analysis, and Display (MIRIAD) software package with natural weighting for maximum sensitivity. The final CLEANed image has a synthesized beam with a FWHM resolution of $0.52'' \times 0.43''$ at a position angle (PA) of 55° east of north and an root-mean-square noise level of 0.67 mJy beam^{-1}. The primary beam of the SMA is $35''$ FWHM at 342 GHz.

The final SMA map resolved HXMM01 into two major components separated by $2.95''$ (25 kpc) along with filamentary structures connecting the two (Figure 1). The northern component (X01N) has a flux density of 9.3 ± 0.7 mJy, while the southern component (X01S) has a flux density of 8.3 ± 0.7 mJy. The filaments have a total flux density of 9.4 ± 1.0 mJy, accounting for $\sim$35% of the entire system (27.0 ± 1.4 mJy).

1.3 Near-IR Imaging

We obtained 40×80 s K_S-band images on 2012 February 4 (UT) with the Keck II laser guide-star adaptive-optics system[41] (PI: Cooray). We used the NIRC2 camera at 0.04″ pixel^{-1} scale (40″ field), and we dithered with 2−3″ steps in a nine-point dithering pattern. The atmospheric seeing at 0.5 μm was ∼0.8″. An $R = 17.6$ magnitude star 67″ SW of HXMM01 served as the tip-tilt reference star. The estimated Strehl ratio at the source position is ∼14% under normal conditions. Because of the large separation from the tip-tilt star, we had to rotate the camera to PA = 121.2° and offset HXMM01 to the upper right quadrant of the detector to keep the tip-tilt star within the range available to the field steering mirrors.

The images were reduced following the standard procedures[42]. Before individual frames were finally combined, the camera distortion was corrected using the on-sky distortion solution from observations of the globular cluster M 92[§]. Five frames were rejected because of poor AO corrections, so the final image combined 35×80-s frames. It has a FWHM resolution of 0.18″, as measured from the stellar source located 10″ W of HXMM01. The image was flux and astrometry calibrated against the K_S-band image from the VISTA Hemisphere Survey (VHS). The K_S-band image is shown in Figure 1. To highlight HXMM01, we added contours to the Keck image after subtracting the foreground galaxies.

HXMM01 was observed as part of our *Hubble* Space Telescope (*HST*) Cycle 19 SNAP Program (ID: 12488, PI: Negrello). Four images of 63 s were taken in the F110W filter by the Wide Field Camera 3 (WFC3) on 2012 September 4 (UT). We applied a four-point dithering pattern with a parallelogram primary pattern shape, a point spacing of 0.572″, and a line spacing of 0.365″. After passing the images through the standard WFC3 pipeline, we used the Image Reduction and Analysis Facility (IRAF) MULTIDRIZZLE package on the four frames for distortion correction, image registration, cosmic-ray rejection, and image combination. A square interpolation kernel was used in the drizzling process and the final reduced image has a pixel scale of 0.06″ and a FWHM resolution of 0.24″.

Finally, we obtained seeing-limited J and K_S-band images of HXMM01 with the Long-slit Intermediate Resolution Infrared Spectrograph (LIRIS), which is mounted on the Cassegrain focus of the 4.2-m William *Herschel* Telescope (WHT) (PI: Perez-Fournon). The observations took place on 2011 January 15 (UT). We took 90×20-s exposures in K_S-band and 135×40-s exposures in J-band. The images were reduced with the IRAF LIRIS data reduction package LIRISDR[¶]. The FWHM resolutions are 0.9″ and 0.8″ for the final J and K_S-band images. The astrometry was calibrated against 2MASS. The images were flux calibrated against the public VHS J and K_S-band images of the same region. The WHT images are much deeper than the VHS images. The two components of HXMM01 are easily separated in the WHT images and are detected at 4−9σ.

[§]http://www2.keck.hawaii.edu/inst/nirc2/forReDoc/post_observing/dewarp/

[¶]http://www.iac.es/galeria/jap/lirisdr/LIRIS_DATA_REDUCTION.html

Besides the morphology information from the high-resolution Keck and *HST* images, all of these images were used to obtain photometry for both the foreground galaxies and HXMM01 so that we can study their spectral energy distributions (SEDs). Wherever detected, the two components of HXMM01 show consistent configurations in different images. However, we do not detect the submillimeter and CO $J = 1 \rightarrow 0$ filaments between the two main components in these images. X01S appears much bluer in F110W$-K_{\rm S}$ than X01N, which is a result of the "clump" that we identified at the southern end of X01S (Figure 1). The clump is also detected in the archival Canada-France-Hawaii telescope optical images, and its SED is consistent with either a less obscured galaxy at $z = 2.3$ or a physically unrelated contaminating source.

1.4 Millimeter Photometry

Millimeter continuum is useful to constrain Rayleigh-Jeans tail of the thermal dust emission. We observed HXMM01 at 1.2 mm with the Max-Planck Millimetre Bolometer 2 (MAMBO2[43]) at the IRAM 30 m telescope on 2011 January 19 and 22 as part of the pool observations, achieving a 1σ sensitivity of $\sim$0.9 mJy (PI: Perez-Fournon). HXMM01 was clearly detected with a flux density of 11.2 ± 0.9 mJy. We note that some MAMBO2 observations taken around the same time could have been affected by a "drift" signal of $\sim$5 mJy, but it is unlikely that the HXMM01 observations were influenced, because (1) we obtained consistent measurements from two scans that were carried out three days apart and (2) the 1.2 mm flux density is consistent with the best-fit modified blackbody from the photometry at other far-IR/millimeter wavelengths (see § 4.2). We also detected the continuum of HXMM01 at 2.1 mm with a flux density of 1.2 ± 0.2 mJy as part of the CO $J = 4 \rightarrow 3$ observations with the PdBI (see § 1.8).

1.5 Archival X-ray-to-Radio Imaging

HXMM01 was observed on 2006 July 4, 2006 July 22, 2008 July 30, and 2009 Jan 1 by *XMM-Newton* as part of the Large Scale Structure Survey[44]. Two pointings (XMM-LSS_42 and 45) in AO5 covered HXMM01. We combined the MOS observations that contains HXMM01 and obtained a stacked image with a total integration time of 85 ks. It is unfortunate that HXMM01 happened to lie on the edges of both pointings where the sensitivity is low. No source is detected at the position of HXMM01. In § 4.9, we provide upper limits on the X-ray fluxes and luminosities.

To build the panchromatic SEDs of HXMM01 and the foreground galaxies, we obtained imaging data from the Galaxy Evolution Explorer[45] (*GALEX*, FUV, NUV), the Canada-France-Hawaii Telescope Legacy Survey Wide[46] (CFHTLS-Wide, u^*, g', r', i', and z'), the VHS (J, H, and $K_{\rm S}$), the *Spitzer* Wide-area Infrared Extragalactic Survey[47] (SWIRE; 3.6, 4.5, 5.6, 8.0, 24, 70, and 160 μm), the Wide-field Infrared Survey Explorer all-sky survey[48] (*WISE*; 3.4, 4.6, 12 and 22 μm), and the FIRST survey[49] (21 cm), in addition to our *HST* (F110W), WHT (J, and $K_{\rm S}$), Keck/NIRC2 ($K_{\rm S}$), *Herschel* PACS and SPIRE (70 to 500 μm), SMA (880 μm), MAMBO2 (1.2 mm), and IRAM/PdBI images (2 mm). In § 2, we describe how we obtained the deblended

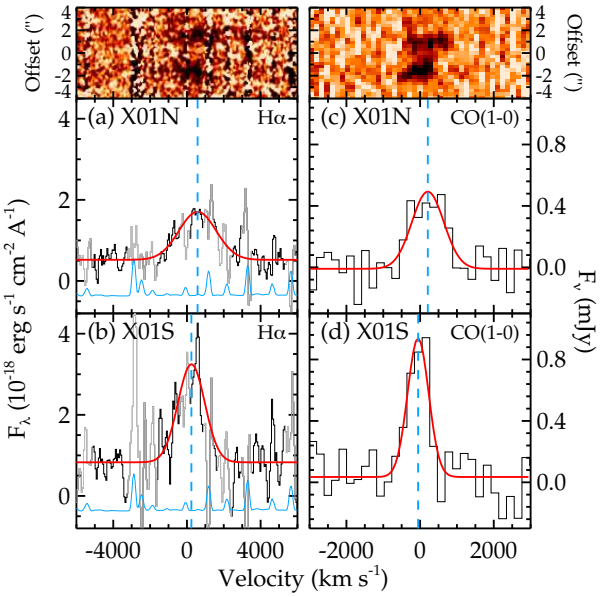

Figure S1 – Keck/NIRSPEC near-IR and JVLA centimeter spectra of HXMM01. The velocities were computed against the systemic redshift at $z = 2.308$. On the top, we use the 2D spectra to show the velocity offset between the two components of HXMM01. The vertical axis is along the spatial direction and NNE is up (slit PA = 24.3° and 22° for the Hα and CO $J = 1\rightarrow0$ spectra, respectively). The CO $J = 1\rightarrow0$ 2D spectrum is extracted from the JVLA data cube using a 1.2″ pseudo-slit centered on the two main components. The middle and bottom left panels show the Hα spectra of X01N and X01S, along with the best-fit Gaussians (red curves). Regions affected by strong sky lines are plotted in light grey. A scaled sky spectrum is shown in blue. The dashed vertical lines indicate the center of the best-fit Gaussian. The right panels show the CO $J = 1\rightarrow0$ spectra along with the best-fit Gaussians (red curves).

photometry for HXMM01 and the foreground galaxies.

1.6 Redshift from CO $J = 3\rightarrow2$

We used CARMA to search for CO $J = 3\rightarrow2$ emission towards HXMM01 in the 3 mm band[50] (PI: Riechers). The search was guided by the photometric redshift of 2.7 ± 0.5 (Dave Clements) obtained from fitting the 250–1100 μm SED. Observations were carried out under good to excellent 3 mm weather conditions during 4 tracks between September 07 and 10 in 2010 in the D array configuration (6–127 m baselines), yielding an on-source (total) observing time of 7.8 hr (13.3 hr). Each of the tracks targeted a different frequency setting, using the 3 mm receivers and the CARMA spectral line correlator with an effective bandwidth of 3.7 GHz per sideband (IF range: 1.2–4.9 GHz) at 5.208 MHz (15.6 km s^{-1} at 100 GHz) spectral resolution, yielding a contiguous frequency coverage of 85–112 GHz to search for redshifted CO line emission. The nearby quasar

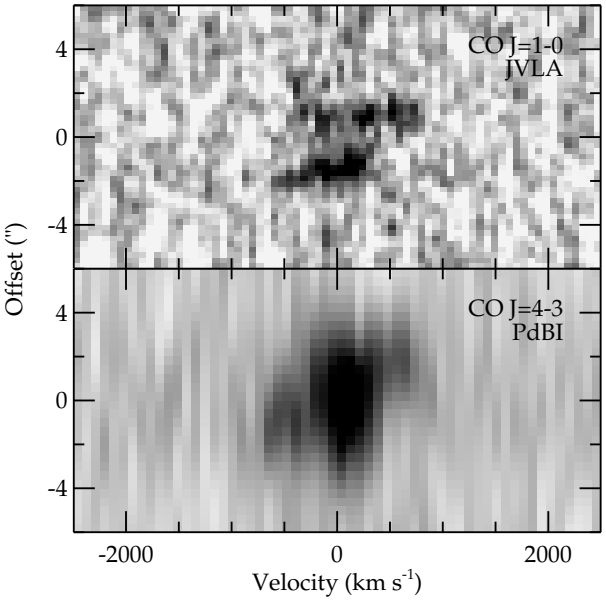

Figure S2 – JVLA and PdBI CO spectra of HXMM01. The velocities were computed against the systemic redshift at $z = 2.308$. The CO $J = 1 \rightarrow 0$ 2D spectrum at the top is extracted from the JVLA data cube using a 1.2″ pseudo-slit at PA = 22°, while the CO $J = 4 \rightarrow 3$ spectrum at the bottom is extracted from the PdBI data cube using a 2.4″ pseudo-slit at the same PA. Both panels clearly show the different velocity structures of the two components, strengthening the findings in Fig. S1.

J0239–025 was observed every 15 minutes for complex gain calibration. Pointing was performed at least every 2–4 hr on nearby stars and radio quasars, using both optical and radio modes. The bandpass shape was derived from observations of the bright quasar 3C 84. Absolute fluxes were bootstrapped relative to Uranus. The resulting calibration is estimated to be accurate within $\sim$15%.

We used the MIRIAD package for data reduction and analysis. The combined data were imaged using "natural" baseline weighting, resulting in a synthesized clean beam size of $5.6'' \times 4.2''$ (PA 149°) and a rms noise level of 0.75 mJy beam^{-1} per 405.4 MHz (1162 km s^{-1}) frequency bin.

A systematic inspection of the data cube both in the uv and image planes yields a single, broad emission feature close to the phase center at the position of the SMA continuum emission. We interpret this feature as the CO J = 3→2 line emission ($\nu_{\rm rest}$=345.7959899 GHz, redshifted to $\nu_{\rm obs} = 104.56 \pm 0.03$ GHz) towards HXMM01, yielding a redshift of $z = 2.3073 \pm 0.0010$. We subsequently confirmed this redshift by detecting CO J = 1→0 line emission with the GBT and JVLA (see next subsection) and CO J = 4→3 line emission with the IRAM/PdBI at the same redshift (see § 1.8).

1.7 CO J = 1→0 Spectroscopy

We obtained single-dish CO J = 1→0 data at the 100 m diameter GBT for 4.85 hr on 2010 November 21 and 2.10 hr on 2011 January 14 (Program ID: GBT09C-70; PI: Harris). We used the Zpectrometer cross-correlation spectrometer[51] attached to the facility Ka-band correlation receiver. With instantaneous frequency coverage from 26.5 to 37.7 GHz, the spectrometer covers redshifts $2.1 \leq z \leq 3.5$ for the 115.27 GHz CO J = 1→0 transition. Over the spectrometer's band the FWHM beam size ranged from 27″ to 16″, and the correlator's 20 MHz bandwidth corresponded to a spectral resolution of 234 to 157 km s^{-1}. Cross-correlation produces spectra that are the difference in power between the receiver's two beams on the sky, which are separated by 78″. The correlation receiver architecture, electrical phase switching, and optically switching the source between the receiver inputs with sub-reflector motion on a 10 sec cycle greatly reduces electronic and atmospheric fluctuations. We switched the telescope between two targets close on the sky on an 8 min cycle to compensate for optical imbalance between the two beams. We established the absolute flux scale with hourly pointing on quasar 0108+0135; spectra of this source before and after pointing also provided average corrections for pointing drift and atmospheric transmission. The flux density of 0108+0135 was 1.8 Jy during the observations, calibrated against 3C 48, whose flux density is 0.74 Jy at 34.8 GHz, following the Astronomical Almanac (2011). The CO J = 1→0 line from HXMM01 is clearly detected at 34.85 ± 0.01 GHz, corresponding to $z = 2.3074 \pm 0.0008$. The FWHM beam size of the GBT is 22″ at this frequency.

We also obtained high-resolution CO J = 1→0 data with the JVLA (PI: Ivison). Observations were carried out dynamically during good weather conditions in 2012 January (DnC configuration), March (C configuration), and June (B configuration). The bright compact calibration

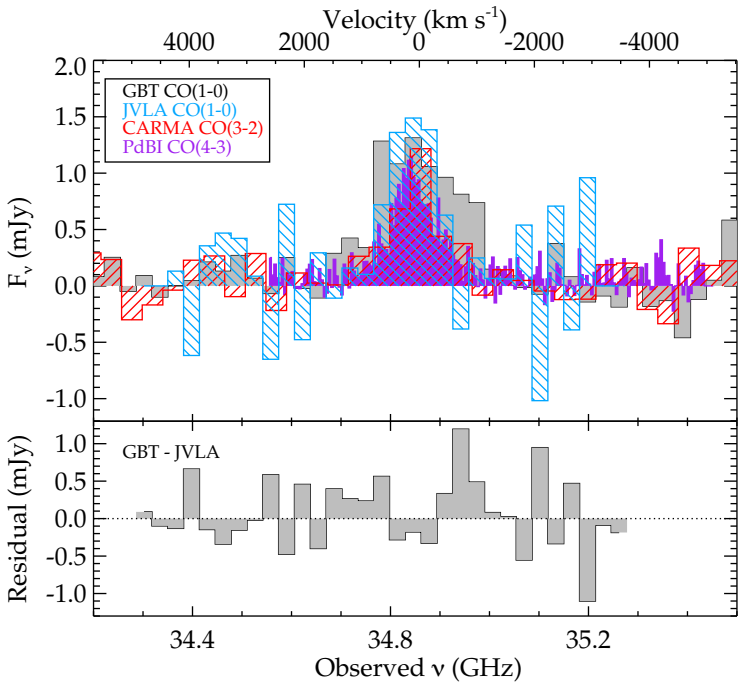

Figure S3 – Comparison of CO line profiles. The PdBI CO $J = 4 \rightarrow 3$ and CARMA CO $J = 3 \rightarrow 2$ lines have been scaled by ν_{1-0}/ν_J in frequency and $(\nu_{1-0}/\nu_J)^2$ in flux density. The top axis shows the velocity relative to the systemic redshift of $z = 2.308$. The JVLA, CARMA, and PdBI spectra are extracted using 6″, 12″, and 15″-diameter apertures.

source, J0241−0815, was observed every few minutes to determine accurate complex gain solutions and bandpass corrections. 3C 48 was also observed to set the absolute flux scale, and the pointing accuracy was checked locally every hour. At the redshift of HXMM01, the receivers gave 1,024 MHz coverage with 512 × 2 MHz channels. The data were reduced using AIPS[11]. Because the AIPS/SETJY task adopts a higher flux for 3C 48 (S = 0.83 Jy at 34.8 GHz; i.e., 12% higher than that assumed for the GBT/Zpectrometer data), we scaled the GBT flux densities by the ratio of the two adopted 3C 48 flux densities (0.83 Jy/0.74 Jy) for a fair comparison. In the end, the channels were imaged over a $512 \times 512\times$ 0.1″ field, with natural weighting (ROBUST = 5), to form a 512^3 cube centered on HXMM01. Integrating over the 80 channels ($\sim 1,400\,\mathrm{km\,s^{-1}}$) that contain line emission yielded an r.m.s. noise level of 25 μJy beam^{-1}.

The CLEANed and velocity-integrated CO J = 1→0 map is shown in Figure 1 as blue contours. The beam is 0.83″×0.77″ at PA = 72°. Similarly to the SMA data, HXMM01 was resolved into two main components separated by ∼2.8″. The northern component shows an extension of low-surface-brightness filament towards the southern component. This filament has a line flux of $S_{\rm CO}\Delta V = 0.20 \pm 0.04$ Jy km s^{-1}, which is 2.5 times lower than that of the northern component. We extracted the CO spectra with 1.2″ and 2.0″-diameter apertures centered on the northern and southern components, respectively. Based on the best-fit Gaussian models of the velocity-integrated map, we then applied aperture corrections of 1.7× and 1.3× to the spectra of the northern and southern components, respectively. The two CO J = 1→0 components have significantly different redshifts and line profiles (Figure S1 and Table S1), indicating that they are two galaxies that are merging, instead of lensed images from a single source. This result is confirmed by the Keck Hα spectra (§ 1.9; see Figure S1) and the PdBI CO J = 4→3 data cube (§ 1.8; see Figure S2).

The JVLA spectrum integrated over a 6″-diameter aperture gives a line width of $\Delta V_{\rm FWHM} = 840 \pm 160\,\mathrm{km\,s^{-1}}$ and a line flux of $S_{\rm CO}\Delta V = 1.7 \pm 0.3$ Jy km s^{-1} at $z = 2.3079 \pm 0.0007$. In comparison, the CO J = 1→0 measurements from the GBT are: $\Delta V_{\rm FWHM} = 1670 \pm 140\,\mathrm{km\,s^{-1}}$ and $S_{\rm CO}\Delta V = 2.3\pm0.3$ Jy km s^{-1} at $z = 2.3074\pm0.0008$. As shown in Figure S3, the discrepancy between JVLA and GBT CO J = 1→0 line intensity is insignificant (0.6 ± 0.4 Jy km s^{-1}), and it mostly occurs in the wings, where the S/N is low. While some flux from the extended wings could have been resolved out in the high-resolution JVLA data, we primarily adopt the JVLA measurements when deriving the properties of HXMM01 to be conservative about its gas mass.

1.8 CO J = 4→3 Spectroscopy and 2 mm continuum

We observed HXMM01 with the PdBI tuned at 139.414 GHz, as part of the millimeter follow-up of the HerMES gravitational lens candidates (PI: Cox). The Widex correlator bandwidth (3.6 GHz) allows covering the CO J = 4→3 emission line and the underlying 2 mm continuum. The observations were carried out in September 2010 in the compact D array configuration with 4 antennae, with a total on-source integration of 3.7 hours. System temperatures during the observations were in the range 70−140 K, and precipitable water vapor around 2 mm. Data analysis was

carried out with GILDAS. Maps were obtained using natural weighting, resulting in a synthesized beam of $8.9'' \times 3.1''$ and PA = 140 deg. The noise level is 0.8 mJy beam^{-1} in 20 MHz channels.

We detected a bright CO $J = 4 \rightarrow 3$ line peaking at a redshift $z = 2.3081 \pm 0.0002$, consistent with the redshift obtained from the CARMA, the JVLA, and the GBT observations. The line is detected at a high significance (30σ; Figure S3). The spatially resolved velocity structure is consistent with that observed in the JVLA data cube (Figure S2), further confirming that HXMM01 is a merger.

In addition to the CO $J = 4 \rightarrow 3$ line, we also detected the continuum at 2.1 mm with a flux density of 1.2 ± 0.2 mJy.

1.9 K-band Hα Spectroscopy

We obtained a K-band spectrum of HXMM01 on 2010 November 21 (UT) with the Near Infrared Spectrometer (NIRSPEC[52]) on the Keck II telescope (PI: Casey). We used the low-resolution mode with NIRSPEC-7 (K') filter and a cross-dispersion angle of 35.49°, which gave a wavelength range of $1.96-2.40\,\mu$m. The $0.76''$ ($R \sim 1450$) slit was oriented at a PA of 24.3° east of north to capture both components of HXMM01. We took eight frames of 600 s each (10 coadds of 60 s exposures), adopting the standard "ABBA" dithering sequence with a step of $15''$ along the slit. HIP 34499 ($K_{\rm S}$ = 10.6 AB) was observed using the same setup for telluric correction.

We used the REDSPEC IDL package[‖] to perform spectral and spatial rectification on all of the frames. Most of the sky background and fringing patterns were removed by subtracting the two nodding positions, and the residual sky background was removed by shifting the rectified and median-combined A−B 2D spectrum along the spatial direction by $15''$ and subtract it from the unshifted 2D spectrum. One-dimensional spectra were extracted with $1.8''$-wide apertures. We detected a broad Hα line as well as the continuum in both X01N and X01S. We calibrated the flux of the spectra with the $K_{\rm S}$-band photometry from the WHT $K_{\rm S}$-band image (§ 1.3). In total, the Hα lines contribute to ∼12% ($2.7 \pm 0.4\,\mu$Jy) of the $K_{\rm S}$-band flux of HXMM01. Consistent with the results from the CO lines, the Hα spectra also show different redshifts and line profiles for X01N and X01S (Fig. S1). Table S1 summarizes the measurements from the best-fit Gaussians.

1.10 Optical Spectroscopy of Foreground Galaxies

We obtained optical spectra of the two foreground lensing galaxies (G1 and G2 in Figure 1) on 2011 March 1 (UT) with the Low Resolution Imaging Spectrometer[53] on the Keck I telescope (PI: Bridge). We took two 540 s exposures through a $1.5''$ slit at PA = 275.2° (i.e., aligned with G1 and G2). On the blue side, we used the 600 groove mm^{-1} grism blazed at 4000 Å; while on

[‖] http://www2.keck.hawaii.edu/inst/nirspec/redspec.html

the red side, we used the 400 groove mm^{-1} grating blazed at 8500 Å tilted to a central wavelength of 7779 Å. The 560 nm dichroic was used. The wavelength coverages are 3000−5600 Å and 5400−10200 Å, and the spectral resolutions are 6 and 10 Å (FWHM) for the blue and red, respectively. The spectra were taken at an airmass of ∼1.75, and the atmospheric seeing was ∼0.6″ at 0.5 μm.

The data were reduced and extracted with IRAF. From the [O III] $\lambda\lambda$4959,5007 lines of G1, and Hα and [N II]$\lambda\lambda$6548,6583 lines of G2, we determined their redshifts of $z_{\rm G1} = 0.6546 \pm 0.0001$ and $z_{\rm G2} = 0.5020 \pm 0.0001$, respectively.

2 Deblending the Photometry

Because of the proximity of HXMM01 to the foreground galaxies on the sky, we need to deblend the various sources in some of the images to obtain robust photometry. Between the $u*$-band and IRAC 8.0 μm, we model the various components in individual images with GALFIT[54]. We use high-S/N PSF stars near HXMM01 that are unsaturated and relatively isolated. For images between the $u*$-band and the $K_{\rm S}$-band, we use a star 10″ south-southwest of HXMM01. For the IRAC images, we use a star 7.7′ north of HXMM01. We model the foreground galaxies as Sérsic profiles. The foreground galaxy G2 shows more complex structures (e.g., the tail to the SW) than G1, so we include multiple Sérsic profiles with their relative positions fixed to those determined from the high-resolution *HST* image. Instead of masking out HXMM01, we include multiple Gaussians to approximate its clumpy morphology and fit them simultaneously with the foreground galaxies. We obtain reasonably good fit with $\chi^2/{\rm dof} \sim 1.0$ for all of the images except for the *HST*-F110W image ($\chi^2/{\rm dof} \sim 1.4$). Although the *HST* image is the most challenging to model because of the detailed structures of the foreground galaxies, the photometry is very robust because the components do not overlap as much as the seeing-limited images and IRAC images.

The total fluxes of G1 and G2 are obtained from the best-fit models. We then subtract the models of G1 and G2 from the original images to reveal emission from HXMM01 (Figure S4). Finally, we measure the fluxes from the residual images with two apertures placed around the northern and southern components. As shown in Figure S4, the apertures are large enough to enclose all of the detectable flux, yet are small enough to avoid including too much background noise. It is clear that X01S is much brighter than X01N in the optical bands but the two are similarly bright in the near-IR, suggesting either less dust obscuration in X01S or that the southern clump of X01S is not part of the system (Figure 1). Since we cannot separate the two parts of X01S at wavelengths other than the *HST*/F160W and Keck/$K_{\rm S}$-band, we decided to include both in the photometry. Future *HST* imaging and spectroscopy are needed to reveal the nature of the southern clump of X01S.

GALEX barely detected one source at the position of HXMM01 in the NUV band. Considering the spatial resolution in the NUV band is 6″ in FWHM, G1, G2, and HXMM01 are blended

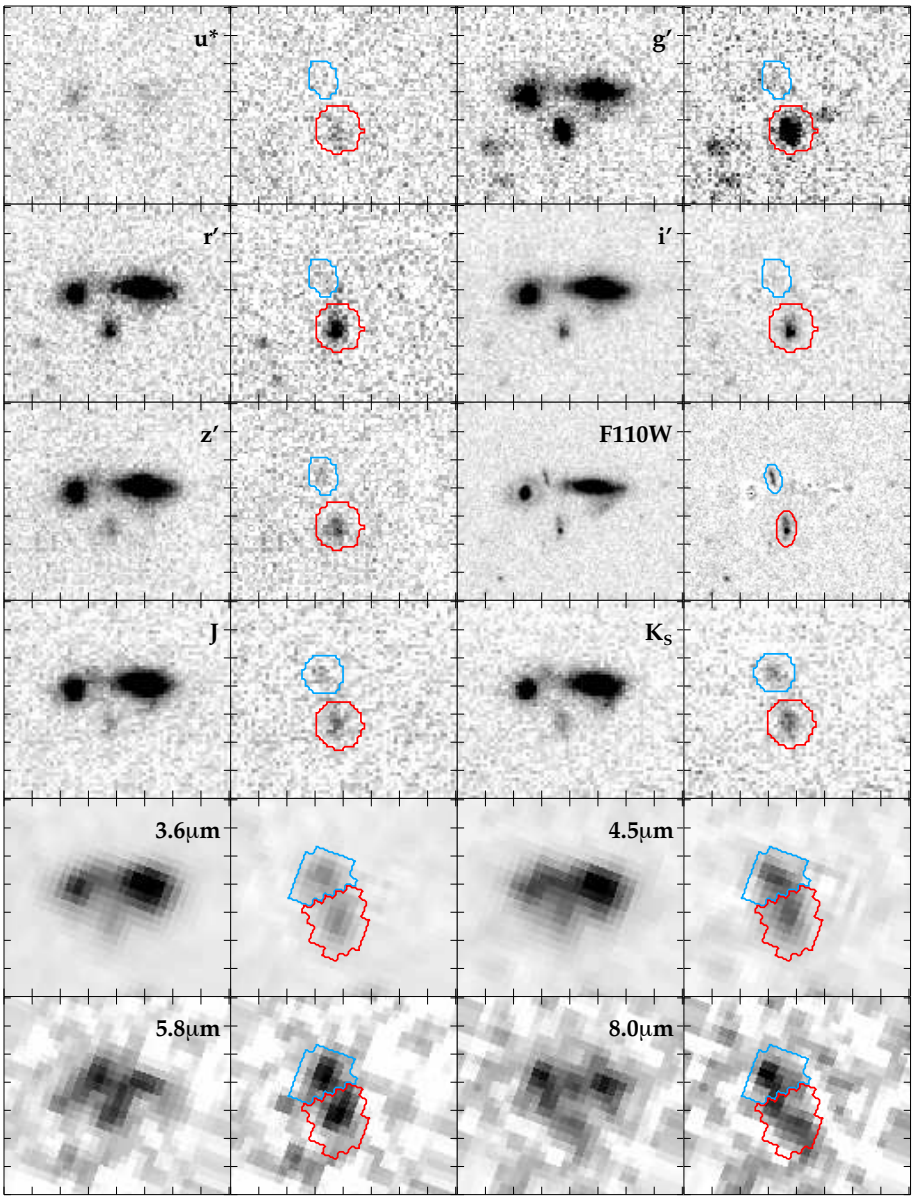

Figure S4 – Deblending the photometry from the $u*$-band to 8 μm. All images are $16'' \times 14''$ and are aligned in astrometry so that N is up and E is to the left. The tickmarks are spaced at intervals of $2''$. For each filter, the original image is shown on the left and the residual image after subtracting the two foreground galaxies is shown on the right. The apertures used for our photometry are outlined in the residual images (blue – X01N; red – X01S).

together. The source has a total NUV flux density of $0.67 \pm 0.18\,\mu$Jy and is undetected at FUV. We can safely assume that the contribution of HXMM01 is negligible because both filters are shorter than the Lyman break at $z = 2.31$. Unable to deblend the flux between G1 and G2, we assume the 3σ upper limit ($0.54\,\mu$Jy) at NUV for all of the components. We do not use the FUV upper limits because of its limited constraints on the SED.

Deblending is unreliable for the *WISE* images because of the low S/N and resolution. The best-fit SED models of G1 and G2 (see § 3.1) suggest that most of the 11 μm emission and $\sim$25% of the 22 μm could arise from the polycyclic aromatic hydrocarbons (PAHs) and dust continuum of the foreground galaxies, although the models have large uncertainties due to the lack of constraints at long wavelengths. Considering the large uncertainties involved in estimating the foreground contamination at 11 μm and that the other three *WISE* channels overlap and are consistent with higher quality *Spitzer* data, we opt not to use the *WISE* data for the SED modeling.

At wavelengths longer than $\sim$20 μm, deblending is impossible because of the limited spatial resolution of the available data, except at 880 μm where we have the high resolution SMA data. However, we do not need to subtract the foreground galaxies here, because HXMM01 should be dominating the total flux as the SEDs of the foreground galaxies rapidly decline at these wavelengths. Indeed, the foreground galaxies are not detected in our SMA image. HXMM01 is resolved in the SMA image (Figure 1), so we can directly measure the 880 μm fluxes of X01N and X01S. Without high-resolution data at other wavelengths, we assume the fluxes can be divided in the same proportion among the three components as at 880 μm — 34% in X01N, 31% in X01S, and 35% in the filaments (see § 1.2). Clearly, we have assumed that the dust temperatures from the three components are the same, which may not be true in reality.

We opt not to use the SWIRE 160 μm data because HXMM01 is severely blended with another brighter source to the NW. And due to the complex structures in the background of that image, we cannot extract a reliable flux for HXMM01. Instead, we use the PACS 160 μm data that is much deeper and has much better spatial resolution.

We did not obtain useful measurements of HXMM01 in the VHS H-band because of image artifacts and at MIPS 160 μm because of blending and complex structures in the background. HXMM01, G1, and G2 are undetected at FUV, NUV, MIPS 70 μm, and 21 cm. The MIPS 70 μm upper limit is consistent with the 2σ detection made with the PACS 70 μm image.

We list the deblended photometry in Tables S2 and S3. In the photometric errors, we have included the following flux calibration uncertainties: 3% for CFHTLS[46], 2% for VISTA and WHT[55], 3% for IRAC[56], 4% for MIPS 24 μm[57], 5% for MIPS 70 μm[58], 5% for PACS 70, 110 and 160 μm, 7% for SPIRE[59] (confusion noises are also included in the total errors for PACS and SPIRE[60]), 10% for SMA, and 15% for MAMBO2 and PdBI.

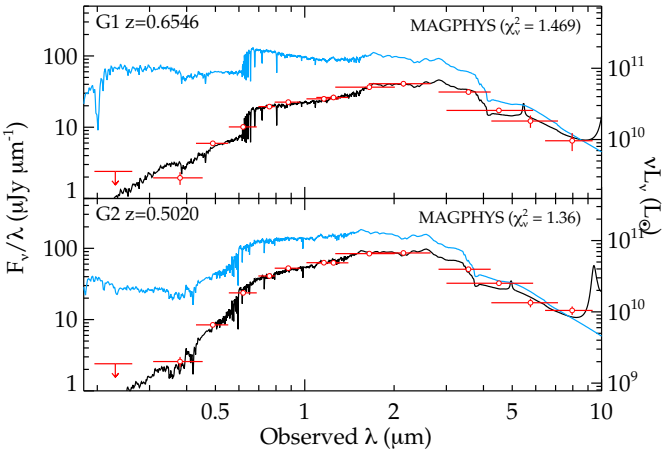

Figure S5 – SEDs of the foreground galaxies G1 (top) and G2 (bottom). The best-fit stellar population synthesis models are plotted as black solid curves, and the corresponding intrinsic models without dust extinction are plotted in blue. Labeled are the spectroscopic redshifts that we obtained from the Keck/LRIS and the reduced χ^2 values of the best-fit models.

3 Lens Modeling

The observed configuration and morphology of HXMM01 do not provide constraints on the lensing potential, except that we can rule out models that produce multiple images of a single source (i.e., the strong lensing cases). To constrain the lensing magnifications that the foreground galaxies impose on HXMM01, we need to estimate the dark matter halo masses of G1 and G2 from their stellar masses. The standard method to measure stellar masses is to fit the observed UV-to-near-IR SEDs with stellar population synthesis models.

3.1 Stellar and Halo Masses of the Lensing Galaxies

We use the MAGPHYS software[9] to fit evolutionary population synthesis models to the SEDs. We used the default 25,000 model spectra computed using the 2007 version of the Bruzual and Charlot stellar population synthesis code[61]. The model spectra were computed at ages between 0.1 and 13.5 Gyr for various metallicities and dust absorption optical depths, assuming a Chabrier IMF[31] and exponentially declining or constant star formation histories with random bursts. Models with ages greater than the age of the Universe at the redshift of the object are excluded. The software treats dust attenuation of young and old stars separately, following a two-component model[62] assuming that stars younger than 10^7 yr are embedded in their birth clouds so these stars are attenuated by dust in both the birth cloud and the interstellar medium (ISM). It then consistently computes the re-radiated mid-to-far-IR dust emission in stellar birth clouds and the ambient diffuse ISM, based on the power of the attenuated stellar emission at shorter wavelengths. A χ^2 and

the corresponding probability $\exp(-\chi^2/2)$ are computed for each model. Finally, a marginalized likelihood distribution is built for each parameter. We adopt the best-fit parameter as the median of the likelihood distribution, which is usually close to the value corresponding to the minimum χ^2, and its $\pm1\sigma$ confidence interval as the 16th$-$84th percentile range.

We fix the galaxies to their spectroscopic redshifts from the Keck/LRIS spectra (§ 1.10; $z_{\rm G1} = 0.6546$ and $z_{\rm G2} = 0.5020$). Both G1 and G2 are undetected in the SMA 880 μm image, so we use the 3σ upper limit at 880 μm to put some constraints on the dust emission, and therefore, on the amount of dust extinction.

Figure S5 shows the best-fit models. Using the 16th, 50th, and 84th percentiles of the marginalized likelihood distributions, we estimated stellar masses of $\log(M_{\rm star}/{\rm M}_\odot) = 10.72 \pm 0.06$ and 10.86 ± 0.10 for G1 and G2, respectively. The stellar masses correspond to halo masses of $\log(M_{\rm halo,G1}/{\rm M}_\odot) = 12.56 \pm 0.26$ and $\log(M_{\rm halo,G2}/{\rm M}_\odot) = 12.81 \pm 0.33$, adopting the $M_{\rm star}-M_{\rm halo}$ relation at their redshifts from abundance matching[30] and including the intrinsic scatter of the relation.

3.2 Magnification Factors

We constrain the lensing magnification factors (μ) in a Monte-Carlo fashion. We start by generating a distribution of 5000 pairs of stellar masses for G1 and G2 that follow the posterior likelihood distributions from MAGPHYS. We assign dark matter halo masses using the $M_{\rm star}-M_{\rm halo}$ relations from abundance matching[30]. The 0.15 dex intrinsic scatter in the $M_{\rm star}-M_{\rm halo}$ relation at fixed halo mass is also included. We adopt the singular isothermal ellipsoid (SIE) density profile for the lenses and assume them to follow the centroid, ellipticity, and PA of the light distribution, as supported by observations[66,67]. For each pair of virial masses (baryonic+dark matter), we compute the image plane magnification map for sources at $z = 2.31$ with LENSTOOL (Figure S6*a*)[68]. We then compute the luminosity weighted magnification factors for HXMM01, X01N, and X01S using the PSF/beam-deconvolved light-distribution models. We compute μ for the $K_{\rm S}$-band, 880 μm, and CO $J = 1\rightarrow0$ images separately because there are significant spatial offsets in the sources across these wavelengths. After excluding the strong lensing cases where models generate more images then observed, the distributions of magnification factors and source-plane projected separations are fit with log-normal functions to find the medians and the 1σ uncertainties (Figure S6*b*).

We find that the total luminosity-weighted magnification of HXMM01 is only about 1.6. The northern component (X01N) is more magnified ($\mu \sim 1.8$) than the southern component (X01S; $\mu \sim 1.4$). The source-plane separation of the two components in $K_{\rm S}$-band is just 2.2″ (19 kpc), which is 35% smaller than the image plane separation (3.4″ = 28 kpc). Table S1 lists the results. We find similar results using the NFW profile[69], although the estimated magnifications are $\sim$20% lower than using the SIE profile. Despite the observed spatial offsets among the gas, dust, and stellar components, their magnification factors agree within 10%. Therefore, we can safely ignore

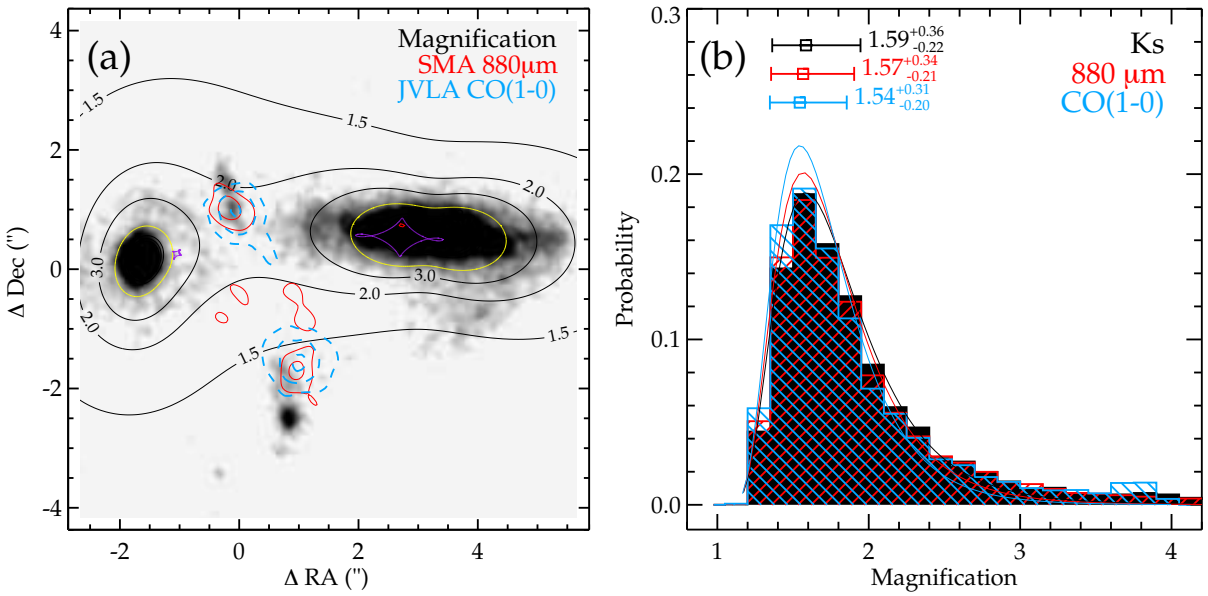

Figure S6 – Magnification map and magnification factors. (a) Similar to Figure 1, but here we overlay the image-plane magnification map as black contours on the *HST*/WFC3 F110W image. The magnification map is computed assuming the median halo masses and SIE profiles. Critical curves are in yellow and caustics are in purple. (b) Distributions of the total magnification factors for the K_S-band (black), the dust continuum at 880 μm (red), and the molecular gas as probed by CO $J = 1 \rightarrow 0$ (blue). On the top, we show their $\pm 1\sigma$ confidence intervals along with the median magnification (boxes with error bars) from log-normal fits (solid curves).

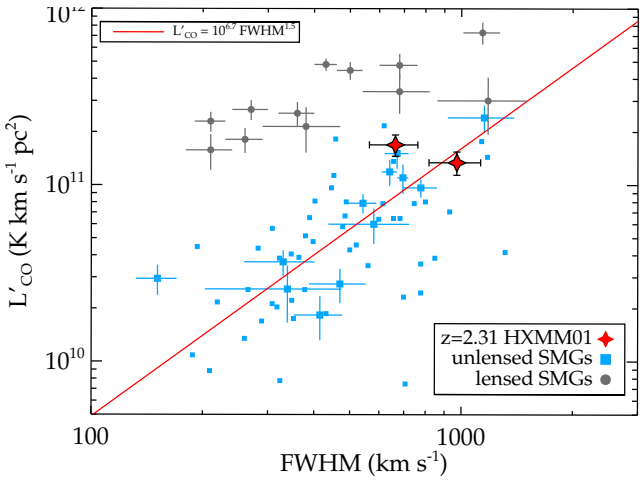

Figure S7 – $L'_{\rm CO}$ **vs. FWHM for CO** J **= 1→0 of lensed and unlensed SMGs**. The two red stars show the measurements of X01N and X01S from the JVLA data cube, which have not been corrected for lensing. The big blue squares with error bars are unlensed and lensing-corrected SMGs with CO $J = 1\rightarrow0$ measurements[11,24,51,63], and the small blue squares are mostly SMGs with higher J CO line measurements converted to CO $J = 1\rightarrow0$ using mean observed ratios[64]. The red line shows the best-fit relation for unlensed SMGs. The grey filled circles with error bars are the GBT CO $J = 1\rightarrow0$ measurements[65] of the brightest lensed SMGs in the H-ATLAS survey. The two main components of HXMM01 lie on the same relation as unlensed SMGs, suggesting that they are mildly ($\mu \lesssim 2$) magnified, consistent with our lens modeling results (§ 3.2).

differential magnification in this source.

As shown in Figure S7, the CO $J = 1 \rightarrow 0$ luminosities and line widths of HXMM01 also suggest only mild ($\mu \sim 1.8$) magnifications from the foreground galaxies, because their CO $J = 1 \rightarrow 0$ luminosities are similar to those of unlensed SMGs with the same line widths.

4 Physical Properties of HXMM01

In this section we determine the physical properties of HXMM01 from the SEDs, the high-resolution CO and dust continuum images, the CO spectra, and the X-ray images. Because the lensing magnifications are small ($\mu < 2$) and have large uncertainties, we list the lens-amplified properties in units of μ^{-1} instead of correcting for μ. Our derived physical parameters are summarized in Table S1 along with the estimated magnification factors, so that readers can easily correct for μ to obtain the intrinsic properties. Note that some parameters are unaffected by lensing — e.g., surface densities, temperatures, and mass or luminosity ratios — so magnification correction is unnecessary.

4.1 Stellar Mass and Star Formation Rate

We modeled the full SEDs of X01N and X01S with MAGPHYS in the same fashion as for the lensing galaxies (§ 3.1), although there are more data points to constrain the dust re-radiated emission. We treat the two components of HXMM01 separately because of their distinct optical-to-near-IR SEDs. As we discussed in § 2, X01N and X01S cannot be separated in the images at wavelengths between 24 μm $< \lambda <$ 1200 μm except at 880 μm, so we had to assume that the two galaxies have the same SED shape at these wavelengths and assign their flux densities based on their proportion in the SMA image — 34% in X01N, 31% in X01S, and 35% in the filaments.

The best-fit SEDs are compared with the photometry in Figure S8a. As expected, the level of dust extinction is high, with a rest-frame V-band optical depth of $\tau_V \sim 3 - 4$ for young stars in birth clouds (i.e., only $\lesssim 5\%$ of the V-band emission leaks out). We find that the total stellar masses are around $(1.6 \pm 0.1) \times 10^{11}\,\mu^{-1}\,\mathrm{M}_\odot$ and $(7.4 \pm 0.2) \times 10^{10}\,\mu^{-1}\,\mathrm{M}_\odot$ for X01N and X01S, respectively. As we noted earlier, the compact clump at the southern end of X01S dominates its emission at wavelengths shorter than the K_S-band (Figs. 1 and S4), making it appear much bluer than X01N. The clump may also be the reason why the model for X01S poorly fits the IRAC data points. Assuming that the clump is not part of X01S and that the intrinsic SED of X01S is similar to that of X01N, we estimate a stellar mass of $(1.7 \pm 0.3) \times 10^{11}\,\mu^{-1}\,\mathrm{M}_\odot$ using the clump-subtracted K_S-band flux ($\sim$8.1 μJy) and the mass-to-light ratio of X01N. In the following we adopt this latter stellar mass estimate for X01S.

By design, the mid-to-far-IR SEDs of the two components are identical. Integrating the best-fit SEDs between rest-frame 8 and 1000 μm, the IR luminosities implies an instantaneous SFR

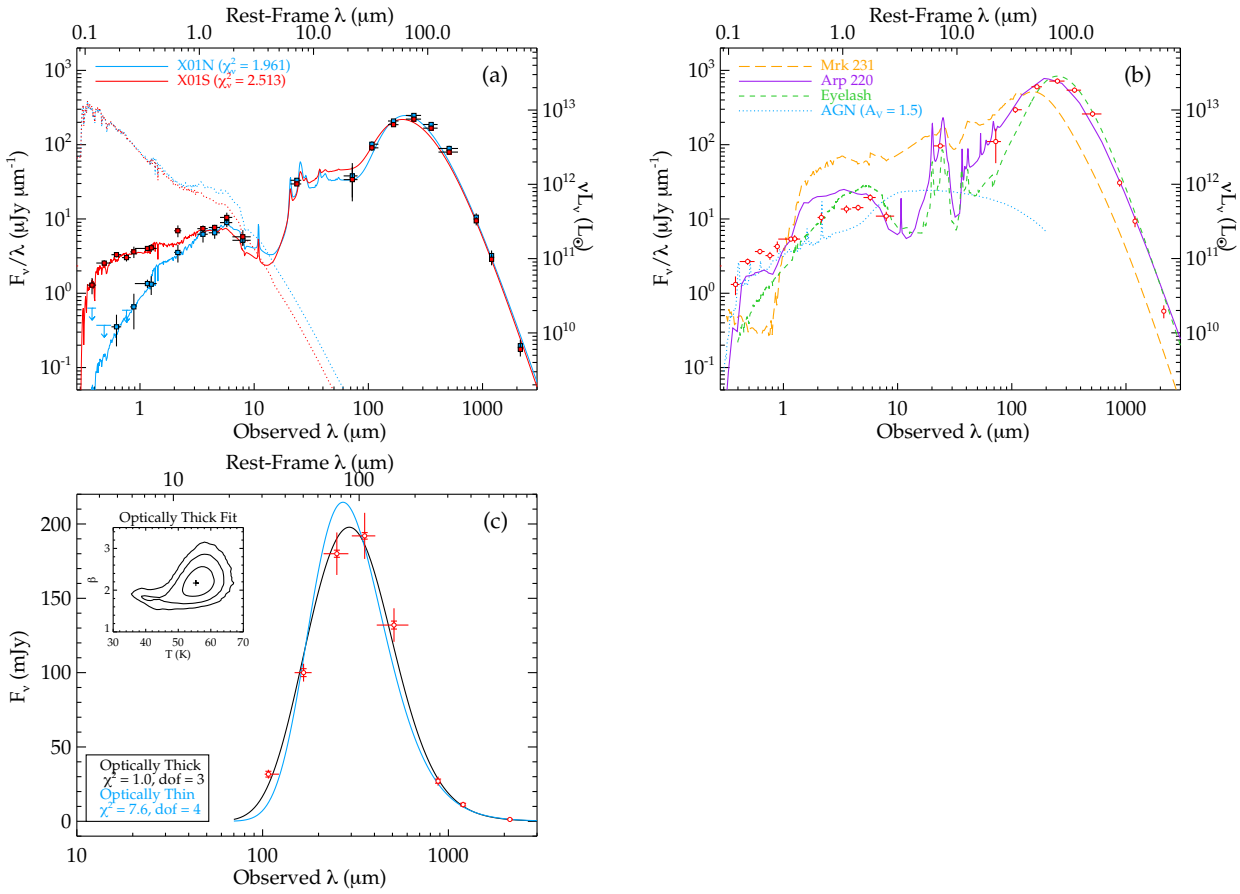

Figure S8 – SED modeling of HXMM01. (a) The blue and red data points are the deblended photometry of X01N and X01S, respectively. Their best-fit MAGPHYS models (blue and red curves) are overlaid to show the quality of the fit. The dotted curves are the dust-free intrinsic (i.e., stellar only) MAGPHYS models. (b) The red data points show the total photometry of HXMM01 including the filament (Table S3). Overplotted are redshifted and scaled templates from the local ULIRGs Mrk 231[70] (orange long dashed curve) and Arp 220[71] (solid purple), the strongly lensed star-forming galaxy "Cosmic Eyelash"[72] at $z = 2.3$ (green dashed), and the AGN composite SED[73] (blue dotted). The AGN template has been scaled to the Hα flux of HXMM01 and reddened[74] by $A_V = 1.5$ magnitude so that it is consistent with the optical SED of HXMM01. (c) Blackbody fit to the far-IR ($160\,\mu\mathrm{m} < \lambda < 2.1$ mm) SED. The optically thick model (black) provides a superior fit than the optically thin model (blue); the corresponding χ^2 values are labelled. The smaller error bars show only the uncorrelated statistical uncertainties, while the full error bars include both confusion noise (for PACS and SPIRE) and systematic flux calibration uncertainties. The inset shows the $1, 2$, and 3σ contours in the $T - \beta$ plane for the optically thick model.

of $1100 \pm 100\ \mu^{-1}\,\mathrm{M}_\odot\ \mathrm{yr}^{-1}$ for each of the components, when converted using the Kennicutt[75] calibration for a Chabrier[31] initial mass function:

$$\mathrm{SFR}/M_\odot\ \mathrm{yr}^{-1} = 10^{-10}\ L_{\mathrm{IR}}/L_\odot \quad (1)$$

4.2 Dust Properties

As the far-IR SED is dominated by thermal dust emission, we fit the photometric data between $160 \leq\ \lambda\ \leq 2{,}100\,\mu$m with a modified blackbody in both the general optically thick form ($S_\nu\ \propto\ (1 - e^{-\tau})\ B_\nu(T)$, where $\tau\ =\ (\nu/\nu_0)^\beta$ and $B_\nu(T)$ is the Planck function) and the optically thin form ($S_\nu\ \propto\ \nu^\beta\ B_\nu(T)$). We treat all components of HXMM01 together because they are spatially resolved only at $880\,\mu$m. We fit for the characteristic dust temperature T_{dust}, the power law slope of dust opacity β, the total far-IR luminosity L_{IR}, and the wavelength $\lambda_0 \equiv c/\nu_0$ below which the emission is optically thick ($\tau \geq 1$). Given that the dust opacity per unit mass follows the same power law as the optical depth, $\kappa_d \propto \nu^\beta$, and a normalization[76,77] of $\kappa_d(125\mu\mathrm{m}) = 2.64 \pm 0.29\ \mathrm{m}^2\ \mathrm{kg}^{-1}$, we can also estimate the dust mass M_{dust} assuming all dusts are at a single temperature T_{dust}.

We use an affine-invariant Markov Chain Monte Carlo (MCMC) sampler (emcee)[78] to compute 500,000 steps from 250 samplers. Our approach fully marginalizes over the optical depth, and takes covariances between the input photometry into account. The autocorrelation length is <50 steps for all parameters, i.e., the chains converge very well. The resulting constraints are shown in Figure S8*c*.

The general optically thick model provides a better fit to the data. Our best fits have $\chi^2 = 1$ for 3 degrees of freedom and $\chi^2\ =\ 7.6$ for 4 degrees of freedom for the optically thick and the optically thin models, respectively. The former is therefore preferred and the best-fit parameters are tabulated in Table S1. The total IR luminosity of $L_{\mathrm{IR}} = (3.2 \pm 0.2) \times 10^{13}\ \mu^{-1}\ \mathrm{L}_\odot$ is enormous. This implies a FIR-derived SFR of $3200 \pm 200\ \mu^{-1}\ \mathrm{M}_\odot\ \mathrm{yr}^{-1}$.

The best-fit dust temperature is ∼55 K, which is much higher than the average 35 K for normal star-forming galaxies and lies at the higher end of the dust temperature distribution[79] of hyper-luminous IR galaxies at $z \sim 2$. This temperature is higher than the previously derived value ($T_{\mathrm{dust}}\ =\ 43 \pm 1$ K)[5], because Wardlow et al. used the optically thin model with β fixed to 1.5, which has been ruled out because of the new PACS photometry. As a reference, our optically thin model with β relaxed gives $T_{\mathrm{dust}} = 36 \pm 2$ K.

High dust temperatures indicate efficient star formation[7,80], i.e., starbursts. Using the Stephen-Boltzmann law, we have

$$T_{\mathrm{dust}}^{4+\beta} \propto L_{\mathrm{IR}}/M_{\mathrm{dust}} \propto \mathrm{SFR}/(Z_{\mathrm{gas}} M_{\mathrm{gas}}) = \mathrm{SFE}/Z_{\mathrm{gas}}, \quad (2)$$

where SFE $\equiv$ SFR/$M_{\rm gas}$ is the global star formation efficiency. Combined with the lack of AGN activity in HXMM01 (§ 4.8), the high dust temperature clearly indicates that HXMM01 is a starburst galaxy. This conclusion is supported by the star formation efficiencies per dynamical timescale estimated in § 4.5.

4.3 Sizes of Star Forming Regions and Gas Reservoirs

Our high resolution data allow us to directly measure the extent of the dusty star forming regions at rest frame 265 μm and the gas reservoirs probed by CO J = 1→0. As both components of HXMM01 are resolved in the SMA and JVLA CO J = 1→0 images, we use Gaussian models convolved with the beam to fit the intensity maps. The source sizes are estimated using the best-fit HWHMs along the major (a) and minor (b) axes ($A = \pi ab$). We find that the total size of the dusty star forming regions ($\Sigma A_{880} \sim 20\,\mu^{-1}$ kpc^2) is about five times smaller than that of the CO J = 1→0 gas reservoirs ($\Sigma A_{\rm CO} = 95\,\mu^{-1}$ kpc^2), similar to the lensed SMG HATLAS12−00 that has high-resolution CO J = 1→0 and dust maps[42].

Since it is blackbody radiation, we can also estimate the cross-section of the dust emitting region using the Stefan-Boltzmann law:

$$A_{\rm SB} = \frac{L_{\rm IR}}{4\sigma T_{\rm dust}^4} = 1.8 \times 10^{-6}\ {\rm kpc}^2\ L_{\rm IR} T_{\rm dust}^{-4} \quad (3)$$

where $L_{\rm IR}$ is in $L_\odot$ and $T_{\rm dust}$ is in K. We estimate a dust cross section of $A_{\rm SB} = 6.3 \pm 1.5\,\mu^{-1}$ kpc^2 using the optically thick model. This cross section based on Stefan-Boltzmann law is three times smaller than the total area measured directly from the 880 μm image, suggesting that the dust filling factor is only about 30%. In other words, a lot of structures have not been resolved, consistent with the low CO brightness temperatures ($T_b = 2-3$ K) that we measured from the JVLA CO J = 1→0 data cube.

4.4 CO→H_2 Conversion Factor

Assuming the dust mass from the optically thick case, $M_{\rm dust} = (2.9 \pm 0.7) \times 10^9\ \mu^{-1}\ M_\odot$, and the Milky Way gas-to-dust mass ratio[81] of 140, we would expect a total gas mass of $(4.1 \pm 1.0) \times 10^{11}\ \mu^{-1}\ M_\odot$. When compared with the CO J = 1→0 luminosity from the JVLA ($L'_{\rm CO} = (4.5 \pm 0.8) \times 10^{11}\ \mu^{-1}$ K km s^{-1} pc^2), we obtain a CO→H_2 conversion factor of $\alpha_{\rm CO,dust} = 0.91 \pm 0.27\ M_\odot$/(K km s^{-1} pc^2).

On the other hand, through hydrodynamical simulations coupled with radiation transfer codes, the mean velocity-integrated CO brightness temperature ($W_{\rm CO}$) from spatially resolved CO J = 1→0 data and gas metallicity ($Z/Z_\odot$) has been used to provide a calibration to estimate $\alpha_{\rm CO}$[22]:

$$\alpha_{\rm CO,N12} = 10.7 W_{\rm CO}^{-0.32} (Z/Z_\odot)^{-0.65} \quad (4)$$

where $W_{\rm CO} = L'_{\rm CO}/A_{\rm CO}$ is in K km s^{-1}, and the resulting $\alpha_{\rm CO,N12}$ is in $M_\odot$/(K km s^{-1} pc^2). Using the sizes measured from the JVLA CO J = 1→0 map and assuming that $Z = Z_\odot$ for SMGs based on previous measurements of gas metallicity from metallicity-diagnostic optical line ratios[82], we estimate that $\alpha_{\rm CO,N12} = 0.71 \pm 0.04$ and 0.88 ± 0.04 for X01N and X01S, respectively. Note that these estimate are unaffected by lensing, because lensing preserves the brightness temperature.

Both of the above methods yield results consistent with the normally assumed value of $\alpha_{\rm CO} \sim 0.8 - 1.0$ for mergers[21,23,83,84]. Such a conversion factor implies tremendous amount of molecular gas and very high gas-to-baryon fractions for both components of HXMM01 (Figs. 2 and 3).

4.5 Star Formation Efficiency

In the form of the Kennicutt-Schmidt law that involves a dynamical timescale[12], the star formation efficiency per dynamical timescale is defined as

$$\epsilon_{\rm SF} \equiv \frac{\Sigma_{\rm SFR}}{\Sigma_{\rm gas}/\tau_{\rm dyn}}. \quad (5)$$

With the measurements from the previous subsections, we can now estimate $\epsilon_{\rm SF}$ using the SFR surface density ($\Sigma_{\rm SFR}$), the gas surface density ($\Sigma_{\rm gas}$), and the dynamical timescale ($\tau_{\rm dyn}$).

Since the half-light radius of a two-dimensional Gaussian is equal to the HWHM, we estimate the surface densities of the SFR and the gas mass using the measured source sizes

$$\Sigma_{\rm SFR} = 0.5{\rm SFR}/A_{880} \quad \text{and} \quad \Sigma_{\rm gas} = 0.5M_{\rm H2}/A_{\rm CO}, \quad (6)$$

There are usually two dynamical timescale estimates — the free-fall timescale ($\tau_{\rm ff}$) and the rotational timescale ($\tau_{\rm rot}$). We estimate the free-fall timescale as

$$\tau_{\rm ff} = \sqrt{R_{1/2}^3/(2GM_{\rm vir})} = 0.74R_{1/2}/\Delta V_{\rm FWHM}, \quad (7)$$

where we have used the virial mass within $R_{1/2}$ ($M_{\rm vir}$)[87,88]

$$M_{\rm vir} = 5\sigma^2 R_{1/2}/G = 2.1 \times 10^5\ M_\odot\ \Delta V_{\rm FWHM}^2 R_{1/2}, \quad (8)$$

where $\Delta V_{\rm FWHM}$ is in km s^{-1} and $R_{1/2}$ in kpc. Note that some authors[3,17] used a virial coefficient greater than 5 to account for non-virialized systems, giving a formula $M_{\rm vir} = 2.8 \times 10^5\ M_\odot\ \Delta V_{\rm FWHM}^2 R_{1/2}$. To be consistent, we have corrected their dynamical masses to our scaling in Figure 3.

On the other hand, the rotational timescale is defined as

$$\tau_{\rm rot} = R_{1/2}/v_c = 1.36R_{1/2}/\Delta V_{\rm FWHM}, \quad (9)$$

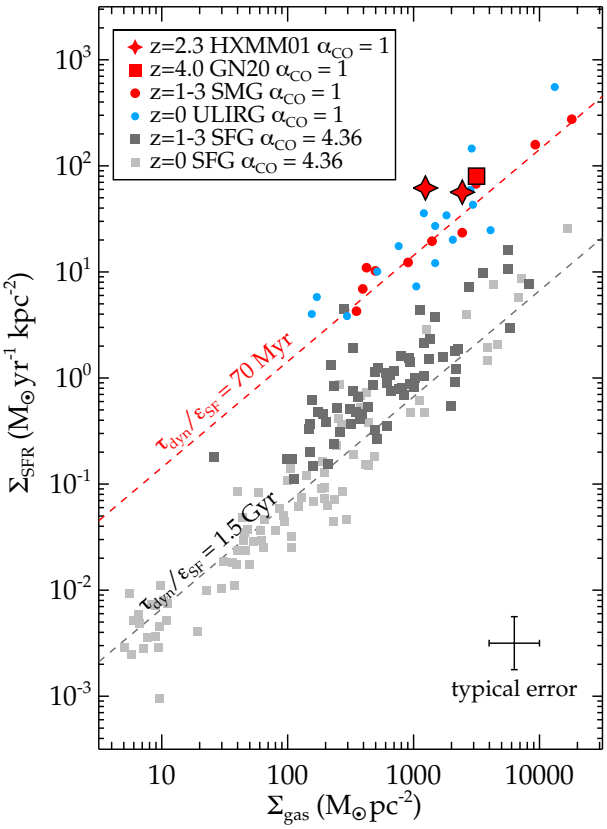

Figure S9 – Star-formation-rate surface density vs. molecular gas surface density for quiescently star-forming and starburst galaxies. The two components of HXMM01 and the $z = 4.05$ SMG GN20[85] are the red stars and the red square, respectively. The 1σ uncertainties of these measurements are smaller than the size of the symbol. Other objects are compiled from the literature[8,86], and their typical 1σ uncertainties are indicated by the cross at the lower right. HXMM01 is consistent with the sequence defined by local[8] and high-redshift starbursts[12], which have ∼20 times higher SFR surface density at a given gas surface density than quiescently star-forming galaxies. The dashed lines indicate constant gas consumption timescales for the star-forming disks ($\tau_{\rm disk} \equiv \tau_{\rm dyn}/\epsilon_{\rm SF} = \Sigma_{\rm gas}/\Sigma_{\rm SFR}$) of 70 Myr (red) and 1.5 Gyr (grey). Note that $\tau_{\rm disk}$ are smaller than the consumption timescale for the entire gas reservoirs ($\tau_{\rm gas}$), because usually only part of the gas reservoir is actively forming stars.

where we have used the isotropic virial formula to estimate the characteristic circular velocity, $v_c = \sqrt{3/8\ln2}\Delta V_{\rm FWHM} = 0.736\Delta V_{\rm FWHM}$[86], because the velocity gradient is unresolved in our data.

We opt to use the mean of the two very similar timescales as an estimate of the dynamical timescale. For the half-light radius ($R_{1/2}$), we use the CO J = 1→0 HWHM along the major axis from the JVLA data. And for simplicity, we have considered the observed $R_{1/2}$ as unmagnified, although it could be magnified *at most* by a factor of μ, and likely only by $\sqrt{\mu}$. Nevertheless, $\epsilon_{\rm SF}$ depends only linearly on $R_{1/2}$ because all of the other parameters involved in the equation, $\Sigma_{\rm SFR}$, $\Sigma_{\rm gas}$, and $\Delta V_{\rm FWHM}$, are preserved by lensing.

With Eqn. 5, we estimate remarkably high star formation efficiencies: $\epsilon_{\rm SF}$ ∼10% and 40% for X01N and X01S, respectively. This is more than an order-of-magnitude higher than the mean efficiency of quiescently star-forming galaxies like our Milky Way ($\epsilon_{\rm SF}$ ∼2%). It is worth noting that the ∼40% star formation efficiency of X01S is similar to that of the hyperluminous quasar host galaxy[89] SDSS J114816.64+525150.3 at $z = 6.4$, which represents another example of maximum starbursts[13].

The SFR and gas surface densities of HXMM01 are broadly consistent with the starburst sequence defined by local ultra-luminous IR galaxies (ULIRGs) and luminous SMGs. Likewise, its SFR surface density is much higher than the quiescently star-forming galaxies at the same epoch[86] (Figure S9). Along with the $z = 4.05$ SMG GN20[63], HXMM01 seems to also show a small offset from the best-fit relation of other SMGs. This is likely because for most of the SMGs the sizes of star formation and gas reservoir were estimated from maps of higher-J CO lines[12]. Lack of CO J = 1→0 data, it was assumed that high-J CO emission have the same physical extent as the star-forming regions and the CO J = 1→0 emission, although accumulating evidence suggests that total gas reservoirs traced by CO J = 1→0 emission are more extended than the star formation regions traced by submillimeter continuum and the denser gas reservoirs traced by higher-J CO lines in both local ULIRGs[90] and SMGs[11,24,63]. Because the higher-J CO line luminosities were converted to those of CO J = 1→0 using empirically determined mean brightness temperature ratios, the gas surface densities are overestimated for the SMGs that lack CO J = 1→0 maps. This alluded bias is unlikely to be present in studies that have used CO J = 1→0 maps for size measurements[11,24,42,63,91].

4.6 Local Disk Stability

As discussed in § 4.3, the CO J = 1→0 gaseous reservoir of HXMM01 is five times larger than the dust emitting regions embedded within. The reservoir is possibly part of a massive gas inflow bringing copious amounts of cold gas to fuel the starbursts. Similar to extended CO disks in some low redshift galaxies[92–94], it is possible that the gas reservoir of HXMM01 is rotationally supported, suppressing its fragmentation to form giant molecular clouds and stars. To address the

stability of the reservoir, we make use of the Toomre's criterion[95]:

$$Q = \frac{\kappa\sigma}{\pi G\Sigma_{\rm gas}} \tag{10}$$

which compares the gas surface density $\Sigma_{\rm gas}$ to the critical surface density value at which the disk becomes unstable and begins to fragment. In the above, σ is the gas velocity dispersion and κ is the epicylic frequency, which for flat rotation curves can be written as $\kappa = \sqrt{2}v_c/R_{1/2}$, where $v_c = \sqrt{3/8{\rm ln}2}\Delta V_{\rm FWHM} = 1.73\sigma$ is the characteristic circular velocity[86], and $R_{1/2}$ is the CO J = 1→0 disk radius. Combined with $\Sigma_{\rm gas} = 0.5M_{\rm H2}/(\pi R_{1/2}^2)$ and Equation 8, we obtain:

$$Q = 0.97\frac{M_{\rm vir}}{M_{\rm gas}} \tag{11}$$

We estimate $Q > 1$ ($Q = 3-6$) for both X01N and X01S, because their virial masses are a few times greater than their gas masses. Therefore, the disks appear locally stable at least at the half-light radii. This is consistent with our finding that intense star formation occurs only in the inner parts of the disks (§ 4.3), since Q decreases at smaller radii mainly as a result of the higher gas surface densities.

4.7 CO Excitation

The velocity-integrated J=x→y to J=1→0 brightness temperature ratio is defined as,

$$r_{x-y/1-1} = \frac{\int T(x-y)\,{\rm d}v}{\int T(1-0)\,{\rm d}v} = \frac{\int S_\nu(x-y)\,{\rm d}v}{\int S_\nu(1-0)\,{\rm d}v}\left(\frac{\nu_{1-0}}{\nu_{3-2}}\right)^2 , \tag{12}$$

where $S_\nu(x-y)$ is the flux density in the J=x→y transition and ν_{x-y} is the frequency of the transition.

Combining the JVLA CO J = 1→0, CARMA CO J = 3→2, and PdBI CO J = 4→3 data, we measure CO brightness temperature ratios of $r_{3-2/1-0} = 0.64 \pm 0.16$, and $r_{4-3/1-0} = 0.49 \pm 0.09$, similar to other SMGs[11,51,64] and $z > 2$ quasar host galaxies[64,96]. This indicates that we would have obtained a similar amount of total molecular gas mass for HXMM01 if we had converted high-J CO lines to CO J = 1→0 using the mean observed brightness temperature ratios for SMGs.

4.8 Could HXMM01 be AGN Dominated?

It would be useful to know if any of our estimates are biased by the presence of AGNs. Broad Balmer lines seen in HXMM01 are common in SMGs[82,97]. They could indicate either AGN broad line regions or strong outflows. The clear stellar light bump sampled by the IRAC bands (Figure S8a) indicates insignificant AGN contribution. Assuming that AGN activity is responsible for the broad Hα lines, we can scale the composite AGN SED[73,98,99] to the observed Hα flux and

compare it to the observed SED. As shown by the blue dotted curve in Figure S8*b*, the AGN's hot dust bump would start to dominate the SED above $\sim$5 μm and produce a red [5.8]$-$[8.0] color, contradicting the observations. Dust extinction does not compromise this result: the more dust extinction we impose to the AGN SED, the redder [5.8]$-$[8.0] color we get because the AGN SED has to be scaled to the same Hα flux. Therefore, the broad Hα lines in HXMM01 are likely produced by strong starburst-driven outflows and HXMM01 is unlikely dominated by AGNs.

4.9 X-ray Constraints on AGN luminosity and Cooling Rate of Hot Gas

Here we attempt to put some constraints on the AGN luminosity and hot gas cooling rate using the stacked X-ray image from the XMM-LSS survey (§ 1.5), although the constraints are limited because of the shallowness of the X-ray data.

We measure the count rates with a 15″-radius aperture and the background with an annulus between 75″ and 90″ (to avoid an X-ray source 60″ to the NE). After correcting for the aperture loss (70%) and the vignetting ($\sim$50% at 10′ off-axis**), we measure a 3σ upper limit of 1.1×10^{-3} count s^{-1} at 0.5$-$10 keV. With WebPIMMS††, the above count rates yield a flux of 6.6×10^{-15} erg s^{-1} cm^{-2} for a Raymond Smith plasma model of 0.2 $Z_{\odot}$ metallicity and a temperature of 7.1×10^{6} K (i.e., the virial temperature of a 10^{13} M$_{\odot}$ halo at $z = 2.3$), and a 0.5$-$10 keV flux of 9.7×10^{-15} erg s^{-1} cm^{-2} for a power-law model with a photon index of 2. Therefore, the X-ray constraint on the cooling luminosity is $L_{0.5-10\ \mathrm{keV}} < 2.8 \times 10^{44}$ erg s^{-1}, while the constraint on the AGN luminosity is $L_{0.5-10\ \mathrm{keV}} < 4.2 \times 10^{44}$ erg s^{-1}. These luminosities imply a cooling rate of $\dot{M} < 1,900$ M$_{\odot}$ yr^{-1}, and an Eddington-limited black hole mass of $M_{\mathrm{BH,Edd}} \equiv (L_{\mathrm{bol}}/1.45 \times 10^{38}\ \mathrm{erg\ s^{-1}})\ \mathrm{M_{\odot}} < 4.8 \times 10^{7}$ M$_{\odot}$, assuming a bolometric correction of 6%[100]. Note that this is not an upper limit on the black hole mass, because sub-Eddington accretion would allow higher mass black holes.

It is not surprising that we did not detect HXMM01 in the relatively shallow X-ray observations. In the 2 Ms *Chandra* Deep Field North[101], only two of the 20 radio-detected SMGs show X-ray fluxes greater than $f_{0.5-8\ \mathrm{keV}} > 5 \times 10^{-15}$ erg s^{-1} cm^{-2}.

4.10 Space Density of Hyper-Luminous Mergers Like HXMM01?

HXMM01 has an intrinsic flux density of $S_{880} \simeq 17 \pm 4$ mJy. From the 880 μm single-dish source count function[102], the surface density of sources brighter than 17 mJy at 880 μm is $\sim$0.1 deg^{-2}. The estimated volume density of such bright objects[102] is 10^{-8} to 10^{-7} Mpc^{-3}, much lower than volume density of SMGs with 880 μm flux densities of a few mJy ($\sim 10^{-6}$ to 10^{-5} Mpc^{-3}). Due to the lack of bright SMGs in LABOCA surveys, the volume and surface density estimates are uncertain by at least a factor of a few in either direction. Follow-up CO observations

** *XMM-Newton* Users' Handbook, §3.2.2.2

†† http://heasarc.gsfc.nasa.gov/Tools/w3pimms.html

of 2 to 10 mJy SMGs find that $16 \pm 13\%$ (2 out of 12) are resolved[17] into physically related galaxy pairs with separations greater than 12 kpc. Assuming that the resolved merger fraction of fainter galaxies is applicable to more luminous sources, we estimate the surface density of bright SMG mergers like HXMM01 is $\sim$0.01 deg^{-2}.

We can make another estimate using the source count at 500 μm. The surface density of *Herschel*-selected sources[5] with $S_{500} > 100$ mJy that are not associated with either a local ($z < 0.1$) star-forming galaxy or a bright radio-loud AGN is 0.14 ± 0.04 deg^{-2}. Follow-up high resolution imaging of ten such sources[5] reveals that HXMM01 is the only well-separated SMG merger with $S_{500} > 100$ mJy in HerMES. So the expected surface density of well-resolved SMG pairs is ~ 0.014 deg^{-2}, or $\sim$1 per 100 deg^2, consistent with our estimate from the 880 μm source count function.

Since various *Herschel* surveys (HerMES, H-ATLAS, HeLMS) have mapped $\sim$1000 deg^2 and the South Pole Telescope[103,104] has mapped $\sim$2500 deg^2 in submillimeter and millimeter wavelengths, there should be $\sim$35 bright SMG mergers in these data. To identify such mergers among $\sim$480 bright lensing candidates, a good strategy is to use the CO luminosity vs. line width relation to weed out highly magnified sources (see Figure S7).

In Table 1 we list the resolved SMG mergers reported in the literature. We note that there are many SMGs that have properties indicative of late-stage mergers. But here we only include the objects that have been resolved into multiple components. These mergers have projected separations spanning from 4 to 30 kpc and are at redshifts between $1 < z < 4$. References for the sources in Table 1 are as follows: SMM J02399−0136[105,106]; SMM J09431+4700[17,24,107]; SMM J105141[17]; SMM J123707[17,82,88,107]; SMM J123711[18,108,109]; SMM 163650[88]; 4C 60.07[110,111].

Acknowledgement

HF, AC and JLW are supported by NASA funds for US participants in *Herschel* through an award from JPL. LW and AS acknowledge support from the Science and Technology Facilities Council [grant number ST/I000976/1]. AH acknowledges support from the NSF grants AST-0503946 to the University of Maryland and AST-0708653 to Rutgers University. MN acknowledges financial support from ASI/INAF agreement I/072/09/0.

Some of the data presented herein were obtained at the W.M. Keck Observatory, which is operated as a scientific partnership among the California Institute of Technology, the University of California and the National Aeronautics and Space Administration. The Observatory was made possible by the generous financial support of the W.M. Keck Foundation. The authors wish to recognize and acknowledge the very significant cultural role and reverence that the summit of Mauna Kea has always had within the indigenous Hawaiian community. We are most fortunate to have the opportunity to conduct observations from this mountain.

Support for CARMA construction was derived from the states of California, Illinois, and Maryland, the James S. McDonnell Foundation, the Gordon and Betty Moore Foundation, the Kenneth T. and Eileen L. Norris Foundation, the University of Chicago, the Associates of the California Institute of Technology, and the National Science Foundation. Ongoing CARMA development and operations are supported by the National Science Foundation under a cooperative agreement, and by the CARMA partner universities.

Based on observations obtained with MegaPrime/MegaCam, a joint project of CFHT and CEA/DAPNIA, at the Canada-France-Hawaii Telescope (CFHT) which is operated by the National Research Council (NRC) of Canada, the Institut National des Science de l'Univers of the Centre National de la Recherche Scientifique (CNRS) of France, and the University of Hawaii. This work is based in part on data products produced at TERAPIX and the Canadian Astronomy Data Centre as part of the Canada-France-Hawaii Telescope Legacy Survey, a collaborative project of NRC and CNRS.

Based on observation obtained as part of the VISTA Hemisphere Survey, ESO Progam, 179.A-2010 (PI: McMahon).

The William *Herschel* Telescope is operated on the island of La Palma by the Isaac Newton Group in the Spanish Observatorio del Roque de los Muchachos of the Instituto de Astrofsica de Canarias. The WHT observations are part of the International Time Programme 2010-2011 (PI Perez-Fournon).

Based on observations carried out with the IRAM Plateau de Bure Interferometer and 30m Telescope. IRAM is supported by INSU/CNRS (France), MPG (Germany) and IGN (Spain).

Quantity	HXMM01	X01N	X01S	Unit
		Lens Modeling		
$\mu(K_S)$	$1.59^{+0.36}_{-0.22}$	$1.78^{+0.55}_{-0.32}$	$1.35^{+0.19}_{-0.12}$	...
$\mu(880\mu m)$	$1.57^{+0.34}_{-0.21}$	$1.83^{+0.60}_{-0.35}$	$1.34^{+0.18}_{-0.12}$	...
μ(CO1-0)	$1.54^{+0.31}_{-0.20}$	$1.84^{+0.63}_{-0.36}$	$1.37^{+0.20}_{-0.13}$	...
Separation	$2.22^{+0.32}_{-0.45}/18.7^{+2.7}_{-3.8}$	...	...	arcsec/kpc
		MAGPHYS SED Modeling		
$M_{\rm stellar}$	...	$(1.6 \pm 0.1) \times 10^{11}$	$(1.7 \pm 0.3) \times 10^{11}$	μ^{-1} M$_\odot$
$L_{\rm IR}$	...	$(1.1 \pm 0.1) \times 10^{13}$	$(1.1 \pm 0.1) \times 10^{13}$	μ^{-1} L$_\odot$
		Optically Thick Blackbody Fit		
$T_{\rm dust}$	55 ± 3	...	...	K
β	2.17 ± 0.23	...	...	...
λ_0	180 ± 30	...	...	μm
$M_{\rm dust}$	$(2.9 \pm 0.7) \times 10^9$	...	...	μ^{-1} M$_\odot$
$L_{\rm IR}$	$(3.2 \pm 0.2) \times 10^{13}$	...	...	μ^{-1} L$_\odot$
$A_{\rm SB}$	6.3 ± 1.5	...	...	μ^{-1} kpc^2
$\alpha_{\rm CO,gas/dust}$	0.91 ± 0.27	...	...	M$_\odot$/(K km s^{-1} pc^2)
		SMA 880 μm Image		
A_{880}	...	9.8 ± 0.6	8.9 ± 0.7	μ^{-1} kpc^2
$\Sigma_{\rm SFR}$	...	56 ± 6	62 ± 8	M$_\odot$ yr^{-1} kpc^{-2}
		JVLA CO(1→0) Datacube		
$z_{\rm CO1-0}$	2.3079 ± 0.0007	2.3103 ± 0.0007	2.3074 ± 0.0004	...
$\Delta V_{\rm FWHM}$	840 ± 160	970 ± 150	660 ± 100	km s^{-1}
$L'_{\rm CO1-0}$	$(4.5 \pm 0.8) \times 10^{11}$	$(1.3 \pm 0.2) \times 10^{11}$	$(1.7 \pm 0.2) \times 10^{11}$	μ^{-1} K km s^{-1} pc^2
$A_{\rm CO}$	...	27 ± 2	68 ± 2	μ^{-1} kpc^2
$W_{\rm CO}$	...	4900 ± 780	2500 ± 350	K km s^{-1}
$\alpha_{\rm CO,N12}$	[0.8]	0.71 ± 0.04	0.88 ± 0.04	M$_\odot$/(K km s^{-1} pc^2)
$M_{\rm gas}$	$(3.6 \pm 0.6) \times 10^{11}$	$(9.5 \pm 1.0) \times 10^{10}$	$(1.5 \pm 0.1) \times 10^{11}$	μ^{-1} M$_\odot$
$\Sigma_{\rm gas}$	...	1730 ± 200	1090 ± 100	M$_\odot$ pc^{-2}
$R_{1/2}$	...	3.0 ± 0.1	4.6 ± 0.1	kpc
$M_{\rm vir}$	...	$(5.9 \pm 1.9) \times 10^{11}$	$(4.3 \pm 1.3) \times 10^{11}$	M$_\odot$
$\tau_{\rm rot}$	...	4.0 ± 0.8	9.3 ± 1.5	Myr
$\tau_{\rm ff}$	...	2.2 ± 0.4	5.1 ± 0.8	Myr
$\epsilon_{\rm SF}$	...	0.10 ± 0.03	0.41 ± 0.10	...
$M_{\rm gas}/M_{\rm baryon}$	0.52 ± 0.05	0.37 ± 0.04	0.47 ± 0.06	
		GBT CO(1→0) Spectrum		
$z_{\rm CO1-0}$	2.3074 ± 0.0008	...	...	...
$\Delta V_{\rm FWHM}$	1670 ± 140	...	...	km s^{-1}
$L'_{\rm CO1-0}$	$(6.2 \pm 0.9) \times 10^{11}$	...	...	μ^{-1} K km s^{-1} pc^2
		CARMA CO(3→2) Spectrum		
$z_{\rm CO3-2}$	2.3073 ± 0.0010	...	...	...

Continued on next page

Quantity	HXMM01	X01N	X01S	Unit
$\Delta V_{\rm FWHM}$	980 ± 200	...	...	km s^{-1}
$L'_{\rm CO3-2}$	$(2.9 \pm 0.5) \times 10^{11}$	...	...	μ^{-1} K km s^{-1} pc^2
$r_{3-2/1-0}$	0.64 ± 0.16	...	...	...
PdBI CO(4→3) Spectrum				
$z_{\rm CO4-3}$	2.3081 ± 0.0002	...	...	...
$\Delta V_{\rm FWHM}$	880 ± 50	...	...	km s^{-1}
$L'_{\rm CO4-3}$	$(2.2 \pm 0.1) \times 10^{11}$	...	...	μ^{-1} K km s^{-1} pc^2
$r_{4-3/1-0}$	0.49 ± 0.09	...	...	...
Keck Hα Spectra				
$z_{\rm H\alpha}$	...	2.3144 ± 0.0019	2.3107 ± 0.0015	...
$\Delta V_{\rm FWHM}$	...	$2,370 \pm 340$	$1,750 \pm 320$	km s^{-1}
$L_{\rm H\alpha}$	...	$(9.4 \pm 1.6) \times 10^{42}$	$(1.4 \pm 0.2) \times 10^{43}$	μ^{-1} erg s^{-1}

Table S1 – Observed and Derived Properties of HXMM01. Refer to text for definitions of the parameters.

Instrument	Band	λ (μm)	G1 (μJy)	G2 (μJy)
GALEX	NUV	0.23	< 0.5	< 0.5
CFHT	u*	0.38	0.7 ± 0.2	1.0 ± 0.2
CFHT	g'	0.49	2.9 ± 0.2	4.1 ± 0.2
CFHT	r'	0.62	6.3 ± 0.3	14.7 ± 0.5
CFHT	i'	0.76	14.8 ± 0.5	31.1 ± 1.0
CFHT	z'	0.88	19.8 ± 0.9	46.3 ± 1.6
HST	F110W	1.16	29.0 ± 0.8	72.9 ± 1.7
WHT	J	1.25	32.7 ± 1.2	78.2 ± 1.8
VISTA	H	1.65	60.2 ± 5.3	138.5 ± 6.5
WHT	K	2.15	87.7 ± 7.7	184.6 ± 5.1
IRAC	3.6 μm	3.56	111.1 ± 7.9	180.6 ± 7.9
IRAC	4.5 μm	4.51	77.6 ± 6.2	146.5 ± 7.3
IRAC	5.8 μm	5.76	70.4 ± 13.9	99.3 ± 13.7
IRAC	8.0 μm	7.96	51.4 ± 14.6	106.8 ± 14.7

Table S2 – Photometry of the foreground lensing galaxies.

Instrument	Band	λ (μm)	HXMM01 (μJy)	X01N (μJy)	X01S (μJy)
GALEX	NUV	0.23	< 0.5	< 0.5	< 0.5
CFHT	u*	0.38	0.5 ± 0.1	< 0.2	0.5 ± 0.1
CFHT	g'	0.49	1.3 ± 0.1	< 0.2	1.2 ± 0.1
CFHT	r'	0.62	2.3 ± 0.2	0.2 ± 0.1	2.1 ± 0.2
CFHT	i'	0.76	2.5 ± 0.3	< 0.5	2.3 ± 0.3
CFHT	z'	0.88	3.8 ± 0.6	0.6 ± 0.3	3.2 ± 0.6
HST	F110W	1.16	6.2 ± 0.4	1.6 ± 0.2	4.7 ± 0.3
WHT	J	1.25	6.8 ± 0.8	1.7 ± 0.5	5.1 ± 0.6
VISTA	H	1.65	...	...	...
WHT	K	2.15	22.5 ± 3.4	7.6 ± 2.0	14.9 ± 2.7
IRAC	3.6 μm	3.56	48.5 ± 5.8	22.2 ± 5.0	26.4 ± 2.8
IRAC	4.5 μm	4.51	64.1 ± 5.9	29.4 ± 5.0	34.7 ± 2.9
IRAC	5.8 μm	5.76	111.9 ± 12.0	51.5 ± 6.8	60.4 ± 9.6
IRAC	8.0 μm	7.96	86.9 ± 12.7	41.1 ± 6.7	45.8 ± 10.6
MIPS	24 μm	23.67	2280 ± 100	...	...
PACS	70 μm	72.35	7970 ± 3890	...	...
PACS	100 μm	107.42	31700 ± 3000	...	...
PACS	160 μm	166.15	102100 ± 6000	...	...
SPIRE	250 μm	250.94	180300 ± 14300	...	...
SPIRE	350 μm	354.27	192100 ± 15500	...	...
SPIRE	500 μm	509.47	131600 ± 11300	...	...
SMA	880 μm	876.50	27000 ± 3000	9300 ± 1200	8300 ± 1100
MAMBO	1.2 mm	1200.00	11200 ± 1900	...	...
PdBI	2.1 mm	2141.31	1220 ± 240	...	...
VLA	1.4 GHz	2.1×10^5	< 420	< 420	< 420

Table S3 – Photometry of HXMM01.

References

31. Chabrier, G. Galactic Stellar and Substellar Initial Mass Function. *PASP* **115**, 763–795 (2003).

32. Komatsu, E. *et al*. Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation. *Astrophys. J. Supp.* **192**, 18–64 (2011).

33. Griffin, M. J. *et al*. The Herschel-SPIRE instrument and its in-flight performance. *Astron. Astrophys.* **518**, L3 (2010).

34. Pilbratt, G. L. *et al*. Herschel Space Observatory. An ESA facility for far-infrared and submillimetre astronomy. *Astron. Astrophys.* **518**, L1 (2010).

35. Negrello, M. *et al*. The Detection of a Population of Submillimeter-Bright, Strongly Lensed Galaxies. *Science* **330**, 800– (2010).

36. Ho, P. T. P., Moran, J. M. & Lo, K. Y. The Submillimeter Array. *Astrophys. J.* **616**, L1–L6 (2004).

37. Perley, R. A., Chandler, C. J., Butler, B. J. & Wrobel, J. M. The Expanded Very Large Array: A New Telescope for New Science. *Astrophys. J.* **739**, L1 (2011).

38. Poglitsch, A. *et al*. The Photodetector Array Camera and Spectrometer (PACS) on the Herschel Space Observatory. *Astron. Astrophys.* **518**, L2 (2010).

39. Levenson, L. *et al*. HerMES: SPIRE Science Demonstration Phase maps. *Mon. Not. R. Astron. Soc.* **409**, 83–91 (2010).

40. Viero, M. P. *et al*. HerMES: Cosmic Infrared Background Anisotropies and the Clustering of Dusty Star-Forming Galaxies. *arXiv:1208.5049* **1208**, 5049 (2012).

41. Wizinowich, P. L. *et al*. The W. M. Keck Observatory Laser Guide Star Adaptive Optics System: Overview. *PASP* **118**, 297–309 (2006).

42. Fu, H. *et al*. A Comprehensive View of a Strongly Lensed Planck-Associated Submillimeter Galaxy. *Astrophys. J.* **753**, 134 (2012).

43. Kreysa, E. *et al*. Bolometer array development at the Max-Planck-Institut für Radioastronomie. *Infrared Physics and Technology* **40**, 191–197 (1999).

44. Pierre, M. *et al*. The XMM-LSS survey. Survey design and first results. *Journal of Cosmology and Astro-Particle Physics* **09**, 011 (2004).

45. Morrissey, P. *et al*. The Calibration and Data Products of GALEX. *Astrophys. J. Supp.* **173**, 682–697 (2007).

46. Gwyn, S. D. J. The Canada-France-Hawaii Telescope Legacy Survey: Stacked Images and Catalogs. *Astron. J.* **143**, 38 (2012).

47. Lonsdale, C. J. *et al*. SWIRE: The SIRTF Wide-Area Infrared Extragalactic Survey. *PASP* **115**, 897–927 (2003).

48. Wright, E. L. *et al*. The Wide-field Infrared Survey Explorer (WISE): Mission Description and Initial On-orbit Performance. *Astron. J.* **140**, 1868–1881 (2010).

49. Becker, R. H., White, R. L. & Helfand, D. J. The FIRST Survey: Faint Images of the Radio Sky at Twenty Centimeters. *Astrophys. J.* **450**, 559 (1995).

50. Riechers, D. A. Molecular Gas in Lensed $z > 2$ Quasar Host Galaxies and the Star Formation Law for Galaxies with Luminous Active Galactic Nuclei. *Astrophys. J.* **730**, 108 (2011).

51. Harris, A. I. *et al*. CO J = 1-0 Spectroscopy of Four Submillimeter Galaxies with the Zpectrometer on the Green Bank Telescope. *Astrophys. J.* **723**, 1139–1149 (2010).

52. McLean, I. S. *et al*. Design and development of NIRSPEC: a near-infrared echelle spectrograph for the Keck II telescope. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, vol. 3354, 566–578 (1998).

53. Oke, J. B. *et al*. The Keck Low-Resolution Imaging Spectrometer. *PASP* **107**, 375 (1995).

54. Peng, C. Y., Ho, L. C., Impey, C. D. & Rix, H.-W. Detailed Decomposition of Galaxy Images. II. Beyond Axisymmetric Models. *Astron. J.* **139**, 2097–2129 (2010).

55. Cohen, M., Wheaton, W. A. & Megeath, S. T. Spectral Irradiance Calibration in the Infrared. XIV. The Absolute Calibration of 2MASS. *Astron. J.* **126**, 1090–1096 (2003).

56. Reach, W. T. *et al*. Absolute Calibration of the Infrared Array Camera on the Spitzer Space Telescope. *PASP* **117**, 978–990 (2005).

57. Engelbracht, C. W. *et al*. Absolute Calibration and Characterization of the Multiband Imaging Photometer for Spitzer. I. The Stellar Calibrator Sample and the 24 mum Calibration. *PASP* **119**, 994–1018 (2007).

58. Gordon, K. D. *et al*. Absolute Calibration and Characterization of the Multiband Imaging Photometer for Spitzer. II. 70 mum Imaging. *PASP* **119**, 1019–1037 (2007).

59. Swinyard, B. M. *et al*. In-flight calibration of the Herschel-SPIRE instrument. *Astron. Astrophys.* **518**, L4 (2010).

60. Nguyen, H. T. *et al*. HerMES: The SPIRE confusion limit. *Astron. Astrophys.* **518**, L5 (2010).

61. Bruzual, G. & Charlot, S. Stellar population synthesis at the resolution of 2003. *Mon. Not. R. Astron. Soc.* **344**, 1000–1028 (2003).

62. Charlot, S. & Fall, S. M. A Simple Model for the Absorption of Starlight by Dust in Galaxies. *Astrophys. J.* **539**, 718–731 (2000).

63. Carilli, C. L. *et al.* Imaging the Molecular Gas in a Submillimeter Galaxy at z = 4.05: Cold Mode Accretion or a Major Merger? *Astrophys. J.* **714**, 1407–1417 (2010).

64. Carilli, C. & Walter, F. Cool Gas in High Redshift Galaxies. *arXiv:1301.0371* (2013).

65. Harris, A. I. *et al.* Blind Detections of CO J = 1–0 in 11 H-ATLAS Galaxies at z = 2.1–3.5 with the GBT/Zpectrometer. *Astrophys. J.* **752**, 152 (2012).

66. Koopmans, L. V. E., Treu, T., Bolton, A. S., Burles, S. & Moustakas, L. A. The Sloan Lens ACS Survey. III. The Structure and Formation of Early-Type Galaxies and Their Evolution since z ˜ 1. *Astrophys. J.* **649**, 599–615 (2006).

67. Sluse, D., Chantry, V., Magain, P., Courbin, F. & Meylan, G. COSMOGRAIL: the COSmological MOnitoring of GRAvItational Lenses. X. Modeling based on high-precision astrometry of a sample of 25 lensed quasars: consequences for ellipticity, shear, and astrometric anomalies. *Astron. Astrophys.* **538**, 99 (2012).

68. Jullo, E. *et al.* A Bayesian approach to strong lensing modelling of galaxy clusters. *New Journal of Physics* **9**, 447 (2007).

69. Navarro, J. F., Frenk, C. S. & White, S. D. M. The Structure of Cold Dark Matter Halos. *Astrophys. J.* **462**, 563 (1996).

70. Chary, R. & Elbaz, D. Interpreting the Cosmic Infrared Background: Constraints on the Evolution of the Dust-enshrouded Star Formation Rate. *Astrophys. J.* **556**, 562–581 (2001).

71. Silva, L., Granato, G. L., Bressan, A. & Danese, L. Modeling the Effects of Dust on Galactic Spectral Energy Distributions from the Ultraviolet to the Millimeter Band. *Astrophys. J.* **509**, 103–117 (1998).

72. Swinbank, A. M. *et al.* Intense star formation within resolved compact regions in a galaxy at z = 2.3. *Nature* **464**, 733–736 (2010).

73. Hopkins, P. F., Richards, G. T. & Hernquist, L. An Observational Determination of the Bolometric Quasar Luminosity Function. *Astrophys. J.* **654**, 731–753 (2007).

74. Calzetti, D. *et al.* The Dust Content and Opacity of Actively Star-forming Galaxies. *Astrophys. J.* **533**, 682–695 (2000).

75. Kennicutt, J., Robert C. Star Formation in Galaxies Along the Hubble Sequence. *ARA&A* **36**, 189–232 (1998).

76. Dunne, L. *et al.* The SCUBA Local Universe Galaxy Survey - I. First measurements of the submillimetre luminosity and dust mass functions. *Mon. Not. R. Astron. Soc.* **315**, 115–139 (2000).

77. James, A., Dunne, L., Eales, S. & Edmunds, M. G. SCUBA observations of galaxies with metallicity measurements: a new method for determining the relation between submillimetre luminosity and dust mass. *Mon. Not. R. Astron. Soc.* **335**, 753–761 (2002).

78. Foreman-Mackey, D., Hogg, D. W., Lang, D. & Goodman, J. emcee: The MCMC Hammer. *arXiv* **1202**, 3665 (2012).

79. Hwang, H. S. *et al.* Evolution of dust temperature of galaxies through cosmic time as seen by Herschel. *Mon. Not. R. Astron. Soc.* **409**, 75–82 (2010).

80. Hayward, C. C. *et al.* How to distinguish starbursts and quiescently star-forming galaxies: the 'bimodal' submillimetre galaxy population as a case study. *Mon. Not. R. Astron. Soc.* **424**, 951–970 (2012).

81. Draine, B. T. *et al.* Dust Masses, PAH Abundances, and Starlight Intensities in the SINGS Galaxy Sample. *Astrophys. J.* **663**, 866–894 (2007).

82. Swinbank, A. M. *et al.* The Rest-Frame Optical Spectra of SCUBA Galaxies. *Astrophys. J.* **617**, 64–80 (2004).

83. Hodge, J. A. *et al.* Evidence for a clumpy, rotating gas disk in a submillimeter galaxy at z=4. *arXiv:1209.2418* (2012).

84. Magnelli, B. *et al.* A Herschel view of the far-infrared properties of submillimetre galaxies. *Astron. Astrophys.* **539**, 155 (2012).

85. Daddi, E. *et al.* Two Bright Submillimeter Galaxies in a z = 4.05 Protocluster in Goods-North, and Accurate Radio-Infrared Photometric Redshifts. *Astrophys. J.* **694**, 1517–1538 (2009).

86. Tacconi, L. J. *et al.* PHIBSS: molecular gas content and scaling relations in z$\sim$1-3 normal star forming galaxies. *arXiv:1211.5743* (2012).

87. Spitzer, L. *Dynamical evolution of globular clusters* (Princeton, NJ, Princeton University Press, 1987, 191 p., 1987).

88. Tacconi, L. J. *et al.* Submillimeter Galaxies at $z \sim 2$: Evidence for Major Mergers and Constraints on Lifetimes, IMF, and CO-H2 Conversion Factor. *Astrophys. J.* **680**, 246–262 (2008).

89. Walter, F. *et al.* A kiloparsec-scale hyper-starburst in a quasar host less than 1gigayear after the Big Bang. *Nature* **457**, 699–701 (2009).

90. Telesco, C. M., Dressel, L. L. & Wolstencroft, R. D. The genesis of starbursts and infrared emission in the centers of galaxies. *Astrophys. J.* **414**, 120–143 (1993).

91. Swinbank, A. M. *et al*. The Interstellar Medium in Distant Star-forming Galaxies: Turbulent Pressure, Fragmentation, and Cloud Scaling Relations in a Dense Gas Disk at z = 2.3. *Astrophys. J.* **742**, 11 (2011).

92. Lucero, D. M. & Young, L. M. Radio Continuum and Star Formation in CO-rich Early-Type Galaxies. *Astron. J.* **134**, 2148–2159 (2007).

93. Koda, J. *et al*. Starbursting nuclear CO disks of early-type spiral galaxies. *Astron. Astrophys.* **431**, 887–891 (2005).

94. Boissier, S., Prantzos, N., Boselli, A. & Gavazzi, G. The star formation rate in disc galaxies: thresholds and dependence on gas amount. *Mon. Not. R. Astron. Soc.* **346**, 1215–1230 (2003).

95. Toomre, A. On the gravitational stability of a disk of stars. *Astrophys. J.* **139**, 1217–1238 (1964).

96. Riechers, D. A. *et al*. CO(J = 1-0) in $z > 2$ Quasar Host Galaxies: No Evidence for Extended Molecular Gas Reservoirs. *Astrophys. J.* **739**, L32 (2011).

97. Harrison, C. M. *et al*. Energetic galaxy-wide outflows in high-redshift ultra-luminous infrared galaxies hosting AGN activity. *arXiv:1205.1801* (2012).

98. Vanden Berk, D. E. *et al*. Composite Quasar Spectra from the Sloan Digital Sky Survey. *Astron. J.* **122**, 549–564 (2001).

99. Richards, G. T. *et al*. Spectral Energy Distributions and Multiwavelength Selection of Type 1 Quasars. *Astrophys. J. Supp.* **166**, 470–497 (2006).

100. Elvis, M. *et al*. Atlas of quasar energy distributions. *Astrophys. J. Supp.* **95**, 1–68 (1994).

101. Alexander, D. M. *et al*. The X-Ray Spectral Properties of SCUBA Galaxies. *Astrophys. J.* **632**, 736–750 (2005).

102. Karim, A. *et al*. An ALMA survey of submillimetre galaxies in the Extended Chandra Deep Field South: High resolution 870um source counts. *arXiv:1210.0249* (2012).

103. Vieira, J. D. *et al*. Extragalactic Millimeter-wave Sources in South Pole Telescope Survey Data: Source Counts, Catalog, and Statistics for an 87 Square-degree Field. *Astrophys. J.* **719**, 763–783 (2010).

104. Reichardt, C. L. *et al*. Galaxy Clusters Discovered via the Sunyaev-Zel'dovich Effect in the First 720 Square Degrees of the South Pole Telescope Survey. *Astrophys. J.* **763**, 127 (2013).

105. Valiante, E. *et al*. A Mid-Infrared Spectroscopic Study of Submillimeter Galaxies: Luminous Starbursts at High Redshift. *Astrophys. J.* **660**, 1060–1071 (2007).

106. Ivison, R. J. *et al*. Gas, dust and stars in the SCUBA galaxy, SMMJ02399-0136: the EVLA reveals a colossal galactic nursery. *Mon. Not. R. Astron. Soc.* **404**, 198–205 (2010).

107. Tacconi, L. J. *et al*. High-Resolution Millimeter Imaging of Submillimeter Galaxies. *Astrophys. J.* **640**, 228–240 (2006).

108. Bothwell, M. S. *et al*. High-resolution CO and radio imaging of ULIRGs: extended CO structures and implications for the universal star formation law. *Mon. Not. R. Astron. Soc.* **405**, 219–233 (2010).

109. Riechers, D. A. *et al*. Imaging the Molecular Gas Properties of a Major Merger Driving the Evolution of a z = 2.5 Submillimeter Galaxy. *Astrophys. J.* **733**, L11 (2011).

110. Papadopoulos, P. P. *et al*. CO (4-3) and Dust Emission in Two Powerful High-Z Radio Galaxies, and CO Lines at High Redshifts. *Astrophys. J.* **528**, 626–636 (2000).

111. Ivison, R. J. *et al*. Interferometric imaging of the high-redshift radio galaxy, 4C60.07: an SMA, Spitzer and VLA study reveals a binary AGN/starburst. *Mon. Not. R. Astron. Soc.* **390**, 1117–1126 (2008).