TESI DI DOTTORATO

UNIVERSITÀ DEGLI STUDI DI NAPOLI "FEDERICO II"

DIPARTIMENTO DI INGEGNERIA BIOMEDICA,
ELETTRONICA E DELLE TELECOMUNICAZIONI

DOTTORATO DI RICERCA IN
INGEGNERIA ELETTRONICA E DELLE TELECOMUNICAZIONI

THE ROLE OF THE WIDELY LINEAR PROCESSING IN THE NARROWBAND AND WIDEBAND SYSTEMS

ANGELA SARA CACCIAPUOTI

Il Coordinatore del Corso di Dottorato	Il Tutore
Ch.mo Prof. Giovanni POGGI	Ch.mo Prof. Luigi PAURA

A. A. 2007–2008

“*To my parents Ferdinando and Imma,*
to my brother Gino and to Marcello.”

Acknowledgments

I would like to thank my advisor Professor Luigi Paura who inspired me to enter the Ph.D. course and provided invaluable assistance always. I’d also like to express him my natural affection for his patience, goodness and trust.

I would also like to thank Professor Giacinto Gelli for his help in my time as Ph.D. student.

I gratefully acknowledge all the people of the Telecommunication Group for profound impact on both my professional and personal level.

A special thanks to Sara Parrilli, who offered me her intimate friendship.

To my always friends Gabriella and Valeria for all.

Thanks to my parents and to my brother, for their unwavering support and unreserved love.

Finally, to Marcello for his love and understanding.

Contents

List of Figures

Notations

$\mathbb{C}$	denotes the field of complex numbers
$\mathbb{R}$	denotes the field of real numbers
$\mathbb{Z}$	denotes the field of integer numbers
$\mathbb{C}^{m\times n}$	denotes the field of $m \times n$ complex matrices
$\mathbb{R}^{m\times n}$	denotes the field of of $m \times n$ real matrices
a	scalar
a^*	complex coniugate of a scalar a
$\|a\|$	absolute value of the number a
$\mathrm{Re}[a]$	real part of a complex scalar a
$\mathrm{Im}[a]$	imaginary part of a complex scalar a
$\angle a$	argument of a complex number in $[0, 2\pi]$
$\boldsymbol{a}$	vector
$\boldsymbol{a}^*$	complex conjugate of vector $\boldsymbol{a}$
$\boldsymbol{a}^T$	transpose of vector $\boldsymbol{a}$

$\boldsymbol{a}^H$	Hermitian (conjugate transpose) of vector $\boldsymbol{a}$
$\|\mathbf{a}\| = (\mathbf{a}^H\mathbf{a})^{1/2}$	Euclidean norm
$\boldsymbol{A}$	matrix
$\boldsymbol{A}^T$	transpose of matrix $\boldsymbol{A}$
$\boldsymbol{A}^H$	Hermitian (conjugate transpose) of matrix $\boldsymbol{A}$
$\boldsymbol{A}^{-1}$	inverse of matrix $\boldsymbol{A}$
$\boldsymbol{A}^\dagger$	Moore-Penrose generalized inverse of matrix $\boldsymbol{A}$
$\mathbf{I}_m \in \mathbb{R}^{m\times m}$	the $m \times m$ identity matrix
$\mathbf{0}_m$	the null vector
$\mathbf{O}_{n\times m}$	the $n \times m$ null matrix
$E[\cdot]$	expectation operator
$Var[\cdot]$	variance
$\mathbf{A} = \text{diag}(A_{11}, \ldots, A_{nn})$	diagonal matrix with elements A_{ii}
trace$(\cdot)$	trace operator
rank$(\cdot)$	rank
$\mathcal{N}(\mathbf{A})$	space null of matrix $\boldsymbol{A} \in \mathbb{C}^{m\times n}$
$\mathcal{R}(\mathbf{A})$	range (column space) of matrix $\boldsymbol{A} \in \mathbb{C}^{m\times n}$
$\mathcal{R}^\perp(\mathbf{A})$	orthogonal complement of the column space of matrix $\boldsymbol{A} \in$
$i \triangleq \sqrt{-1}$	the imaginary unit

Introduction

Wireless Communication is one of the most vibrant areas in the communication field today. This is due to a confluence of several factors. First, there has been an explosive increase in demand for wireless multimedia services. Second the dramatic progress in micro-processor technology has enabled small-area and low-power implementation of sophisticate signal processing algorithms and coding techniques.

The design of a wireless communication system differs notably from wired system design due to the nature of the wireless channel, which is an unpredictable and difficult communication medium. A frequently occurring problem with electromagnetic wave propagation in such a medium is the signal multipath. Signal multipath occurs when the transmitted signal arrives at the receiver via multiple propagation paths with different delays and it commonly results in intersymbol interference (ISI). Moreover, a signal that propagates through a wireless channel may experience random fluctuations in time if the transmitter, the receiver, or surrounding objects are moving, due to changing reflections and attenuation. Thus, the characteristics of the channel appear to change randomly with time, adding further complexity to the design of reliable communication systems.

In addition, the increasing demand for wireless communications is making the radio spectrum a scarce resource that has to be managed as efficiently as possible since it must be allocated to many different applications and systems.

In this thesis, we focus on the physical layer strategies that allow the effectiveness sharing of the available resources. In the first part of this thesis, we consider narrowband systems, namely systems in which the performances are significantly limited by intersymbol interference (ISI) because of the time dispersive nature of the wireless channel. The second part focuses on wideband systems that use DS-CDMA or MC-CDMA techniques, for which the principal cause of performance degradation is the interference due to multiple access (MAI). The common elements among these topics is the signal process-

ing technique used to recovery the information data: the *Widely-Linear* (WL) technique.

More specifically, in the first part, starting from the basic concepts related to the multiple access techniques, we present the general framework of the equalization problem in a digital communication system and we explain why it is mandatory. In this context, we also introduce the WL receiver technique, since when the transmitted symbol sequence is *improper or noncircular* [1] (as it happens in many modulation schemes of practical interest) it is well-known [2, 3] that a better estimate of the transmitted symbols can be obtained by resorting to widely-linear (WL) estimators, which augment the degrees-of-freedom at the equalizer designer's disposal.

Successively we consider one of the earliest blind channel equalization technique: the Godard or constant modulus (CM) algorithm [4, 5, 6]. This technique is blind since it does not utilize training sequences to cancel or reduce the intersymbol interference (ISI) introduced by frequency-selective transmission channels, avoiding so to waste the available bandwidth resources. The CM cost function exhibits a multimodal surface whose characteristics allow one to gain important insights about the expected behaviors of the equalizer. Therefore, detailed studies of the stationary points of the CM cost function are conducted in the letterature. Such studies have shown that, in the absence of noise and when the symbol sequence is a *proper* [7] random process, all the local minima of the L-CM cost function are global ones, allowing one to exactly suppress the ISI. Nevertheless, in this thesis we show that if the symbol sequence is an improper random process, as it happens in many cases of practical interest, the L-CM equalizer is not able to fully suppress the ISI. The reason for such a behavior is the presence of undesired local minima of the L-CM cost function, which have been analytically determined.

Moreover, we have already underline that when the transmitted symbol sequence is improper, widely-linear (WL) equalizers outperform linear ones. As a consequence, we provide a general and unified framework [8, 9] to design WL equalizers for both real- and complex-valued improper modulations, by deriving the conditions assuring perfect symbol recovery in the absence of noise and providing some insights into the synthesis and analysis of blind WL-CM equalizers. Specifically, since also the WL-CM cost function exhibits undesired global minima we propose to resort to a constrained WL-CM equalizer assuring perfect symbol recovery in the absence of noise.

In the second part of this thesis, we focus on wideband systems as DS-CDMA and MC-CDMA ones. During the last two decades, a great bulk of

research activities has been devoted to multiuser detection (MUD) for both direct-sequence (DS) [10] and multicarrier (MC) [11] code-division multiple-access systems, since it allows one to achieve a dramatic performance improvement over simpler single-user detection schemes in those environments where the multiple-access interference (MAI) is the predominant performance-limiting factor.

With reference to a DS-CDMA system, we present a performance analysis of L and WL receivers based on the minimum output energy criterion (MOE), both in the known-channel case [12, 13] and in the unknown channel case [14]. Specifically, we start carrying out a detailed study of the conditions on channel and codes that assure perfect MAI suppression in absence of noise for WL-MOE in both underloaded and overloaded downlink configurations, since there was a lack of this issue in the letterature. Moreover, the ideal implementation of the L-MOE and WL-MOE receivers requires perfect knowledge of two quantities [15]: the *autocorrelation matrix* (ACM) of the received signal, and the *received signature* of each user to be demodulated. These two quantities can be estimated in practice from a finite number of samples at the receiver. Since the received signature is a distorted version of the transmitted one due to the effects of the unknown channel response, the channel estimation (CE) is a necessary step for the implementation of both the L- and WL-MOE receivers. To gain more insight about these points, at first we evaluate the performance degradation due to finite-sample ACM estimation in the known-channel case. In particular this analysis is carried out with reference to two different data-estimated implementations of the L-MOE and WL-MOE receivers: the SMI (sample matrix inversion) receiver (referred to as L-SMI and WL-SMI), which employs a sample estimate of the data autocorrelation matrix, and the SUB (subspace) receiver (referred to as L-SUB and WL-SUB), which exploits the properties of the eigenvalue decomposition (EVD) of the data autocorrelation matrix to reduce the effects of estimation errors. The results of this analysis are easily interpretable formulas, which allow one to obtain clear insights about the effects of different parameters on performances. Moreover, the results of the analysis shows that the WL-MOE receiver is more sensitive than its linear counterpart to finite sample-size effects associated to ACM estimation, and it generally requires subspace-based implementation to achieve in practice the performance gains predicted by theory.

Successively, we extend [14] the previous analysis incorporating the effects due to CE on the synthesis of the L- and WL-MOE receivers. The conventional method for CE is to periodically transmit training sequences of data

that are known *a priori* to the receiver. However, such a scheme might lead to a significant waste of power and bandwidth resources in mobile communication systems, especially when channel conditions require the use of long training sequences and/or frequent repetition of training. Consequently, the past few decades have witnessed a huge number of contributions in the area of *blind* CE approaches. Blind CE approaches relying on second-order statistics (SOS) of the received data are particularly attractive since they require far fewer samples than traditional methods based on higher-order statistics [16]. Among existing SOS-based approaches, the subspace CE method is one of the most studied blind algorithm for its robustness and our analysis is developed with reference to this method. We derive easily interpretable formulas that with reference to subspace-based receivers implementations, show that for moderate-to-high values of the SNR, the errors in estimating the L-SUB-CE and WL-SUB-CE receivers are essentially due to ACM estimation. This is not true for the L-SMI-CE and WL-SMI-CE receivers, for which CE errors undesirably combine with ACM errors. Therefore, when considering finite sample-size implementation, the blind WL-MOE receiver is able to assure a significant performance gain (for low-to-moderate values of the SNR) with respect to its linear counterpart only when it is built by resorting to the more sophisticated subspace-based implementation. In this case, for a given channel length, it allows one to work with an increased number of users, making so it a viable choice in heavily-congested DS-CDMA networks.

Finally, in the last part of this thesis we deal with MC-CDMA wireless networks employing *frequency-domain spreading* because, at high data-rates (of the order of several hundred megabits/s), the common single-carrier DS-CDMA technology becomes impractical, due to both severe multipath-induced intersymbol interference (ISI) and synchronization difficulties. In particular, the analysis of L and WL multiuser detection for both CP- and ZP-based configurations is carried out [17, 18]. The problem is to derive mathematical conditions that guarantee perfect symbol recovery in the absence of noise for either CP-based or ZP-based MC-CDMA downlink transmissions, which employ frequency-domain symbol-spreading. This issue is important also for the synthesis of MMSE receivers, since the performances of MMSE detectors strongly depend on the existence of the corresponding ZF solutions. The conditions that we derive are channel-independent and are expressed in terms of relatively simple system design constraints, regarding the maximum number of allowable users and their spreading sequences. Specifically, with reference to FIR L-MUD receiving structures, it is known in letterature that perfect sym-

bol recovery is guaranteed in a ZP-based downlink, for any FIR channel of order smaller that the cyclic prefix length, as long as the number of users is smaller than the number of subcarriers (underloaded systems) and the code vectors are linearly independent. In general, a similar feature does not hold for CP-based transmissions. Thus, we show that universal L-ZF-MUD can be guaranteed even for the underloaded CP-MC-CDMA downlink, provided that the spreading codes are judiciously designed. On the other hand, a detailed study of the conditions assuring FIR WL-MUD perfect symbol recovery in both CP- and ZP-based systems is lacking. Consequently, we show also that, if appropriate complex-valued spreading codes are employed, universal WL-ZF multiuser detectors can be designed even for overloaded CP-MC-CDMA and ZP-MC-CDMA systems.

The outline of the thesis is the following:

Chapter 1 presents the general framework. We introduce the strategies for allocating the available resources among users sharing a common wireless communication channel.

Chapter 2 addresses the basic concepts associated with equalizer design, a fundamental step to introduce the Widely-Linear processing on which this thesis is based.

Chapter 3 addresses the constant modulus (CM) criterion applied to narrowband systems. More in detail, we analyze the CM-cost function under the general assumptions that improper modulation schemes are employed. Such an analysis allows one to determine a broad family of undesired minima of the L-CM cost function, which do not lead to perfect symbol recovery in the absence of noise. Successively we deal with the problem of designing WL equalizers for both real- and complex-valued improper modulation schemes, by proposing a constrained widely-linear constant modulus equalizer able to recover perfectly the symbols in the absence of noise.

Chapter 4 establishes finite-sample performance results for WL- multiuser receivers in DS-CDMA systems, as well as their comparison with conventional L- ones. Specifically, the analysis is carried out with reference to Minimum Output-Energy (MOE) criterion: first we compare the ideal signal-to-interference-plus-noise-ratio (SINR) performances of the WL-MOE

and L-MOE receivers and then the SINR degradation of the data-estimated WL-MOE receivers is accurately evaluated and compared with that of its linear counterpart both in the perfectly-known and unknown channel cases.

Chapter 5 focuses on multiuser detection for MC-CDMA systems employing cyclic-prefixed (CP) or zero-padded (ZP) transmission techniques. For both systems, we consider the linear and the widely-linear receiving structures, showing that, under certain assumptions, L-FIR and WL-FIR universal zero-forcing (ZF) multiuser detectors can be synthesized. Thus, in the absence of noise, it is assured a perfect symbol recovery for each user, regardless of the underlying frequency-selective channel. Finally, some spreading code examples satisfying the design rules are provided as well.

Chapter 1

Multiple Access Systems

In this chapter we shortly describe the strategies for allocating the available resources among users sharing a common wireless communication channel. In particular we focus our attention only on the so-called channel-partitioning access methods, in which a fixed allocation of the channel resources, frequency, time or spreading code are implemented. The three basic fixed-assignment multiple access methods are TDMA, FDMA and CDMA. The choice of an access method has a great impact on the performances and QOS provided by a communication channel.

1.1 Multiple User Environments

A multiuser system is any system in which the available system resources must be divided among different users. This means that several users may access a common channel to communicate with other users. Since the users are not necessarily in a common location the term multiple access is used. The strategies for allocating the available resources among users sharing a common communication channel depend on the traffic properties, on the network topologies, and on the channel characteristics [19, 20, 21, 22]. We will concentrate on the physical layer strategies that allow the sharing of the available resources with acceptable performance. A multiuser channel is any channel that must be shared among multiple users. There are two different types of multiuser channels: the *uplink* channel and the *downlink* channel. A downlink channel, also called a broadcast channel or forward channel, has one transmitter sending to many receivers. Since the signals transmitted to all users originate from the downlink transmitter, the transmitted signal is the sum of signals transmitted

to all users. Thus, the total signaling dimensions and power of the transmitted signal must be divided among the different users. Synchronization of the different users is relatively easy in the downlink since all signals originate from the same transmitter, although multipath in the channel can destroy synchronization. Another important characteristic of the downlink is that both signal and interference are distorted by the same channel. In particular, the kth user's signal and all interfering signals pass through the kth user's channel to arrive at the kth user's receiver. This is a fundamental difference between the uplink and the downlink, since in the uplink signals from different users are distorted by different channels. Examples of wireless downlinks include all radio and television broadcasting, the transmission link from a satellite to multiple ground stations, and the transmission link from a base station to the mobile terminals in a cellular system.

An uplink channel, also called reverse channel, has many transmitters sending signals to one receiver, where each signal must be within the the total system bandwidth. However, in contrast to the downlink, in the uplink each user has an individual power constraint associated with its transmitted signal. In addition, since the signals are sent from different transmitters, these transmitters must coordinate themselves if signal synchronization is required. Moreover, the signals of the different users in the uplink are transmitted through different channels. Examples of wireless uplinks include laptop wireless LAN cards transmitting to a wireless LAN access point, transmissions from ground stations to a satellite, and transmissions from mobile terminals to a base station in cellular systems. Most communication systems are bi-directional, and hence consist of both uplinks and downlinks. The radio transceiver that sends to users over a downlink channel and receives from these users over an uplink channel is often refered to as an access point or base station. It is generally not possible for radios to receive and transmit on the same frequency band due to the interference that results. Thus, bi-directional systems must separate the uplink and downlink channels, typically using time or frequency dimensions. This separation is called duplexing. In particular, time-division duplexing (TDD) assigns orthogonal timeslots to a given user for receiving from an access point and transmitting to the access point, and frequency-division duplexing (FDD) assigns separate frequency bands for transmitting to and receiving from the access point. An advantage of TDD is that bi-directional channels are typically symmetrical in their channel gains, so channel measurements made in one direction can be used to estimate the channel in the other direction. This is not necessarily the case for FDD in

frequency-selective fading: if the frequencies assigned to each direction are separated by more than the coherence bandwidth of the multipath wireless channel, then these channels will exhibit independent fading.

1.2 Multiple Access Techniques for narrowband and wideband systems

In this section we describe the basic properties of the different access techniques following the lines outlined in [19, 21, 22, 23]. The multiple access in a communication system, is done by dividing the signaling dimensions along the time, frequency and/or code space axes. Efficient allocation of signaling dimensions between users is a key design aspect of both uplink and downlink channels, since bandwidth is usually scarce and/or very expensive. In general, applications with continuous transmission and delay constraints require channel-partitioning access methods. This would be the case, for example, with digitized voice traffic, data file transfer or fax-simile transmission. Typical access methods in this context are time-division, frequency-division, code-division, or hybrid combinations of these techniques. However, if the transmission to be transmitted is intermittent or bursty in nature, channel-partitioning access methods can result in communication resources being wasted for much of the duration of the connection. Random access methods, using some form of random channel allocation which does not guarantee channel access, provide a more efficient and flexible way of managing channel access for communicating short bursty messages. In general, the choice of whether to use channel-partitioning or random access, and which specific fixed or random access technique to apply, will depend on the system applications, the traffic characteristics of the users in the system, the performance requirements, and the characteristics of the channel and other interfering systems operating in the same bandwidth. In the sequel we do not analyze random access techniques.

In addition to time and bandwidth, another resource available in wireless systems is space. For example, if one transmitter-receiver pair is sufficiently far away from another, then the mutual interference between them is attenuated enough so as to be negligible. Thus, wireless resources can be utilized more efficiently by employing spatial reuse, which forms the basis for cellular communication systems. On the other hand directional antennas add an additional angular dimension which can also be used to channelize the signal space: this technique is called space-division multiple access (SDMA). In the following subsections, we shortly describe time-division multiple access

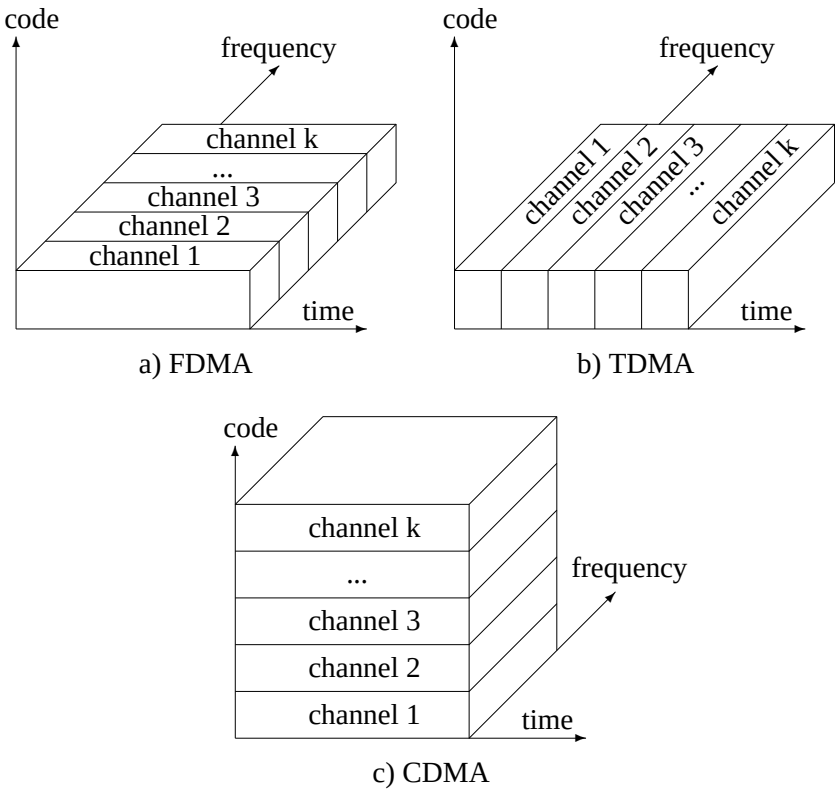

Figure 1.1: The three basic channel-partitioning techniques

(TDMA), frequency-division multiple access (FDMA) and code-division multiple access (CDMA) (for details about CDMA technique see chapter 4), that are pictorially represented in Figure 1.1.

1.2.1 Frequency Division Multiple Access (FDMA)

FDMA is the simplest and oldest form of multiplexing. In FDMA systems, the band available for the service is divided into several non-overlapping sub-channels and each user is assigned a different frequency channel. Once assigned, a sub-channel is held by a user until the user transmits. The principle at the base of this technique is that even if the total system bandwidth is

large and therefore subject to frequency-selective fading, dividing the available bandwidth into subchannels under the assumption that the subchannels are sufficiently narrowband, they will not experience frequency-selective fading. Between two adjacent sub-channels there is typically a guard interval to compensate for imperfect filters, adjacent channel interference, and spectral spreading due to Doppler and therefore to facilitate the separation of different user signals at the receiver. To limit the waste of resources due to the insertion of the guard bands, it is necessary to have a good frequency synchronization. Moreover, it is difficult to assign multiple channels to the same user under FDMA, since this requires the receivers to simultaneously demodulate signals received over multiple frequency channels.

1.2.2 Time Division Multiple Access (TDMA)

In a TDMA system, a number of users share the same frequency band by taking assigned turns in using the channel. Therefore, with this technique, the system dimensions are divided along the time axis into nonoverlapping channels, and each user is assigned a different cyclically-repeating time slot. These TDMA channels occupy the entire system bandwidth, as opposed to FDMA, where each user gets only a portion of the bandwidth. Therefore, in a FDMA system, each subchannel is typically flat-fading, which implies that it is not needed an equalization filter at the receiver. Conversely, the TDMA channel is typically frequency-selective, and thus the receiver must implement some form of equalization. The time slots are also organized in frames, where each frame contains the time slot for a certain number of users plus overhead bits carrying signaling information. With TDMA a transmit controller assigns time slots to users and an assigned time slot is held by a user until the user releases it. At the receiving end, a receiver station synchronizes to the TDMA signal frame and extracts the time slot designated for that user. This modus operandi implies that transmission is not continuous for any user. Therefore, digital transmission techniques which allow for buffering are required. The fact that transmission is not continuous simplifies overhead functions such as channel estimation, since these functions can be done during the time slots occupied by other users. TDMA also has the advantage that it is simple to assign multiple channels to a single user by simply assigning him multiple time slots, in order to provide for example different transmission rates to different users. This is particularly useful for networks supporting multimedia applications, where different media require different rates.

The heart of a TDMA system is synchronization that is necessary to main-

tain orthogonal time slots in the received signals (not needed for FDMA systems). For example multipath channels can destroy time-division orthogonality in both uplinks and downlinks if the multipath delays are a significant fraction of a time slot. However, also for flat-fading channels, the synchronization is needed al least for uplink channels. In fact, in the uplink channel the users transmit over different channels with different respective delays. To maintain orthogonal time slots in the received signals, the different uplink transmitters must be synchronized. Therefore, between any two adjacent time slots often there are time guard intervals.

Moreover, we can underline that TDMA and FDMA have a different behavior respect to the interference. In fact, in a FDMA system narrowband interference affects only one subchannel, whereas in TDMA, the same interference affects all the channels. At the same time, whereas all the power of the narrowband interference acts over one subchannel, in a TDMA system the interference power is split among all the subchannels and then each subchannel has to cope only with a portion of the interference power. Hence, TDMA is more robust to narrowband interference than FDMA. Conversely, by dual arguments, if the interference is impulsive in time, it is better to use FDMA rather than TDMA. Finally we remark that TDMA, in combination with FDMA, is used in the GSM, PDC, IS-54, and IS-136 digital cellular phone standards.

1.2.3 Code Division Multiple Access (CDMA)

FDMA and TDMA strategies assign to different users only portions of the available frequency or time. A different philosophy is followed by CDMA methods where, in principle, every user can get the whole bandwidth and the whole time. Specifically, in CDMA, which we will review with more details in Chapter 4, the information signals of different users are modulated by orthogonal or non-orthogonal spreading codes. The resulting spread signals simultaneously occupy the same time and bandwidth. The receiver uses the spreading code structure to separate out the different users. It is simple to allocate multiple channels to one user with CDMA by assigning to that user multiple codes. These characteristics can be utilized for example in a cellular system to improve the handoff procedure. Specifically in cellular systems, when a mobile passes from one cell to the next, it has to switch the link from the old base station to the new one. This procedure is known as handoff. In 2G systems, where adjacent cells transmit over nonoverlapping frequency bands, the handoff is hard, meaning that the receiver has to switch from the frequency band to the other, as it crosses the boundary from one cell to the next. Con-

versely, adjacent cells in a CDMA cellular network use the same frequency band. Hence, when a mobile moves from one cell to the next, it can communicate with both cells and even combine the two signals advantageously. Only when a reliable link has been established with the new station does the mobile user stop communicating with the previous station. This technique is called soft handoff. Turning to the analysis of CDMA systems, we note that downlinks typically use orthogonal spreading codes such as Walsh-Hadamard codes, although the orthogonality can be degraded by multipath. Uplinks generally use non-orthogonal codes due to the difficulty of user synchronization and the complexity of maintaining code orthogonality in uplinks with multipath.

One of the main advantages of CDMA with respect to TDMA and TDMA is that little dynamic coordination of users in time or frequency is required, since the users can be separated by the code properties alone. In addition, since TDMA and FDMA carve up the signaling dimensions orthogonally, there is a hard limit on how many orthogonal channels can be obtained. This is also true for CDMA using orthogonal codes. However, if non-orthogonal codes are used, there is no hard limit on the number of channels that can be obtained. However, because non-orthogonal codes cause mutual interference between users, the more users that simultaneously share the system bandwidth using non-orthogonal codes, the higher the level of interference, which degrades the performance of all the users. Moreover since the CDMA waveforms occupy the whole spectrum and time available for transmission, the transmission is inherently robust against selective fading and against narrowband interference. In fact, as we will see in the chapter 4, CDMA is built on Spread Spectrum technique that is able to mitigate the performance degradation due to intersymbol and narrowband interference.

Nevertheless, the advantages that we have underline are paid for the increased complexity at the receiver. In addition since CDMA systems are inherently affected by multi-access interference, it is particularly important to adapt the power used on each link in order to limit the detrimental effects of mutual interference. A typical problem is the so-called *near-far effect* that arises in the uplink because the channel gain between a user's transmitter and the receiver is different for different users. Specifically, users near to a base station or an access point, can create a huge interference toward users accessing the same station from the boundary of the cell. To cope with this problem and try to maximize the number of users accessing the system with satisfying QoS, it is necessary to implement a power control (i.e., a feedback mechanism that forces the users to adapt their transmission power depending on the dis-

tance to the base station). Thus, power control is used such that the received signal power of all users is roughly the same. This form of power control, which essentially inverts any attenuation and/or fading on the channel, causes each interferer to contribute an equal amount of power, thereby eliminating the near-far effect. CDMA systems can also use a multiuser detector (MUD) to reduce interference between users. We will see in chapters 4 and 5 some results with reference to these arguments. Finally we note that CDMA is used for multiple access in the IS-95/cdmaOne digital cellular standards, with orthgonal spreading codes on the downlink and a combination of orthogonal and non-orthogonal codes on the uplink . It is also used in the W-CDMA and cdma2000 digital cellular standards .

Now some important remarks are needed. TDMA and FDMA techniques are more appropriate for narrowband systems because user transmissions are restricted to separate narrowband channels. In this scenario a multiuser system is simplified and can be approximated by a collection of point-to-point *non-interfering* links and physical-layer issues are essentially point-to point ones. Therefore the design complexity of the multiple access and interference management are simplified. From this perspective, the description and analysis of a narrowband system is the same for uplink and downlink. On the other hand, the philosophy of a CDMA system is different because all transmission are spread to the entire bandwidth and are hence *wideband*. As a consequence, the multiple access and interference management strategies are different, in the sense that all users share all degrees of freedom and therefore interfere with each other: the system is *interference-limited* rather than *degree-of-freedom-limited*.

Chapter 2

Linear and Widely Linear Equalizers

In this chapter we introduce the basic concepts associated with equalizer design considering a single-input/single-output (SISO) system model. This model, in fact, is the basic mathematical model adopted in the design of communication systems. Successively we describe the basic characteristics of the widely linear (WL) processing and we compare it with the standard linear equalization method.

2.1 Introduction

The wireless radio channel poses a severe challenge as a medium for reliable high-speed communication. It is not only susceptible to noise, interference, and other channel impediments, but these impediments change over time in unpredictable ways due to user movement [19]. Therefore, wireless channels may exhibit frequency selective fading and Doppler shift. Frequency-selective fading gives rise to intersymbol interference (ISI), which can cause an irreducible error floor when the modulation symbol time is on the same order as the channel delay spread. Doppler causes spectral broadening, which leads to adjacent channel interference. In this chapter we focus only on the effects of frequency-selective fading neglecting Doppler effects, following the lines outlined in [19].

In a broad sense, equalization defines any signal processing technique used at the receiver to alleviate the ISI problem caused by delay spread. Signal processing can also be used at the transmitter to make the signal less susceptible

to delay spread: spread spectrum and multicarrier modulation fall in this category of transmitter signal processing techniques. We will examine these methods in chapters 4 and 5 with reference to CDMA systems. Equalizer design must typically balance ISI mitigation with noise enhancement, since both the signal and the noise pass through the equalizer, which can increase the noise power. Equalization techniques fall into two broad categories: linear and nonlinear. The linear techniques are generally the simplest to implement. On the other hand, nonlinear equalizers suffer less from noise enhancement than linear equalizers, but typically entail higher complexity. Among nonlinear equalization techniques, decision-feedback equalization (DFE) is the most common, since it is fairly simple to implement and generally performs well. Moreover, among nonlinear equalization techniques, the optimal equalization technique is maximum likelihood sequence estimation (MLSE). Unfortunately, the complexity of this technique grows exponentially with the length of the channel, and is therefore impractical for most channels of interest. Linear and nonlinear equalizers are typically implemented using a transversal or lattice structure. We focalize on the transversal structure, that is, a filter with N - 1 delay elements and N taps with tunable complex weights.

Most equalizers are implemented digitally after A/D conversion, since such filters are small, cheap, easily tuneable, and very power efficient. This chapter mainly focuses on digital equalizer implementations and moreover, our analysis of equalization is based on the equivalent lowpass representation of bandpass systems.

2.2 Preliminaries

Let consider the block diagram in figure 2.1 of an equivalent lowpass end-to-end system with a digital equalizer. The information symbol $s(k)$ is passed through a pulse shape filter $x(t)$ that improves the spectral properties of the transmitted signal. This pulse shape is under the control of the system designer. Then the signal waveform is transmitted over the ISI channel with impulse response $g(t)$, whose effects are outside the designer's control and generally have a random nature. At the receiver front-end white Gaussian noise $n(t)$ is added to the received signal for a resulting signal $u_a(t)$

$$u_a(t) = s(t) * x(t) * g(t) + n(t) \tag{2.1}$$

with $s(t) = \sum_n s(n)\delta(t - nTs)$, the pulse train of information symbols.

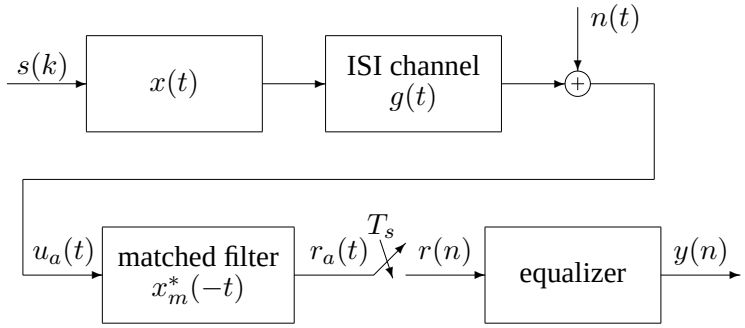

Figure 2.1: End-to-end system (equivalent lowpass representation

In order to obtain a digital version of the received signal, at first $u_a(t)$ passes through an analog matched filter $x_m(t)$ to obtain output $r_a(t)$, which is then sampled via an A/D converter. The purpose of the matched filter is to maximize the signal-to-noise ratio (SNR) of the signal before sampling and subsequent processing. We note that in AWGN the SNR of the received signal is maximized prior to sampling by using a filter that is matched to the pulse shape. This result indicates that SNR prior to sampling is maximized by passing $u_a(t)$ through a filter matched to $x(t) * g(t)$, so ideally we would have $x_m(t) = x(t) * g(t)$. However, since in general the channel impulse response $g(t)$ is time-varying and analog filters are not easily tuneable, it is generally not possible to have $x_m(t) = x(t) * g(t)$. However, $x_m(t)$ is chosen in such a way that good performances are assured. Often $x_m(t)$ is matched to the transmitted pulse shape $x(t)$, which is the optimal pulse shape when the channel is ideal, i.e, $g(t) = \delta(t)$, but this design is clearly suboptimal when $g(t) \neq \delta(t)$. The fact that $x_m(t)$ cannot be matched to $x(t) * g(t)$ can result in significant performance degradation and also makes the receiver extremely sensitive to timing error. These problems are somewhat mitigated by sampling $r_a(t)$ at a rate much faster than the symbol rate and designing the equalizer for this oversampled signal. This process is called fractionally-spaced equalization for which we will see an application in chapter 3. The equalizer output then provides an estimate of the transmitted symbol. This estimate is then passed through a decision device that rounds the equalizer output to a symbol in the alphabet of possible transmitted symbols.

Let $c_a(t)$ denote the combined baseband impulse response of the transmit-

ter, channel, and matched filter:

$$c_a(t) \stackrel{\triangle}{=} x(t) * g(t) * x_m^*(-t). \tag{2.2}$$

Then the matched filter output is given by

$$r_a(t) = s(t) * c_a(t) + w_a(t) = \sum_{k=-\infty}^{-\infty} s(k) c_a(t - kT_s) + w_a(t) \tag{2.3}$$

where $w_a(t) = n(t) * x_m^*(-t)$ is the equivalent baseband noise at the equalizer input and T_s is the symbol time. If we let $c(k) = c_a(kT_s)$ denote samples of $c_a(t)$ every T_s seconds then sampling $r_a(t)$ every T_s seconds yields the discrete time signal $r(k) = r_a(kTs)$ given by

$$\begin{aligned} r(k) &= \sum_{n=-\infty}^{+\infty} s(n) c_a(kT_s - nT_s) + w_a(kT_s) = \sum_{n=-\infty}^{+\infty} s(n) c(k-n) + n(k) \\ &= s(k)c(0) + \sum_{n \neq k} s(n) c(k-n) + n(k) \end{aligned} \tag{2.4}$$

where the first term in (2.4) is the desired data bit, the second term is the ISI, and the third term is the sampled baseband noise $n(k) \stackrel{\triangle}{=} w_a(kT_s)$. We see from (2.4) that we get zero ISI if $c(k-n) = 0$ for $n \neq k$, i.e. $c(n) = \delta(n)c(0)$. In this case (2.4) reduces to

$$r(k) = s(k)c(0) + n(k) \tag{2.5}$$

If the combined baseband impulse response of the transmitter, channel, and matched filter, also called composite channel impulse response, $c_a(t)$ spans L_c symbol periods, i.e., $c_a(t) = 0$ for $t \notin [\,0, L_c\, T_s\,)$, after sampling $r_a(t)$ at baud rate $1/T_s$, the expression of the kth ($k \in \mathbb{Z}$) received time-discrete signal $r(k)$ is given by

$$r(k) = \sum_{n=0}^{L_c - 1} c(n)\, s(k-n) + n(k). \tag{2.6}$$

We underline that the baud-rate sampling has a single-input/single-output (SISO) nature; instead, as we will in chapter 3, the fractionally-spaced equalizers have a single-input/multi-output (SIMO) nature.

2.3 Linear Equalizers

In this section we assume a linear equalizer implemented via an L_e-tap transversal filter (where the equalizer length is expressed in symbol intervals). The length of the equalizer L_e is typically dictated by implementation considerations, since a large L_e usually entails higher complexity. For a given equalizer size L_e the equalizer design must specify the tap weights for a given channel frequency response (note that in the hypothesis of time-varying channel the equalization design must also specify the algorithm for updating these tap weights as the channel varies). Recall that our performance metric in wireless systems is the probability of error, so for a given channel the optimal choice of equalizer coefficients would be the coefficients that minimize probability of error. Unfortunately it is extremely difficult to optimize the equalizer coefficients with respect to this criterion. Therefore we must use an indirect optimization that balances ISI mitigation with the prevention of noise enhancement. We now describe two linear equalizers: the Zero Forcing (ZF) equalizer and the Minimum Mean Square Error (MMSE) equalizer. The former equalizer cancels all ISI, but can lead to considerable noise enhancement. The latter technique minimizes the expected mean squared error between the transmitted symbol and the symbol detected at the equalizer output, thereby providing a better balance between ISI mitigation and noise enhancement. Because of this more favorable balance, MMSE equalizers tend to have better BER performance than equalizers using the ZF algorithm.

Since we are considering linear equalizers, their output $y(k)$ is a linear combination of the input samples $r(k)$. Moreover, we have assumed that a linear equalizer is implemented via an L_e-tap transversal filter, therefore, to compensate for ISI and noise, namely, to produce a reliable estimate of the symbol $s(k)$ a linear equalizer has to jointly elaborate L_e consecutive symbols,

$$y(k) = \sum_{j=0}^{L_e-1} f_j^* \, r(k-j). \tag{2.7}$$

We can note that the equalizer could introduce a equalization delay, but for notational simplicity we neglect it. In the chapter 3 we take in account a possible equalization delay. If we define $\mathbf{r}(k) \stackrel{\triangle}{=} [r(k), r(k-1), \ldots, r(k-L_e+1)]^T \in \mathbb{C}^{L_e}$ and $\mathbf{f} = [f_0, f_1, \ldots, f_{L_e-1}]^T \in \mathbb{C}^{L_e}$ collects all the equalizer's parameters, we can rewrite the input-output relationship of an L-FIR equalizer (2.7) as

$$y(k) = \mathbf{f}^H \mathbf{r}(k). \tag{2.8}$$

Accounting for (2.6), the vector $\mathbf{r}(k)$ can be expressed as

$$\mathbf{r}(k) = \boldsymbol{\mathcal{B}}\,\mathbf{s}(k) + \mathbf{n}(k)\,, \tag{2.9}$$

where $\mathbf{s}(k) \triangleq [s(k),\, s(k-1),\, \ldots,\, s(k-K+1)]^T \in \mathbb{C}^K$, with $K \triangleq L_e + L_c - 1$, whereas

$$\boldsymbol{\mathcal{B}} \triangleq \begin{bmatrix} c(0) & \ldots & c(L_c-1) & 0 & 0 \\ 0 & c(0) & \ldots & c(L_c-1) & 0 \\ \vdots & \ddots & \ldots & \ddots & \vdots \\ 0 & 0 & c(0) & \ldots & c(L_c-1) \end{bmatrix} \in \mathbb{C}^{L_e \times K} \tag{2.10}$$

is a Toeplitz matrix and, finally, $\mathbf{n}(k) \triangleq [n(k),\, n(k-1),\, \ldots,\, n(k-L_e+1)]^T \in \mathbb{C}^{L_e}$.

We remark that the vector $\mathbf{f}$ can be chosen on the base of different criterions, such as, for example, ZF or MMSE criterions that we describe in the following. To simplify the analysis of this criterions we suppose that $n(k) \in \mathbb{C}$ is a sequence of i.i.d. *proper* or circular [1] zero-mean random variables, statistically independent of $s(n)$, with variance $\sigma_n^2 \triangleq \mathrm{E}[|n(k)|^2]$. We can, for example, suppose that the receiver is equipped with a noise-whitening filter. No specific assumption is introduce for the symbols $\{s(k)\}$ because we will see in the Section 2.4 how this choice will influence the equalization design.

2.3.1 Linear ZF Equalizers

The idea of the ZF (zero-forcing) criterion is to set to zero the ISI contribution, as suggested by the name. Therefore the equalizer output $y(k)$ should be equal to the symbol that is transmitted

$$y(k) = s(k). \tag{2.11}$$

In the absence of noise, imposing the zero-forcing condition (2.11), accounting for equations (2.9) and (2.8), leads to the linear equation system

$$\mathbf{f}^H \boldsymbol{\mathcal{B}} = \mathbf{e}_1^T \;\Leftrightarrow\; \boldsymbol{\mathcal{B}}^H \mathbf{f} = \mathbf{e}_1, \tag{2.12}$$

where $\mathbf{e}_1 \triangleq [1, 0, \ldots, 0]^T \in \mathbb{R}^K$. Nevertheless the ZF equalizer is not implementable as a finite impulse response (FIR) filter, in fact (2.12) defines a

linear equation system with a number of equations smaller than the number of unknown quantities, therefore the system has no solution. In fact in the baud-spaced case the column dimension of $\boldsymbol{\mathcal{B}}$ always exceed the row dimension

$$L_e < K = L_e + L_c - 1. \tag{2.13}$$

We will see in the section 3.3 of the chapter 3 that if we adopt a fractionally-spaced equalizer the ZF equalizer can be implemented as a finite impulse response (FIR) filter. An alternative proof of the fact that a baud-spaced ZF equalizer cannot be implemented as a FIR filter, can be conducted considering that the system (2.12) is consistent if and only if (iff) [24]

$$\boldsymbol{\mathcal{B}}^H(\boldsymbol{\mathcal{B}}^H)^{-}\mathbf{e}_1 = \mathbf{e}_1, \tag{2.14}$$

where $(.)^-$ denotes the $\{1\}-$inverse of the matrix $\boldsymbol{\mathcal{B}}^H$. If the channel matrix $\boldsymbol{\mathcal{B}}$ is full-column rank, i.e., rank$(\boldsymbol{\mathcal{B}}) = K$, it results that $\boldsymbol{\mathcal{B}}^H(\boldsymbol{\mathcal{B}}^H)^- = \mathbf{I}_K$ and, then, this system turns out to be consistent. Nevertheless the full-column rank requirement implies that $\boldsymbol{\mathcal{B}}$ must have at least as many as rows as columns, which in the baud-spaced case is never satisfied because the column dimension of $\boldsymbol{\mathcal{B}}$ always exceed the row dimension. Moreover, we note that the ZF equalizer can lead to considerable noise enhancement (see [19] for details). This motivates an equalizer design that better optimizes between ISI mitigation and noise enhancement. One such equalizer is the MMSE equalizer, described in the next subsection.

2.3.2 Linear MMSE Equalizer

In MMSE (minimum-mean-square error) equalization the goal of the equalizer design is to minimize the average mean square error between the transmitted symbol $s(k)$ and its estimate $y(k)$ at the output of the equalizer. In other words the vector $\mathbf{f}$ is chosen to minimize $\mathrm{E}[|y(k) - s(k)|^2]$. Thus, we want to minimize the mean square error:

$$J(\mathbf{f}) \stackrel{\triangle}{=} \mathrm{E}[|y(k) - s(k)|^2] = \mathbf{f}^H \mathbf{R}_{\mathbf{rr}} \mathbf{f} + \mathrm{E}[|s(k)|^2] - 2\mathrm{Re}(\mathbf{f}^H \mathbf{r}_s) \tag{2.15}$$

where $\mathbf{R}_{\mathbf{rr}} \stackrel{\triangle}{=} \mathrm{E}[\mathbf{r}(k)\,\mathbf{r}^H(k)] \in \mathbb{C}^{L_e \times L_e}$ is the autocorrelation matrix of $\mathbf{r}(k)$ and $\mathbf{r}_s \stackrel{\triangle}{=} \mathrm{E}[\,\mathbf{r}(k)\,s^*(k)\,] \in \mathbb{C}^{L_e \times 1}$. It can be shown [19, 25] that the solution of this minimization problem is given by

$$\mathbf{f}_{\text{L-MMSE}} = \mathbf{R}_{\mathbf{rr}}^{-1}\mathbf{r}_s \tag{2.16}$$

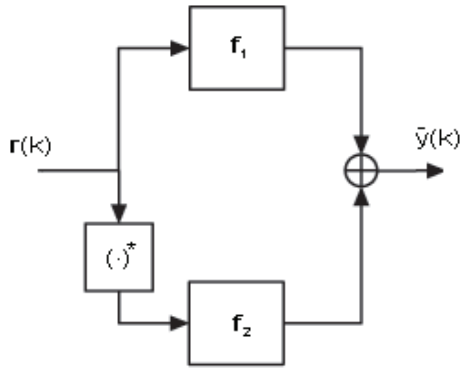

Figure 2.2: Widely-Linear processing scheme

Note that solving for $\mathbf{f}_{\text{L-MMSE}}$ requires the data autocorrelation matrix inversion. Thus, the complexity of this computation is quite high, typically on the order of L_e^2 to L_e^3 operations. Substituting (2.16) in (2.15) we obtain the minimum mean square error

$$J_{\text{L-MMSE}} \stackrel{\triangle}{=} J(\mathbf{f}_{\text{L-MMSE}}) = \mathrm{E}[|s(k)|^2] - \mathbf{r}_s^H \, \mathbf{R}_{\mathbf{rr}}^{-1} \, \mathbf{r}_s \tag{2.17}$$

It is possible to show a classical result of the estimation theory that is the estimate $y(k)_{\text{opt}}$ that minimize the MS error is the regression or the conditional expectation value $\mathrm{E}[s(k)|\mathbf{r}(k)]$. Moreover in [19, 25] it is underlined that the MMSE infinite length equalizer is identical to the ZF filter except for the noise term, so in the absence of noise the two equalizers are equivalent.

2.4 Widely linear processing

In this section, based on [2], we present widely-linear processing, the core of this thesis. In particular, we report here the results of [2] that show that when the data are not proper [7] but are improper [1] the widely-linear processing outperforms the linear one.

In the previous subsection 2.3.2 we have underlined that the estimate that minimize the MS error is the regression or the conditional expectation value. In classical estimation theory this result is usually proven when $s(k)$ and $\mathbf{r}(k)$ are real. However, it remains valid when these quantities are complex valued if we redefine correctly the regression concept. Specifically, we observe that if $s(k)$ and $\mathbf{r}(k)$ are real the regression is linear. For complex data this is no

longer true because the regression must be linear both in $\mathbf{r}(k)$ and $\mathbf{r}(k)^*$ and is called *widely linear* (WL). More specifically, we have seen that in L-MSE criterion the estimate (2.8) is found by processing it with (2.16). We report (2.8) here

$$y(k) = \mathbf{f}^H \mathbf{r}(k), \tag{2.18}$$

and we note that $y(k)$ is a scalar product, therefore $y(k)$ is a linear function of the vector $\mathbf{r}(k)$, as defined in classical linear algebra. Consider now the scalar $y_{\text{wl}}(k)$ defined by

$$y_{\text{wl}}(k) = \mathbf{f}_1^H \mathbf{r}(k) + \mathbf{f}_2^H \mathbf{r}^*(k) \tag{2.19}$$

where $\mathbf{f}_1$ and $\mathbf{f}_2$ are two complex vector, see figure 2.2. This is the general form of "linear" regression for complex random variables. It is clear that $y_{\text{wl}}(k)$ is not a linear function of $\mathbf{r}(k)$. However, the moment of order k of $y_{\text{wl}}(k)$ is completely defined from the moments of order k of $\mathbf{r}(k)$ and $\mathbf{r}^*(k)$, which characterizes a form of linearity. This is why (2.19) is called a *wide sense linear* or *widely linear* filter or system.

To show that taking into account WL systems defined by (2.19) instead of strictly linear ones defined by (2.8) can yield significant improvements in estimation problems using complex data, we report here the results of [2]. The problem of WL mean square estimation (WL-MMSE) is to find the vectors $\mathbf{f}_1$ and $\mathbf{f}_2$ in such a way that the MSE between the estimandum $s(k)$ (or $s(k-d)$ if we consider a equalization delay) and the estimate $y_{\text{wl}}(k)$, is minimum. Preliminarily, we observe that the set of scalar complex random variables

$$z = \mathbf{a}^H \mathbf{x} + \mathbf{b}^H \mathbf{x}^* \tag{2.20}$$

where $\mathbf{a}$ and $\mathbf{b}$ are two complex vector, constitutes a linear subspace over the complex field. If we define a scalar product as $< u_1, u_2 > \stackrel{\triangle}{=} \mathrm{E}[u_1^* u_2]$, this linear subspace becomes a Hilbert subspace. Therefore, we can regard (2.19) as the projection of $y_{\text{wl}}(k)$ onto this subspace and due to the orthogonality principle, one has

$$(\, s(k) - y_{\text{wl}}(k) \,) \perp \mathbf{r}(k); \quad (\, s(k) - y_{\text{wl}}(k) \,) \perp \mathbf{r}^*(k). \tag{2.21}$$

The symbol $\perp$ means that all the components of $\mathbf{r}(k)$ or $\mathbf{r}^*(k)$ are orthogonal to $(\, s(k) - y_{\text{wl}}(k) \,)$ under the previous scalar product. In agreement with the definition of scalar product, (2.21) can be rewritten as

$$\mathrm{E}[s^*(k)\, \mathbf{r}(k)] = \mathrm{E}[y_{\text{wl}}^*(k)\, \mathbf{r}(k)]; \quad \mathrm{E}[s^*(k)\, \mathbf{r}^*(k)] = \mathrm{E}[y_{\text{wl}}^*(k)\, \mathbf{r}^*(k)] \tag{2.22}$$

Substituting (2.19) in (2.22) one obtains:

$$\mathbf{R}_{\mathbf{rr}}\,\mathbf{f}_1 + \mathbf{R}_{\mathbf{rr}^*}\,\mathbf{f}_2 = \mathbf{r}_s \tag{2.23}$$

$$\mathbf{R}^*_{\mathbf{rr}^*}\,\mathbf{f}_1 + \mathbf{R}^*_{\mathbf{rr}}\,\mathbf{f}_2 = \mathbf{r}^*_v \tag{2.24}$$

where $\mathbf{R}_{\mathbf{rr}}$ is the autocorrelation matrix that we have defined in (2.15), $\mathbf{R}_{\mathbf{rr}^*} \triangleq \mathrm{E}[\mathbf{r}(k)\,\mathbf{r}^T(k)]$ is the pseudo-correlation matrix and finally $\mathbf{r}_s = \mathrm{E}[s^*(k)\,\mathbf{r}(k)]$ is defined in (2.15) and $\mathbf{r}_v \triangleq \mathrm{E}[s(k)\,\mathbf{r}(k)]$. From (2.23) and (2.24), we find the solutions expressed as

$$\mathbf{f}_1 = \left[\mathbf{R}_{\mathbf{rr}} - \mathbf{R}_{\mathbf{rr}^*}\,(\mathbf{R}^{-1}_{\mathbf{rr}})^*\,\mathbf{R}^*_{\mathbf{rr}^*}\right]^{-1}\left[\mathbf{r}_s - \mathbf{R}_{\mathbf{rr}^*}\,(\mathbf{R}^{-1}_{\mathbf{rr}})^*\,\mathbf{r}^*_v\right] \tag{2.25}$$

$$\mathbf{f}_2 = \left[\mathbf{R}^*_{\mathbf{rr}} - \mathbf{R}^*_{\mathbf{rr}^*}\,\mathbf{R}^{-1}_{\mathbf{rr}}\,\mathbf{R}_{\mathbf{rr}^*}\right]^{-1}\left[\mathbf{r}^*_v - \mathbf{R}^*_{\mathbf{rr}^*}\,\mathbf{R}^{-1}_{\mathbf{rr}}\,\mathbf{r}_s\right] \tag{2.26}$$

The corresponding MSE is obtained by substituting (2.19) in (2.15) with $\mathbf{f}_1$ and $\mathbf{f}_2$ given by (2.25) and (2.26), or it is deduced from the projection theorem:

$$J_{\text{WL-MMSE}} = \mathrm{E}[|s(k)|^2] - (\mathbf{f}_1^H\mathbf{r}_s + \mathbf{f}_2^H\,\mathbf{r}^*_v) \tag{2.27}$$

Let us define the quantity $\Delta J \triangleq J_{\text{L-MMSE}} - J_{\text{WL-MMSE}}$, that is equal to [2]

$$\Delta J = \left[\mathbf{r}^*_v - \mathbf{R}^*_{\mathbf{rr}^*}\,\mathbf{R}^{-1}_{\mathbf{rr}}\,\mathbf{r}_s\right]^H\left[\mathbf{R}^*_{\mathbf{rr}} - \mathbf{R}^*_{\mathbf{rr}^*}\,\mathbf{R}^{-1}_{\mathbf{rr}}\,\mathbf{R}_{\mathbf{rr}^*}\right]^{-1}\left[\mathbf{r}^*_v - \mathbf{R}^*_{\mathbf{rr}^*}\,\mathbf{R}^{-1}_{\mathbf{rr}}\,\mathbf{r}_s\right]. \tag{2.28}$$

ΔJ is always nonnegative because the matrix $\left[\mathbf{R}^*_{\mathbf{rr}} - \mathbf{R}^*_{\mathbf{rr}^*}\,\mathbf{R}^{-1}_{\mathbf{rr}}\,\mathbf{R}_{\mathbf{rr}^*}\right]$ is positive definite (see [2] for details) and it results $\Delta J = 0$ only when $\mathbf{r}^*_v - \mathbf{R}^*_{\mathbf{rr}^*}\,\mathbf{R}^{-1}_{\mathbf{rr}}\,\mathbf{r}_s = \mathbf{0}$. Therefore the error that is obtained with a L-MMSE (2.17) is greater then the errore that is obtained with a WL-MMSE (2.27) and the advantage of the WL-MMSE procedure over the L-MMSE is characterized by the quantity ΔJ.

Now we consider some case studies.

✓ *Jointly Circular Case*:

this situation is characterized by

$$\mathbf{R}_{\mathbf{rr}^*} = \mathbf{0}; \quad \mathbf{r}_v = \mathbf{0} \tag{2.29}$$

It immediately results from (2.26) and (2.25) that (2.29) implies

$$\mathbf{f}_1 = \mathbf{R}^{-1}_{\mathbf{rr}}\mathbf{r}_s \tag{2.30}$$

$$\mathbf{f}_2 = \mathbf{0} \tag{2.31}$$

(2.30) is the same of (2.16). Thus, the assumption of joint circularity [26] implies that the WL-MMSE (2.19) takes the form (2.8) and is strictly linear. Moreover it results that ΔJ, given by (2.28), is equal to zero. In conclusion, in the case of a joint circularity, the strictly linear system (2.8) is sufficient to reach the best performance. This is one of the arguments justifying the interest of circularity. However, even if circularity appears in many practical situations, there are cases where it cannot be introduced, as we will see in the next chapters.

✓ *Circular observation*:

we suppose that circularity is only valid for the observation and is characterized by

$$\mathbf{R}_{\mathbf{r}\mathbf{r}^*} = \mathbf{0}, \tag{2.32}$$

whereas no specific assumption is introduced for the estimandum $s(k)$. In this case it results that (2.26) and (2.25) become

$$\mathbf{f}_1 = \mathbf{R}_{\mathbf{r}\mathbf{r}}^{-1}\mathbf{r}_s \tag{2.33}$$

$$\mathbf{f}_2^* = \mathbf{R}_{\mathbf{r}\mathbf{r}}^{-1}\mathbf{r}_v \tag{2.34}$$

This means that the term $\mathbf{f}_1^H\mathbf{r}(k)$ in (2.19) is the same as the one obtained when using strictly linear estimation (2.16).

Moreover, from (2.28) it results that

$$\Delta J = \mathbf{r}_v^H\,\mathbf{R}_{\mathbf{r}\mathbf{r}}^{-1}\,\mathbf{r}_v \tag{2.35}$$

Thus, a nonzero vector $\mathbf{r}_v$ necessarily implies an increment of the performance of WL estimation compared with L one.

✓ *Case of a Real Estimandum* $s(k)$:

we suppose then that $s(k)$ is real whereas $\mathbf{r}(k)$ is complex (this case appears in many situation, see chapter 4 for a thorough discussion). From (2.26) and (2.25), we find

$$\mathbf{f}_1 = \mathbf{f}_2^* \tag{2.36}$$

because it results that $\mathbf{r}_s = \mathbf{r}_v$. Substituting (2.36) in (2.19), we obtain

$$y_{\text{wl}}(k) = 2\,\text{Re}(\mathbf{f}_1^H\mathbf{r}(k)). \tag{2.37}$$

Similarly, the estimation error takes the form

$$\Delta J = \text{E}[s(k)^2] \;-\; 2\,\text{Re}(\mathbf{f}_1^H\mathbf{r}_s). \tag{2.38}$$

The main property of the estimate (2.37) is that it is real, although there is no reason for the strictly linear estimate to be real, which is not convenient when estimating a real quantity. The advantage of the structure (2.19) with respect to (2.8) is even more clear when the observation $\mathbf{r}(k)$ is circular. In fact, as seen in the case of Circular observation, the vector $\mathbf{f}_1$ is the same as the one that must be used to realize the L-MMSE $\mathbf{f}$. Thus, by using this vector, the two estimators (2.8) and (2.19) become

$$y(k) = \mathbf{f}_1^H \mathbf{r}(k); \;\; y_{\text{wl}}(k) = 2\,\text{Re}(\mathbf{f}_1^H \mathbf{r}(k)) \tag{2.39}$$

and the corresponding errors (2.17), (2.27), are

$$J_{\text{L-MMSE}} = \text{E}[s^2(k)] - \mathbf{f}_1^H \mathbf{r}_s; \;\; J_{\text{WL-MMSE}} = \text{E}[s^2(k)] - 2\,\mathbf{f}_1^H \mathbf{r}_s \tag{2.40}$$

The quantity $\mathbf{f}_1^H \mathbf{r}_s = \mathbf{f}_1^H\, \mathbf{R}_{\mathbf{rr}}\, \mathbf{f}_1$ is positive because $\mathbf{R}_{\mathbf{rr}}$ is a Hermitian matrix. In conclusion, the wide sense linear estimator (2.19) provides a real estimate and a decrease of the error that is twice as great as the strictly linear estimate, which in general is complex.

From these results, we conclude that widely linear systems can yield significant improvements in estimation performance with respect to strictly linear systems generally used, except when the circularity assumption is introduced. In the next chapter we will exploit these results in the channel equalization scenario.

Chapter 3

Constant Modulus Equalizers for Narrowband Systems

In this chapter, the constant modulus (CM) cost function is analyzed under the general assumptions that improper modulation schemes of practical interest are employed and the baseband equivalent of the channel impulse response is complex-valued. Preliminarily, this study is conducted with respect to the linear finite-impulse (L-FIR) fractionally-spaced (FS) blind equalization of FIR frequency- selective channels. The analysis allows one to determine a broad family of undesired minima of the L-CM cost function, which do not allow perfect symbol recovery in the absence of noise. The results developed herein generalize and subsume as a particular case existing studies of the L-CM cost function, which exclusively consider real-valued binary modulations. Successively the chapter deals with the problem of designing widely-linear (WL) fractionally-spaced (FS) equalizers for both real- and complex-valued improper modulation schemes. Specifically, the synthesis of both WL-FS minimum mean-square error and zero-forcing equalizers is discussed, by deriving the mathematical conditions assuring perfect symbol recovery in the absence of noise. We also propose a constrained widely-linear constant modulus equalizer, able to recover perfectly the symbols in the absence of noise. The effectiveness of the proposed equalizers is corroborated by means of computer simulation results.

3.1 Introduction

In digital communications, blind channel equalization techniques allow one to cancel or reduce the intersymbol interference (ISI) introduced by frequency-selective transmission channels, without wasting the available bandwidth resources due to transmission of training sequences. One of the earliest blind receivers, and perhaps the most widely used in practice, is based on the Godard or constant modulus (CM) algorithm [4, 5, 6]. In his original paper Godard observed by simulations that receivers designed by minimizing the constant modulus cost function, have MSE performance similar to those of nonblind Wiener receivers. This striking observation provides strong support for using CM blind receivers because they do not require the cooperation of the transmitter and also achieve near optimal performance (in the sense of minimizing mean square error of the estimate).

Most early detailed studies of the stationary points of the CM cost function were conducted. These analyses [27, 28] have shown that, in the absence of noise and under certain mathematical conditions, all the local minima of the L-CM cost function are global ones and allow one to exactly suppress the ISI, provided that the transmitted symbol sequence is a *proper* [7] complex random process.

On the other hand, in many cases of practical interest, the symbol sequence is an *improper* [1] random process. In this case and when the channel impulse response is complex-valued, the ISI suppression capabilities of L-CM equalizers turn out to be adversely affected. With reference to BPSK modulation this weakness was evidenced in [29], wherein it was pointed out that, besides containing desired local minima, the infinite-length CM linear equalizer also exhibits *undesired* global minima, which do not lead to perfect symbol recovery in the absence of noise. However, the analysis carried out in [29], cannot be directly extended to multidimensional real-valued modulations, as well as to others improper complex-valued modulation schemes of practical interest. Therefore, at the first in this chapter it is shown [30] how to improve and to generalize the results of [29], by determining a broad family of undesired local minima of the L-CM cost function.

Moreover, we can observe that when the transmitted symbol sequence is improper (see section 2.4), it is well-known [2, 3] that a better (in the sense of second-order statistics) estimate of the transmitted symbols can be obtained by resorting to widely-linear (WL) estimators, which jointly process the received signal and its complex conjugate. As a consequence in this chapter, following the lines outlined in [8, 9], we provide a general and unified framework to de-

sign WL equalizers for both real- and complex-valued improper modulations, by deriving the conditions assuring perfect symbol recovery in the absence of noise and providing some insights into the synthesis and analysis of blind WL-CM equalizers.

Specifically, as confirmed by the simulation results, also the WL-CM cost function exhibits undesired global minima which do not lead to perfect symbol recovery in the absence of noise. To overcome this drawback, [8, 9] propose to resort to a constrained WL-CM equalizer assuring perfect symbol recovery in the absence of noise. The proposed designs generalize and subsume as a particular case some previously proposed WL equalizers [31, 32] targeted at real-valued modulations. The effectiveness of the proposed equalizers is corroborated by means of computer simulation results.

3.2 Signal Model

Let us consider a digital communication system employing linear modulation with symbol period T_s. The complex envelope of the received continuous-time signal, after filtering and ideal carrier-frequency recovering, can be expressed as

$$r_a(t) = \sum_{q=-\infty}^{\infty} s(q)\, c_a(t - q\,T_s) + w_a(t)\,, \tag{3.1}$$

where $\{s(n)\}_{n\in\mathbb{Z}}$ is the sequence of the transmitted symbols, $c_a(t)$ denotes the *composite* impulse response (including transmitting filter, physical channel, receiving filter, and timing offset) of the linear time-invariant signal channel and, finally, $w_a(t)$ represents additive noise at the output of the receiving filter. If the channel impulse response $c_a(t)$ spans L_c symbol periods, i.e., $c_a(t) = 0$ for $t \notin [\,0, L_c\,T_s\,)$, after sampling $r_a(t)$ at rate N/T_s, with $N \geq 1$ being an integer number, the expression of the kth ($k \in \mathbb{Z}$) received data block $\mathbf{r}(k) \stackrel{\triangle}{=} [r^{(0)}(k), r^{(1)}(k), \ldots, r^{(N-1)}(k)]^T \in \mathbb{C}^N$, with $r^{(\ell)}(k) \stackrel{\triangle}{=} r_a(k\,T_s + \ell\,T_s/N)$, is given by

$$\mathbf{r}(k) = \sum_{q=0}^{L_c-1} \mathbf{c}(q)\, s(k-q) + \mathbf{w}(k)\,, \tag{3.2}$$

where $\mathbf{c}(k) \stackrel{\triangle}{=} [\,c^{(0)}(k),\, c^{(1)}(k),\, \ldots,\, c^{(N-1)}(k)\,]^T \in \mathbb{C}^N$, with $c^{(\ell)}(k) \stackrel{\triangle}{=} c_a(k\,T_s + \ell\,T_s/N)$ denoting the ℓth *phase* of the discrete-time channel $c(n) \stackrel{\triangle}{=} c_a(nT_s/N)$ and, similarly, the noise vector $\mathbf{w}(k) \stackrel{\triangle}{=}$

$[\,w^{(0)}(k),\, w^{(1)}(k),\, \ldots,\, w^{(N-1)}(k)\,]^T \in \mathbb{C}^N$ collects the noise phases $w^{(\ell)}(k) \triangleq w_a(k\,T_s + \ell\,T_s/N)$, for $\ell \in \{0, 1, \ldots, N-1\}$. It is worth to note that when $N = 1$, the fractionally spaced model (3.2) degenerates in the common baud-spaced model.

To compensate for ISI and noise, namely, to produce a reliable estimate of the symbol $s(k-d)$, with d denoting a suitable *equalization delay*, an equalizer has to jointly elaborate L_e consecutive symbols, by processing the input vector $\mathbf{z}(k) \triangleq [\,\mathbf{r}^T(k), \mathbf{r}^T(k-1), \ldots, \mathbf{r}^T(k-L_e+1)\,]^T \in \mathbb{C}^{NL_e}$ which, accounting for (3.2), can be expressed as

$$\mathbf{z}(k) = \mathbf{C}\,\mathbf{s}(k) + \mathbf{v}(k)\,, \tag{3.3}$$

where $\mathbf{s}(k) \triangleq [s(k),\, s(k-1),\, \ldots,\, s(k-K+1)]^T \in \mathbb{C}^K$, with $K \triangleq L_e + L_c - 1$, whereas

$$\mathbf{C} \triangleq \begin{bmatrix} \mathbf{c}(0) & \ldots & \mathbf{c}(L_c-1) & \mathbf{0}_N & \mathbf{0}_N \\ \mathbf{0}_N & \mathbf{c}(0) & \ldots & \mathbf{c}(L_c-1) & \mathbf{0}_N \\ \vdots & \ddots & \ldots & \ddots & \vdots \\ \mathbf{0}_N & \mathbf{0}_N & \mathbf{c}(0) & \ldots & \mathbf{c}(L_c-1) \end{bmatrix} \in \mathbb{C}^{(NL_e)\times K} \tag{3.4}$$

is a block Toeplitz matrix and, finally, $\mathbf{v}(k) \triangleq [\mathbf{w}^T(k),\, \mathbf{w}^T(k-1),\, \ldots,\, \mathbf{w}^T(k-L_e+1)]^T \in \mathbb{C}^{NL_e}$.

The following customary assumptions will be considered hereinafter:

A1) $s(n) \in \mathbb{C}$ is a sequence of independent and identically distributed (i.i.d.) zero-mean random variables, whose kurtosis $\kappa_s \triangleq \mathrm{E}[|s(n)|^4] - 2\,\mathrm{E}^2[|s(n)|^2] - \left|\mathrm{E}[s^2(n)]\right|^2 < 0$.

A2) $w(n) \triangleq w_a(nT_s/N) \in \mathbb{C}$ is a sequence of i.i.d. *proper* [1] zero-mean random variables, statistically independent of $s(n)$, with variance $\sigma_w^2 \triangleq \mathrm{E}[|w(n)|^2]$.

The condition $\kappa_s < 0$ imposes that the transmitted symbols are "sub-Gaussian" [5], which is the case commonly encountered in digital communications. Assumption A2 is surely satisfied if the continuous-time filter used at the receiving side has (approximatively) a square root raised-cosine impulse response and, more generally, A2 holds if a whitening matched-filter is employed at the receiver.

3.3 Linear Constant Modulus Equalizer

Blind channel equalization consists of designing a linear FIR (L-FIR) equalizer, which is able to extract the desired symbol $s(k-d)$ (with d denoting the equalization delay) by jointly counteracting intersymbol interference (ISI) and noise, without making use of training sequences. Denoting with L_e the equalizer length (expressed in symbol intervals), the input-output relationship of an L-FIR equalizer is

$$y(k) = \mathbf{f}^H \mathbf{z}(k), \tag{3.5}$$

where $\mathbf{f} \in \mathbb{C}^{NL_e}$ collects all the equalizer's parameters and $\mathbf{z}(k)$ is given by (3.3).

The term perfect equalization means that the equalization output $y(k)$ is equal to the symbol that is transmitted for some fixed suitable equalization delay d, i.e.

$$y(k) = s(k-d), \;\; \text{with} \;\; d \in \{0, 1, \ldots, K-1\}. \tag{3.6}$$

In the absence of noise, imposing the so called "zero-forcing" (ZF) condition $y(k) = s(k-d)$ leads to the linear equation system

$$\mathbf{f}^H \mathbf{C} = \mathbf{e}_d^T \;\Leftrightarrow\; \mathbf{C}^H \mathbf{f} = \mathbf{e}_d, \tag{3.7}$$

where $\mathbf{e}d \triangleq [\overbrace{0, \ldots, 0}^{d}, 1, 0, \ldots, 0]^T \in \mathbb{R}^K$, for any $d \in \{0, 1, \ldots, K-1\}$. This system is consistent if and only if (iff) [24]

$$\mathbf{C}^H (\mathbf{C}^H)^{-} \mathbf{e}_d = \mathbf{e}_d, \tag{3.8}$$

where $(.)^{-}$ denotes the $\{1\}-$inverse of the matrix $\mathbf{C}^H$. If the channel matrix $\mathbf{C}$ is full-column rank, i.e., rank$(\mathbf{C}) = K$, it results that $\mathbf{C}^H (\mathbf{C}^H)^{-} = \mathbf{I}_K$ and, then, this system turns out to be consistent regardless of the equalization delay d. In this case, the *minimal norm* solution, i.e., the solution of the constrained optimization problem

$$\mathbf{f}_{\text{ZF}} = \arg \min_{\mathbf{f} \in \mathbb{C}^{NL_e}} \|\mathbf{f}\|^2 \,, \quad \text{subject to } \mathbf{C}^H \mathbf{f} = \mathbf{e}d \,, \tag{3.9}$$

is given by (see, e.g., [24])

$$\mathbf{f}_{\text{ZF}} = \mathbf{C} \, (\mathbf{C}^H \mathbf{C})^{-1} \mathbf{e}_{\text{d}}. \tag{3.10}$$

Therefore, in the sequel, to guarantee [5, 33] the existence of L-FIR zero-forcing (ZF) designs for $\mathbf{f}$, we assume that

A3) $\mathbf{C}$ is full-column rank, i.e., rank$(\mathbf{C}) = K$.

The full-column rank requirement [5, 33], implies that $\mathbf{C}$ must have at least as many as rows as columns, which in the N/T_s-spaced case results in the following equalizer length requirement:

$$NL_e \geq K \;\; \Rightarrow (N-1)L_e \geq L_c - 1. \tag{3.11}$$

Applying the same argument to a L- baud-spaced equalizer ($N = 1$), we understand that in this case the L-FIR-ZF solution does not exist because the column dimension of $\mathbf{C}$ always exceed the row dimension. Moreover, let $C^{(\ell)}(z)$ denote the $\mathcal{Z}$-transform of the ℓth channel phase $\{c^{(\ell)}(k)\}_{k=0}^{L_c-1}$, for $\ell \in \{0, 1, \ldots, N-1\}$, in [5], is underlined that to guarantee the full-column rank requirement, the polynomials $C^{(\ell)}(z)$ share no common root, i.e., are coprime.

To blindly suppress ISI and noise, one can resort to the constant modulus (CM) criterion [4, 5], where the vector $\mathbf{f}$ is chosen such as to minimize the cost function

$$J_{\text{cm}}(\mathbf{f}) \stackrel{\triangle}{=} \text{E}[(\gamma_s - |y(k)|^2)^2]\,, \tag{3.12}$$

where $\gamma_s \stackrel{\triangle}{=} \text{E}[|s(n)|^4]/\text{E}[|s(n)|^2]$ denotes the dispersion coefficient of the transmitted symbol sequence. The cost function (3.12) exhibits a multimodal surface and its minimization does not lead to closed-form solutions. In practice, stochastic gradient descent (SGD) algorithms are commonly employed to minimize $J_{\text{cm}}(\mathbf{f})$, by starting at some location on the surface and following the trajectory of the steepest descent. Since the characteristics of the CM cost function allow one to gain important insights about the expected behaviors of any SGD algorithm that attempts to minimize $J_{\text{cm}}(\mathbf{f})$, such as the popular CM algorithm (CMA) [5], the derivation and classification of the stationary points of the cost function (3.12) has received a great deal of attention as we have underlined in the introduction 3.1. Moreover, again in the introduction we have noted that there is a lack of detailed analysis of the CM cost function (3.12) when the symbol sequence is an improper random process, i.e., its conjugate correlation function does not vanish identically, $R_{ss^*}(n,m) \stackrel{\triangle}{=} \text{E}[s(n)\,s(n-m)] \neq 0$ for some $n, m \in \mathbb{Z}$. The simplest examples of improper modulation formats are all the real-valued ones, such as ASK, BPSK, differential BPSK (DBPSK), for which the conjugate correlation function $R_{ss^*}(n,m)$ trivially boils down to the autocorrelation function $R_{ss}(n,m) \stackrel{\triangle}{=} \text{E}[s(n)\,s^*(n-m)]$, i.e. $R_{ss^*}(n,m) = R_{ss}(n,m)$ for any $n, m \in \mathbb{Z}$.

In the literature, when real-valued modulations are considered, a common assumption [34, 35, 36, 6] is that the baseband channel impulse response (CIR) and the additive noise are also real-valued: in this case, the CM cost surface essentially exhibits the same characteristics of $J_{\text{cm}}(\mathbf{f})$ given by (3.12) when the transmitted symbols are assumed to be proper complex, i.e, all the local minima of $J_{\text{cm}}(\mathbf{f})$ are desired enabling perfect recovery of the transmitted symbols in the absence of noise. A noticeable exception is [29], wherein the transmission of BPSK symbols over a complex baseband channel is considered. In this case, it was argued in [29] that, besides containing desired local minima, the infinite-length CM linear equalizer also exhibits *undesired* global minima, which do not lead to perfect symbol recovery in the absence of noise. However, the analysis carried out in [29] is exclusively targeted at a BPSK modulation and cannot be directly extended to multidimensional real-valued modulations, as well as to others improper complex-valued modulation schemes of practical interest, such as offset QPSK (OQPSK), offset QAM (OQAM), MSK and its variant Gaussian MSK. Therefore, on the basis of these considerations, with reference to a complex-valued CIR, the aim of this section, following our paper [30] is to provide an accurate derivations and classification of a broad family of undesired local minima of $J_{\text{cm}}(\mathbf{f})$ for a more general class of improper modulation schemes of practical interest. Hence the study of the stationary point of (3.12) is herein carried out under the more general assumptions:

A4) besides fulfilling assumption A1, $s(n)$ is an improper [1] random process, with second-order moments $\sigma_s^2 \triangleq \mathrm{E}[|s(n)|^2]$ and $\zeta_s(n) \triangleq \mathrm{E}[s^2(n)] \neq 0$, $\forall n \in \mathbb{Z}$, whose improper nature arises from the linear dependence existing between $s(n)$ and its conjugate version $s^*(n)$, i.e.,

$$s^*(n) = e^{j\,2\pi\beta n}\, s(n), \quad \forall n \in \mathbb{Z}; \tag{3.13}$$

A5) $c(n)$ is a *complex-valued* channel, that is, neither $c_R(n)$ nor $c_I(n)$ vanish identically.

A large number of digital modulation schemes satisfy assumption A4, including ASK, BPSK, DBPSK, offset QPSK (OQPSK), offset QAM (OQAM), MSK and its variant Gaussian MSK (GMSK) (see [32, 37] for a detailed discussion). Specifically, real modulation schemes, such as ASK and DBPSK, fulfill assumption A1 with $\beta = 0$, i.e., $s^*(n) = s(n)$, whereas for complex modulation formats, such as OQPSK, OQAM, and MSK-type, it results that $\beta = 1/2$, i.e., $s^*(n) = (-1)^n\, s(n)$. Finally, assumption A5 is customary when

one resorts to the equivalent lowpass representations of passband signals and systems.

3.3.1 Analysis of the L-CM cost function

In the absence of noise, accounting for (3.3) and invoking assumptions A1, A3, A4 and A5, after tedious but straightforward algebraic manipulations (see also [5]), it can be shown that minimization of (3.12) with respect to $\mathbf{f}$ is equivalent to the minimization with respect to the *combined channel-equalizer* vector $\mathbf{q} \stackrel{\triangle}{=} \mathbf{C}^H\mathbf{f} = [q_0, q_1, \ldots, q_{K-1}]^T \in \mathbb{C}^K$, of the cost function

$$\begin{aligned}\widetilde{J}_{\text{cm}}(\mathbf{q}) &\stackrel{\triangle}{=} J_{\text{cm}}(\mathbf{C}\,(\mathbf{C}^H\mathbf{C})^{-1}\mathbf{q}) \\ &= \kappa_s \sum_{\ell=0}^{K-1} |q_\ell|^4 + 2\left(\underbrace{\sigma_s^2\,\|\mathbf{q}\|^2}_{\text{E}[|y(k)|^2]}\right)^2 + \left|\underbrace{\sigma_s^2\,\mathbf{q}^H\mathbf{J}^*\mathbf{q}^*}_{\text{E}[y^2(k)]}\right|^2 - 2\,\gamma_s\,\underbrace{\sigma_s^2\,\|\mathbf{q}\|^2}_{\text{E}[|y(k)|^2]} + \gamma_s^2\end{aligned} \tag{3.14}$$

where $\kappa_s \stackrel{\triangle}{=} \text{E}[|s(n)|^4] - 2\,\text{E}^2[|s(n)|^2] - \left|\text{E}[s^2(n)]\right|^2 = \sigma_s^2\,(\gamma_s - 3\,\sigma_s^2) < 0$ is the kurtosis of $s(n)$, whereas $\mathbf{J} \stackrel{\triangle}{=} \text{diag}[1, e^{-j\,2\pi\beta}, \ldots, e^{-j\,2\pi\beta(K-1)}] \in \mathbb{C}^{K\times K}$ is a diagonal unitary matrix, i.e., $\mathbf{J}\,\mathbf{J}^* = \mathbf{J}^*\mathbf{J} = \mathbf{I}_K$. More precisely, if $\overline{\mathbf{q}} \in \mathbb{C}^K$ is a local minimum of $\widetilde{J}_{\text{cm}}(\mathbf{q})$, then, under assumption A3, $\overline{\mathbf{f}} = \mathbf{C}\,(\mathbf{C}^H\mathbf{C})^{-1}\overline{\mathbf{q}} + \mathbf{f}_{\mathcal{N}}$, where $\mathbf{f}_{\mathcal{N}} \in \mathbb{C}^{NL_e}$ is an arbitrary vector belonging to the null space of $\mathbf{C}^H$, is a local minimum of $J_{\text{cm}}(\mathbf{f})$. Furthermore, observe that $\sigma_s^2\,\|\mathbf{q}\|^2$ in (3.14) represents the mean-output-energy $\text{E}[|y(k)|^2]$ of the equalizer output $y(k) = \mathbf{q}^H\mathbf{s}(k)$, whereas $\sigma_s^2\,\mathbf{q}^H\mathbf{J}^*\mathbf{q}^*$ coincides with the second-order moment $\text{E}[y^2(k)]$. It is noteworthy that, compared with expressions of the CM cost functions commonly encountered in the literature, eq. (3.14) is more general. Specifically, when both the transmitted symbols and the CIR are real-valued (i.e., $\beta = 0$ and $\mathbf{C} \in \mathbb{R}^{(NL_e)\times K}$), the cost function (3.14) ends up to that studied in [34, 35, 36, 6]. In fact, in this case, the combined channel-equalizer vector $\mathbf{q}$ turns out to be real-valued, too, i.e., $\mathbf{q} \in \mathbb{R}^K$, and, consequently, the second and third summand in (3.14) can be grouped together. Additionally, the CM cost function (3.14) is different from that studied in [27, 28]. Indeed, in these papers, it is assumed that the transmitted symbols are proper complex: in this case, the kurtosis of $s(n)$ assumes the form $\kappa_s = \sigma_s^2\,(\gamma_s - 2\,\sigma_s^2)$ and, most important, the third summand in (3.14) disappears, i.e., $\text{E}[y^2(k)] = 0, \forall k \in \mathbb{Z}$. Henceforth, the basic difference between (3.14) and the expressions of the CM cost functions considered in [27, 28] and [34, 35, 36, 6] stems from the fact

that, under assumptions A4 and A5, the third summand in (3.14), which arises as a consequence of the improper nature of $s(n)$, is nonzero and different from the second one.

A vector $\overline{\mathbf{q}} \in \mathbb{C}^K$ is a stationary point of $\widetilde{J}_{\text{cm}}(\mathbf{q})$ if it is a solution of $\widetilde{\mathbf{g}}(\mathbf{q}) \stackrel{\triangle}{=} \nabla_{\mathbf{q}^*}[\widetilde{J}_{\text{cm}}(\mathbf{q})] = \mathbf{O}_K$, where $\nabla_{\mathbf{q}^*}[\widetilde{J}_{\text{cm}}(\mathbf{q})]$ denotes the complex gradient of $\widetilde{J}_{\text{cm}}(\mathbf{q})$ with respect to $\mathbf{q}^*$. Accounting for (3.14), one obtains

$$\begin{aligned}\widetilde{\mathbf{g}}(\mathbf{q}) &\stackrel{\triangle}{=} \nabla_{\mathbf{q}^*}[\widetilde{J}_{\text{cm}}(\mathbf{q})] \\ &= 2\,\kappa_s\,\widetilde{\boldsymbol{\Sigma}}(\mathbf{q})\,\mathbf{q} + 4\,\sigma_s^4\,\|\mathbf{q}\|^2\,\mathbf{q} + 2\,\sigma_s^4\,(\mathbf{q}^T\mathbf{J}\,\mathbf{q})\,\mathbf{J}^*\mathbf{q}^* - 2\,\gamma_s\,\sigma_s^2\,\mathbf{q} \quad (3.15)\end{aligned}$$

with $\widetilde{\boldsymbol{\Sigma}}(\mathbf{q}) \stackrel{\triangle}{=} \text{diag}[|q_0|^2, |q_1|^2], \ldots, |q_{K-1}|^2] \in \mathbb{R}^{K\times K}$. A stationary point $\overline{\mathbf{q}}$ is a local minimum of $\widetilde{J}_{\text{cm}}(\mathbf{q})$ if the Hessian matrix $\widetilde{\boldsymbol{\mathcal{H}}}(\mathbf{q}) \stackrel{\triangle}{=} \nabla_{\mathbf{q}}\left\{\nabla_{\mathbf{q}^*}[\widetilde{J}_{\text{cm}}(\mathbf{q})]\right\} \in \mathbb{C}^{K\times K}$ is positive definite for $\mathbf{q} = \overline{\mathbf{q}}$. Accounting for (3.15), one has

$$\begin{aligned}\widetilde{\boldsymbol{\mathcal{H}}}(\mathbf{q}) \stackrel{\triangle}{=} \nabla_{\mathbf{q}}\left\{\nabla_{\mathbf{q}^*}[\widetilde{J}_{\text{cm}}(\mathbf{q})]\right\} &= 4\,\kappa_s\,\widetilde{\boldsymbol{\Sigma}}(\mathbf{q}) + 4\,\sigma_s^4\,\|\mathbf{q}\|^2\,\mathbf{I}_K \\ &+ 4\,\sigma_s^4\,\mathbf{q}\,\mathbf{q}^H + 4\,\sigma_s^4\,\mathbf{J}^*\mathbf{q}^*\mathbf{q}^T\mathbf{J} - 2\,\gamma_s\,\sigma_s^2\,\mathbf{I}_K \quad (3.16)\end{aligned}$$

A useful property of $\widetilde{J}_{\text{cm}}(\mathbf{q})$ can be demonstrated. By virtue of (3.15), the cost function (3.14) can be rewritten as

$$\widetilde{J}_{\text{cm}}(\mathbf{q}) = \frac{1}{2}\,\mathbf{q}^H\widetilde{\mathbf{g}}(\mathbf{q}) - \gamma_s\,\sigma_s^2\,\|\mathbf{q}\|^2 + \gamma_s^2, \quad (3.17)$$

where the identity $\mathbf{q}^H\widetilde{\boldsymbol{\Sigma}}(\mathbf{q})\,\mathbf{q} = \sum_{\ell=0}^{K-1}|q_\ell|^4$ has been used. Thus, if $\overline{\mathbf{q}}$ is a stationary point of $\widetilde{J}_{\text{cm}}(\mathbf{q})$, i.e., $\widetilde{\mathbf{g}}(\overline{\mathbf{q}}) = \mathbf{O}_K$, one obtains

$$\widetilde{J}_{\text{cm}}(\overline{\mathbf{q}}) = \gamma_s\,(\gamma_s - \sigma_s^2\,\|\overline{\mathbf{q}}\|^2)\,, \quad (3.18)$$

which allows one to readily calculate the value of the CM cost function at any stationary point. As a by-product, since $\widetilde{J}_{\text{cm}}(\mathbf{q})$ is a nonnegative function, it follows from (3.18) that the mean-output-energy corresponding to any stationary point cannot be greater than the dispersion constant γ_s of the transmitted symbol sequence

$$\text{E}[|y(k)|^2] = \sigma_s^2\,\|\mathbf{q}\|^2 \leq \gamma_s \quad (3.19)$$

On the basis of (3.15) and (3.16), the following Theorem provides a family of local minima of the CM cost function (3.14).

Theorem 3.1 *The CM cost function (3.14) has local minima at the following vectors:*

$$\overline{\mathbf{q}}_{min,1} = e^{j\,\theta}\,\mathbf{e}_{i_1}, \quad \textit{when } \sigma_s^2 \leq \gamma_s < 2\,\sigma_s^2, \tag{3.20}$$

$$\overline{\mathbf{q}}_{min,2} = e^{j\,\theta}\sqrt{\frac{\gamma_s}{\gamma_s+\sigma_s^2}}\cdot\left[\mathbf{e}_{i_1} - j\,(-1)^{\ell_{i_1,i_2}}\,e^{j\,\pi\beta(i_2-i_1)}\,\mathbf{e}_{i_2}\right],$$

$$\textit{when } \sigma_s^2 < \gamma_s < 2\,\sigma_s^2, \tag{3.21}$$

$$\overline{\mathbf{q}}_{min,3} = e^{j\,\theta}\cdot\left[\rho\,\mathbf{e}_{i_1} - j\,(-1)^{\ell_{i_1,i_2}}\,e^{j\,\pi\beta(i_2-i_1)}\,\sqrt{1-\rho^2}\;\mathbf{e}_{i_2}\right],$$

$$\textit{when } \gamma_s = \sigma_s^2, \tag{3.22}$$

with $\mathbf{e}_{i_j} \stackrel{\triangle}{=} [\overbrace{0,\ldots,0}^{i_j},1,0,\ldots,0]^T \in \mathbb{R}^K$, $\theta \in [0,2\pi)$, $\ell_{i_1,i_2} \in \mathbb{Z}$, $i_1 \neq i_2 \in \{0,1,\ldots,K-1\}$ *and* $0<\rho<1$.

Proof: See Appendix A.1.

Some comments are now in order. First of all, observe that when $\mathbf{q} = \overline{\mathbf{q}}_{\text{min},1}$, the equalizer output is given by

$$y(k) = \overline{\mathbf{q}}_{\text{min},1}^H\,\mathbf{s}(k) = e^{j\,\theta}s(k-i_1) \tag{3.23}$$

i.e., except for an arbitrary phase rotation, perfect symbol recovery is guaranteed: in this case, the CM behaves as a blind ZF equalizer, which completely suppresses ISI. It is worth noting that, unlike the CM cost functions studied in [27], [28], [6, 34, 35, 36], the function (3.14) does not exhibit the ISI-free local minima (3.20) when the transmitted symbols $s(n)$ are "Gaussian" [5], i.e., $\kappa_s = 0 \Leftrightarrow \gamma_s = 3\,\sigma_s^2$, or "super-Gaussian" [5], i.e., $\kappa_s > 0 \Leftrightarrow \gamma_s > 3\,\sigma_s^2$, as well as when $s(n)$ is sub-Gaussian with $-\sigma_s^4 \leq \kappa_s < 0 \Leftrightarrow 2\,\sigma_s^2 \leq \gamma_s < 3\,\sigma_s^2$. This is the reason we have assumed in A1 that the transmitted symbols are sub-Gaussian. Additionally, observe that, accounting for (3.18), it follows that $\widetilde{J}_{\text{cm}}(\overline{\mathbf{q}}_{\text{min},1}) = \gamma_s\,(\gamma_s - \sigma_s^2)$. In contrast, when $\mathbf{q} = \overline{\mathbf{q}}_{\text{min},2}$ or $\mathbf{q} = \overline{\mathbf{q}}_{\text{min},3}$, the equalizer output is contaminated by ISI, since a particular linear combination of the two transmitted symbols $s(k-i_1)$ and $s(k-i_2)$, with $i_1 \neq i_2 \in \{0,1,\ldots,K-1\}$, is extracted in these cases.

Henceforth, different from [27], [28], [34]-[6], the cost function (3.14) exhibits local minima that do not lead to perfect ISI suppression, even in the absence of noise.

In particular, it is worth noting that, when $-2\,\sigma_s^4 < \kappa_s < -\sigma_s^4 \Leftrightarrow \sigma_s^2 < \gamma_s < 2\,\sigma_s^2$, the CM cost function exhibits the undesired local minima (3.21). In this case, relying on (3.18), it results that

$$\widetilde{J}_{\text{cm}}(\overline{\mathbf{q}}_{\text{min},2}) = \widetilde{J}_{\text{cm}}(\overline{\mathbf{q}}_{\text{min},1}) \cdot [\gamma_s/(\gamma_s + \sigma_s^2)] < \widetilde{J}_{\text{cm}}(\overline{\mathbf{q}}_{\text{min},1}), \quad \text{for } \sigma_s^2 < \gamma_s < 2\,\sigma_s^2 \tag{3.24}$$

which shows that, surprisingly enough, the ISI-free local minima (3.20) are not global. On the other hand, when $\kappa_s = -2\,\sigma_s^4 \Leftrightarrow \gamma_s = \sigma_s^2$, a situation occurring when $s(n)$ is constant modulus, accounting for (3.18), the value of $\widetilde{J}_{\text{cm}}(\mathbf{q})$ at the undesired local minima (**??**) is given by

$$\widetilde{J}_{\text{cm}}(\overline{\mathbf{q}}_{\text{min},3}) = 0, \quad \text{for } \gamma_s = \sigma_s^2. \tag{3.25}$$

In this case, $\widetilde{J}_{\text{cm}}(\overline{\mathbf{q}}_{\text{min},1})$ turns out to be zero as well and, hence, both the desired (3.20) and undesired (3.22) local minima are global. It should be observed that, by setting $\beta = 0$, $\ell_{i_1,i_2} = 1$ and $\rho = \cos(\phi)$, with $\phi \in [0, 2\,\pi)$, the expression of the undesired local minima (3.22) ends up to that derived in [29, eq. (15)] for the case of an infinite length CM equalizer, under the simplifying assumption of BPSK symbols with unitary variance. Finally, it is noteworthy that, with reference to real-valued symbols (i.e., $\beta = 0$), it was shown in [31] that the undesired minima (3.21) and (3.22) disappear by minimizing the CM cost function (3.12), provided that the equalizer output $y(k)$ is a *not* strictly linear function of $\mathbf{z}(k)$, namely, $y(k) = \text{Re}\{\mathbf{f}^H\,\mathbf{z}(k)\}$. More generally, if the transmitted symbols are improper complex, one has to use widely-linear equalizing structures , whereby the equalizer output is given by $y(k) = \mathbf{f}^H\mathbf{z}(k) + \mathbf{g}^H\mathbf{z}^*(k)$ and the CM cost function is minimized with respect to both $\mathbf{f}$ and $\mathbf{g}$, where $\mathbf{g} \in \mathbb{C}^{NL_e}$ is not necessarily constrained to be equal to $\mathbf{f}^*$. This issue is the topic of the next section.

3.3.2 Numerical Results

To corroborate our analysis, the results of a Monte Carlo computer simulation are now presented. We consider both QPSK and OQPSK modulations, with $s(n)$ taking equiprobable values in $\{\pm 1, \pm j\}$, and a noise vector $\mathbf{v}(k)$ in (3.3) modeled as a zero-mean complex proper white random process, with autocorrelation matrix $\mathbf{R}_{\mathbf{vv}} \triangleq \text{E}[\mathbf{v}(k)\,\mathbf{v}^H(k)] = \sigma_v^2\,\mathbf{I}_{NL_e}$. The signal-to-noise ratio (SNR) at the equalizer input is defined as SNR $\triangleq [\sigma_s^2/(N\,\sigma_v^2)] \cdot \sum_{q=0}^{L_c-1} \| \mathbf{c}(q) \|^2$ and is set to 25 dB. The received signal $r_a(t)$

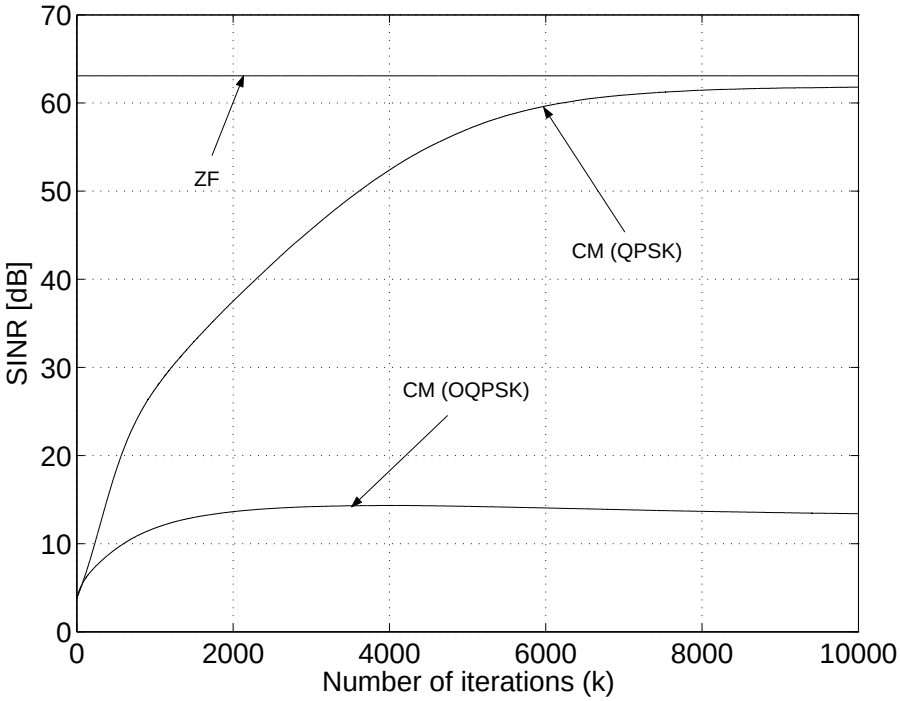

Figure 3.1: SINR versus number of iterations.

is fractionally sampled at rate $2/T_s$ and the $\mathcal{Z}$-transforms of the two 3rd-order polyphase components $\widetilde{c}^{(\ell)}(q)$ are given by $\widetilde{C}^{(\ell)}(z) = (1-0.5\,e^{j\,\theta_{1,\ell}}z^{-1})\,(1-1.2\,e^{j\,\theta_{2,\ell}}z^{-1})$, for $\ell \in \{0,1\}$, where $\theta_{1,0} = 0.7\,\pi$, $\theta_{2,0} = \theta_{1,0} + \pi$, $\theta_{1,1} = \theta_{1,0} + 0.2\,\pi$ and $\theta_{2,1} = \theta_{2,0} + 0.2\,\pi$. The minimization of the CM cost function (3.12) is adaptively carried out by resorting to the stochastic gradient descent algorithm [5], where $\gamma_s = \sigma_s^2 = 1$, $L_e = 5$, double-spike initialization is used and the step-size is continuously adjusted to achieve fast convergence without compromising stability. For each of the 10^3 Monte Carlo trials carried out, both the symbol and noise sequences are randomly and independently generated. Fig. 5.1 reports the signal-to-interference-plus-noise ratio (SINR) at the output of the CM equalizer as a function of the number of iterations, when either QPSK or OQPSK modulations are employed at the transmitter; for the sake of comparison, it is also reported the SINR (which is the same for QPSK and OQPSK modulations) at the output of the minimal-norm L-FIR ZF equalizer (synthesized by assuming perfect knowledge of $\mathbf{C}$). Results show

that, when $s(n)$ is a proper random sequence (QPSK) and, thus, all the local minima of $J_{\text{cm}}(\mathbf{f})$ are desired, the performance of the CM equalizer rapidly improves as the number of iterations increases and becomes comparable to that of the ZF equalizer. In contrast, when the transmitted symbols are improper (OQPSK), due to the presence of the undesired global minima (3.22), the curve of the CM equalizer quickly saturates to a value that is significantly less (of about 50 dB) than the output SINR of the ZF equalizer.

3.4 Widely Linear Constant Modulus Equalizer

In this section, we present some results reported in [8, 9]. Specifically, we provide a general and unified framework to design WL equalizers for both real- and complex-valued improper modulations, by deriving the conditions assuring perfect symbol recovery in the absence of noise and providing some insights into the synthesis and analysis of blind WL-CM equalizers (it can be noted that the theoretical analysis that we have presented in [9] is herein omitted).

A brief characterization of the the second-order statistical properties of $\mathbf{z}(k)$ [see (3.3)] is now in order. Preliminarily, we observe that as a consequence of equation (3.13), the vector $\mathbf{s}(k)^*$ can be expressed as

$$\mathbf{s}^*(k) = e^{j\,2\pi\beta k}\,\mathbf{J}\,\mathbf{s}(k), \tag{3.26}$$

where $\mathbf{J} = \text{diag}[1, e^{-j\,2\pi\beta}, \ldots, e^{-j\,2\pi\beta(K-1)}]$ is the diagonal unitary matrix defined in equation (3.14). Therefore, accounting for the equation (3.3) and assumptions A4, A2, the second-order moment statistics of $\mathbf{z}(k)$ are given by both the autocorrelation matrix

$$\mathbf{R}_{\mathbf{zz}} \triangleq \text{E}[\mathbf{z}(k)\,\mathbf{z}^H(k)] = \sigma_s^2\,\mathbf{C}\,\mathbf{C}^H + \sigma_w^2\,\mathbf{I}_{NL_e}, \tag{3.27}$$

and the *conjugate* correlation matrix

$$\mathbf{R}_{\mathbf{zz}^*}(k) \triangleq \text{E}[\mathbf{z}(k)\,\mathbf{z}^T(k)] = \sigma_s^2\,e^{-j\,2\pi\beta k}\,\mathbf{C}\,\mathbf{J}^*\mathbf{C}^T. \tag{3.28}$$

Since $\mathbf{R}_{\mathbf{zz}^*}(k)$ is nonvanishing, $\forall k \in \mathbb{Z}$, the vector $\mathbf{z}(k)$ is improper [1]. Additionally, observe that, for real modulation schemes (for which $\beta = 0$), such as ASK and DBPSK, the vector $\mathbf{z}(k)$ is wide-sense stationary (WSS), whereas for complex modulation formats (for which $\beta = 1/2$), such as OQPSK, OQAM, and MSK-type, it results that $\mathbf{z}(k)$ is wide-sense conjugate (second-order) cyclostationary [38] with period 2.

Since $\mathbf{z}(k)$ is an improper vector, it is well-known (see section **??** for details) that, compared with L-FIR processing, a WL-FIR estimator, which is linear both in $\mathbf{z}(k)$ and $\mathbf{z}^*(k)$, can assure a better estimate of the symbol $s(k-d)$, with $d \in \{0, 1, \ldots, K-1\}$ (a suitable equalization delay). The weight vector of the resulting WL-FIR estimator depends on both $\mathbf{R}_{\mathbf{zz}}$ and $\mathbf{R}_{\mathbf{zz}^*}(k)$. Therefore, in this section to account for the (possible) time-varying feature of $\mathbf{R}_{\mathbf{zz}^*}(k)$, we consider a slight modification of the classical WL-FIR. Specifically, we preliminarily derive the forms of the WL-MMSE and WL-ZF equalizers, by gaining some new insights. Successively, using some of these results, we analyze the WL-CM cost function and compare it with the L-CM one.

3.4.1 WL-MMSE and WL-ZF Equalizers

Preliminarily, we observe that from (3.26) it follows, accounting (3.3), that

$$\mathbf{z}^*(k) = e^{j\,2\pi\beta k}\,\mathbf{C}^*\,\mathbf{J}\,\mathbf{s}(k) + \mathbf{v}^*(k). \tag{3.29}$$

Thus, the (possible) wide-sense conjugate cyclostationarity of $\mathbf{z}(k)$ can be compensated by performing a derotation of $\mathbf{z}^*(k)$ before constructing the WL-FIR estimator, that is,

$$\begin{aligned} y(k) &= \mathbf{f}_1^H\mathbf{z}(k) + \mathbf{f}_2^H\mathbf{z}^*(k)\,e^{-j\,2\pi\beta k} \\ &= \underbrace{\begin{bmatrix}\mathbf{f}_1^H & \mathbf{f}_2^H\end{bmatrix}}_{\widetilde{\mathbf{f}}^H \in \mathbb{C}^{1\times 2NL_e}} \underbrace{\begin{bmatrix}\mathbf{z}(k) \\ \mathbf{z}^*(k)\,e^{-j\,2\pi\beta k}\end{bmatrix}}_{\widetilde{\mathbf{z}}(k)\in\mathbb{C}^{2NL_e}} = \widetilde{\mathbf{f}}^H\widetilde{\mathbf{z}}(k)\,, \end{aligned} \tag{3.30}$$

where

$$\widetilde{\mathbf{z}}(k) = \underbrace{\begin{bmatrix}\mathbf{C} \\ \mathbf{C}^*\mathbf{J}\end{bmatrix}}_{\widetilde{\mathbf{C}}\in\mathbb{C}^{2NL_e\times K}}\mathbf{s}(k) + \underbrace{\begin{bmatrix}\mathbf{v}(k) \\ \mathbf{v}^*(k)\,e^{-j\,2\pi\beta k}\end{bmatrix}}_{\widetilde{\mathbf{v}}(k)\in\mathbb{C}^{2NL_e}} = \widetilde{\mathbf{C}}\,\mathbf{s}(k) + \widetilde{\mathbf{v}}(k)\,. \tag{3.31}$$

It is worth noting that the conventional linear estimator $y(k) = \mathbf{f}^H\mathbf{z}(k)$ can be obtained from (3.30) by setting $\mathbf{f}_1 = \mathbf{f} \in \mathbb{C}^{NL_e}$ and $\mathbf{f}_2 = \mathbf{0}_{NL_e}$. Let $J_{\text{mse}}(\widetilde{\mathbf{f}}) \triangleq \mathrm{E}[|y(k) - s(k-d)|^2]$, with $d \in \{0, 1, \ldots, K-1\}$ a suitable equalization delay, denote the mean-square error between the equalizer output and the desired symbol $s(k-d)$, according to the MMSE criterion, the weight vector $\widetilde{\mathbf{f}}$ is chosen as follows

$$\widetilde{\mathbf{f}}_{\text{wl-mmse}} \triangleq \begin{bmatrix}\mathbf{f}_{1,\text{wl-mmse}} \\ \mathbf{f}_{2,\text{wl-mmse}}\end{bmatrix} = \arg\min_{\widetilde{\mathbf{f}}\in\mathbb{C}^{2NL_e}} J_{\text{mse}}(\widetilde{\mathbf{f}}) = \sigma_s^2\,\widetilde{\mathbf{R}}_{\mathbf{zz}}^{-1}\,\widetilde{\mathbf{C}}\,\mathbf{e}_d \tag{3.32}$$

where $\mathbf{e}_d \triangleq [\overbrace{0,\ldots,0}^{d},1,0,\ldots,0]^T \in \mathbb{R}^K$ and the autocorrelation matrix of the augmented vector $\widetilde{\mathbf{z}}(k)$ is given by

$$\widetilde{\mathbf{R}}_{\mathbf{zz}} \triangleq \mathrm{E}[\widetilde{\mathbf{z}}(k)\,\widetilde{\mathbf{z}}^H(k)] = \sigma_s^2\,\widetilde{\mathbf{C}}\,\widetilde{\mathbf{C}}^H + \sigma_w^2\,\mathbf{I}_{2NL_e}. \tag{3.33}$$

Moreover, by partitioning $\widetilde{\mathbf{R}}_{\mathbf{zz}}$ and $\widetilde{\mathbf{C}}$ according to the structure of $\widetilde{\mathbf{z}}(k)$, resorting to the inverse of a partitioned matrix and accounting for the expression of $\mathbf{R}_{\mathbf{zz}^*}(k)$, one has

$$\begin{aligned}\mathbf{f}_{1,\text{wl-mmse}} = {} & \sigma_s^2\,\left[\mathbf{R}_{\mathbf{zz}} - \sigma_s^4\,\mathbf{C}\,\mathbf{J}^*\mathbf{C}^T(\mathbf{R}^*_{\mathbf{zz}})^{-1}\mathbf{C}^*\mathbf{J}\,\mathbf{C}^H\right]^{-1} \\ & \left[\mathbf{C} - \sigma_s^2\,\mathbf{C}\,\mathbf{J}^*\mathbf{C}^T(\mathbf{R}^*_{\mathbf{zz}})^{-1}\mathbf{C}^*\,\mathbf{J}\right]\,\mathbf{e}_d\,, \end{aligned} \tag{3.34}$$

$$\mathbf{f}_{2,\text{wl-mmse}} = e^{-j\,2\pi\beta d}\,\mathbf{f}^*_{1,\text{wl-mmse}}\,. \tag{3.35}$$

As it is apparent from (3.35), a particular linear dependence must exist between $\mathbf{f}_{2,\text{wl-mmse}}$ and $\mathbf{f}^*_{1,\text{wl-mmse}}$. As a side remark, observe that the WL-MMSE equalizer given by (3.32) generalizes and subsumes as a particular case the WL-MMSE equalizer derived in [32]: more precisely, when real modulation schemes, such as ASK and DBPSK, are employed at the transmitter and $N = 1$, i.e., the received signal $r_a(t)$ is sampled at the baud rate, the devised WL-MMSE equalizer (3.32) boils down to that proposed in [32].

The performance of the WL-MMSE equalizer (3.32) strongly depends on the existence of WL-ZF solutions, in the absence of noise. This important issue is investigated now.

Following the same lines of the linear case (see section 3.3 for details) it can be shown from (3.30), that in the absence of noise, imposing the ZF condition $y(k) = s(k-d)$ leads to the system of linear equations $\widetilde{\mathbf{f}}^H\widetilde{\mathbf{C}} = \mathbf{e}_d^T \Leftrightarrow \widetilde{\mathbf{C}}^H\widetilde{\mathbf{f}} = \mathbf{e}_d$, which is consistent if and only if (iff) $\widetilde{\mathbf{C}}^H(\widetilde{\mathbf{C}}^H)^-\mathbf{e}_d = \mathbf{e}_d$ (see [24]). If the augmented channel matrix $\widetilde{\mathbf{C}}$ is full-column rank, i.e., $\mathrm{rank}(\widetilde{\mathbf{C}}) = K$, it results that $\widetilde{\mathbf{C}}^H(\widetilde{\mathbf{C}}^H)^- = \mathbf{I}_K$ and, then, this system turns out to be consistent regardless of the equalization delay d. In this case, the *minimal norm* solution, i.e., the solution of the constrained optimization problem

$$\widetilde{\mathbf{f}}_{\text{wl-zf}} = \arg\min_{\widetilde{\mathbf{f}}\in\mathbb{C}^{2NL_e}} \|\widetilde{\mathbf{f}}\|^2\,, \quad \text{subject to } \widetilde{\mathbf{C}}^H\,\widetilde{\mathbf{f}} = \mathbf{e}_d\,, \tag{3.36}$$

is given by (see, e.g., [24])

$$\widetilde{\mathbf{f}}_{\text{wl-zf}} \triangleq \begin{bmatrix}\mathbf{f}_{1,\text{wl-zf}} \\ \mathbf{f}_{2,\text{wl-zf}}\end{bmatrix} = (\widetilde{\mathbf{C}}^H)^\dagger\,\mathbf{e}_d = \widetilde{\mathbf{C}}\,(\widetilde{\mathbf{C}}^H\widetilde{\mathbf{C}})^{-1}\mathbf{e}_d\,. \tag{3.37}$$

It is worth noting that, accounting for (3.32) and for the limit formula for the Moore-Penrose inverse [24], it can be verified that $\lim_{\sigma_w^2/\sigma_s^2 \to 0} \widetilde{\mathbf{f}}_{\text{wl-mmse}} = (\widetilde{\mathbf{C}}^H)^\dagger \, \mathbf{e}_d = \widetilde{\mathbf{f}}_{\text{wl-zf}}$, that is, as the noise variance σ_w^2 vanishes, the WL-MMSE solution approaches to the ZF one. Henceforth, we can maintain that, similarly to (3.35), resorting to the expression of $\widetilde{\mathbf{C}}$ one has

$$\mathbf{f}_{1,\text{wl-zf}} = \mathbf{C} \, (\mathbf{C}^H \, \mathbf{C} \, + \mathbf{J}^* \mathbf{C}^T \mathbf{C}^* \mathbf{J})^{-1} \, \mathbf{e}_d \, , \tag{3.38}$$

$$\mathbf{f}_{2,\text{wl-zf}} = \mathbf{C}^* \, \mathbf{J} \, (\mathbf{C}^H \, \mathbf{C} \, + \mathbf{J}^* \mathbf{C}^T \mathbf{C}^* \mathbf{J})^{-1} \, \mathbf{e}_d, \tag{3.39}$$

from which we desume that the following relation holds

$$\mathbf{f}_{2,\text{wl-zf}} = e^{-j \, 2\pi\beta d} \, \mathbf{f}_{1,\text{wl-zf}}^* \tag{3.40}$$

between the subvectors $\mathbf{f}_{1,\text{wl-zf}}$ and $\mathbf{f}_{2,\text{wl-zf}}^*$ in (3.37). The following Theorem (whose proof is omitted), provides the mathematical conditions assuring the existence of WL-ZF solutions, i.e., conditions assuring that the augmented channel matrix $\widetilde{\mathbf{C}}$ is full-column rank.

Theorem 3.2 *Let $C^{(\ell)}(z)$ denote the $\mathcal{Z}$-transform of the ℓth channel phase $\{c^{(\ell)}(k)\}_{k=0}^{L_c-1}$, for $\ell \in \{0, 1, \ldots, N-1\}$, and assume that at least one polynomial $\{C^{(\ell)}(z)\}_{\ell=0}^{N-1}$ is of maximum order $L_c - 1$. Then, matrix $\widetilde{\mathbf{C}}$ is full-column rank if the following conditions hold:*

C1) $2\,N\,L_e \geq K = L_e + L_c - 1$*;*

C2) the $2\,N$ polynomials $C^{(\ell)}(z)$ and $C^{(\ell)}(z^* \, e^{-j\,2\pi\beta})$, for $\ell \in \{0, 1, \ldots, N-1\}$, are coprime.

Some interesting remark are now in order. First, as regards condition C1, observe that, unlike L-FIR-ZF equalization, WL-FIR-ZF solutions might exist not only when fractionally sampling is performed at the receiver, but also when the received signal $r_a(t)$ is sampled at the baud rate, i.e., $N = 1$; in this case, condition C1 requires that $L_e \geq L_c - 1$ and condition C2 is fulfilled if, $\forall q_1, q_2 \in \{1, \ldots, L_c - 1\}$, there is no pair $(\zeta_{q_1}, \zeta_{q_2})$ of zeros of the $\mathcal{Z}$-transform of $c(n) = c_a(nT_s)$ such that $\zeta_{q_1} = \zeta_{q_2}^* \, e^{-j\,2\pi\beta}$. Second, and most important, note that, in comparison with L-FIR-ZF fractionally spaced equalization, condition C2 imposes a milder constraint on the channel phases $\{c^{(\ell)}(k)\}_{\ell=0}^{N-1}$. Indeed, when $N > 1$, L-FIR-ZF solutions exist if the N polynomials $\{C^{(\ell)}(z)\}_{\ell=0}^{N-1}$ are coprime (see section 3.3 for details). In contrast, Theorem 3.2 states that WL-FIR-ZF solutions exist even when

$\{C^{(\ell)}(z)\}_{\ell=0}^{N-1}$ have a common zero z_0, i.e., $C^{(0)}(z_0) = C^{(1)}(z_0) = \ldots = C^{(N-1)}(z_0) = 0$, provided that, the complex number z_0 is not a common zero of $C^{(\ell)}(z^* \, e^{-j\,2\pi\beta})$, $\forall \ell \in \{0, 1, \ldots, N-1\}$, that is, there exists at least one index $\ell_0 \in \{0, 1, \ldots, N-1\}$ such that $C^{(\ell_0)}(z_0^* \, e^{-j\,2\pi\beta}) \neq 0$. As we will see in Section 3.4.2, conditions C1 and C2 also play a fundamental role for the synthesis of CM-based WL equalizers and, thus, we assume hereinafter that both of them are fulfilled.

As it is apparent from (3.32) and (3.37), the synthesis of both WL-MMSE and WL-ZF equalizers requires the explicit knowledge or estimation of the channel vectors $\{\mathbf{c}(k)\}_{k=0}^{L_c-1}$, which are *unknown* at the receiver. To design a blind ISI-resilient receiver for improper modulation formats, without requiring any training sequence, we resort in the next section to the CM criterion.

3.4.2 Analysis of WL-CM cost function

With reference to the WL-FIR estimator given by (3.30), one might attempt to blindly choose the augmented weight vector $\widetilde{\mathbf{f}}$ by minimizing the *unconstrained* CM cost function

$$J_{\text{wl-cm}}(\widetilde{\mathbf{f}}) \stackrel{\triangle}{=} \mathrm{E}[(\gamma_s - |y(k)|^2)^2] \,, \tag{3.41}$$

where $\gamma_s \stackrel{\triangle}{=} \mathrm{E}[|s(k)|^4]/\sigma_s^2$ is again the dispersion constant.

Note that the classical L-FS-CM cost function $J_{\text{cm}}(\mathbf{f})$ (3.12) can be obtained from (3.41) by setting in (3.30) $\mathbf{f}_1 = \mathbf{f} \in \mathbb{C}^{NL_e}$ and $\mathbf{f}_2 = \mathbf{0}_{NL_e}$. In the section 3.3 we have shown that when noise is absent, the channel impulse response is complex-valued (see A5), and the transmitted sub-Gaussian symbols fulfill assumption A4, besides containing desired local minima, the function $J_{\text{cm}}(\mathbf{f})$ also exhibits *undesired* global minima, namely, they do not lead to perfect source recovery.

On the basis of widely-linear filtering theory [2, 3, 1], it can be argued that the presence of undesired global minima for the L-FS-CM cost function (3.12) is a consequence of the fact that, when the transmitted symbol sequence is improper, a linear estimator cannot take advantage of the additional information available in the conjugate correlation matrix of $\mathbf{z}(k)$. Consequently, it should be concluded that the minimization of the WL-FS-CM cost function (3.41) might lead to a blind receiver whose ISI suppression capabilities are close to those of the WL-FS-MMSE equalizer given by (3.32) [or, in the absence of noise, to those of the WL-FS-ZF equalizer given by (3.37)]. Interestingly enough, as it is confirmed by the simulation results reported in Section 5.5, this

conclusion is not entirely true. Indeed, similarly to $J_{\text{cm}}(\mathbf{f})$, the cost function $J_{\text{wl-cm}}(\widetilde{\mathbf{f}})$ exhibits undesired global minima, whose presence is basically due to the fact that the vector $\widetilde{\mathbf{f}}_{\text{wl-cm}}$ corresponding to a local minimum of $J_{\text{wl-cm}}(\widetilde{\mathbf{f}})$ might not exhibit the conjugate symmetry property (3.35) and (3.40), which characterizes instead the WL-MMSE and WL-ZF equalizers. To overcome this drawback, we propose to resort to the following constrained minimization of $J_{\text{wl-cm}}(\widetilde{\mathbf{f}})$, by imposing in (3.30) that

$$\mathbf{f}_2 = e^{-j\,2\pi\beta d}\,\mathbf{f}_1^*, \tag{3.42}$$

i.e., we consider the following optimization problem

$$\widetilde{\mathbf{f}}_{\text{wl-ccm}} = \arg \min_{\widetilde{\mathbf{f}} \in \mathbb{C}^{2NL_e}} J_{\text{wl-cm}}(\widetilde{\mathbf{f}})\,, \quad \text{subject to } \mathbf{f}_2 = e^{-j\,2\pi\beta d}\,\mathbf{f}_1^*\,, \tag{3.43}$$

which will be referred to as the WL-FS constrained CM (WL-FS-CCM) equalizer. Since CM equalizers do not have closed-form solutions, minimization of (3.43) is adaptively carried out by resorting to the stochastic gradient descent (SGD) algorithm. Specifically, let $\widetilde{\mathbf{f}}_{\text{wl-ccm}}(k) \stackrel{\triangle}{=} [\mathbf{f}_{1,\text{wl-ccm}}^T(k), \mathbf{f}_{2,\text{wl-ccm}}^T(k)]^T \in \mathbb{C}^{2NL_e}$, with $\mathbf{f}_{2,\text{wl-ccm}}(k) = e^{-j\,2\pi\beta d}\ \mathbf{f}_{1,\text{wl-ccm}}^*(k) \in \mathbb{C}^{NL_e}$, denote the estimate of $\widetilde{\mathbf{f}}_{\text{wl-ccm}}$ at iteration k, starting from (3.43), one obtains the updating equation

$$\begin{aligned}\mathbf{f}_{1,\text{wl-ccm}}(k+1) = \mathbf{f}_{1,\text{wl-ccm}}(k) + \mu\, y_{\text{wl-ccm}}^*(k) \\ \cdot (\gamma_s - |y_{\text{wl-ccm}}(k)|^2)\,\mathbf{z}(k)\,,\end{aligned} \tag{3.44}$$

where

$$y_{\text{wl-ccm}}(k) \;=\; \mathbf{f}_{1,\text{wl-ccm}}^H(k)\,\mathbf{z}(k) \;+\; \mathbf{f}_{1,\text{wl-ccm}}^T(k)\,\mathbf{z}^*(k)\,e^{-j\,2\pi\beta(k-d)} \tag{3.45}$$

and $\mu > 0$ denotes the step-size of the algorithm. It should be observed that, when real modulation schemes, such as ASK and DBPSK, are employed at the transmitter and $N = 1$, i.e., the received signal $r_a(t)$ is sampled at the baud rate, the proposed WL-FS-CCM equalizer (3.43) boils down to the single-axis equalizer devised in [31]. The fact that the single-axis equalizer is actually a WL equalizer was not recognized in [31]. The performances of the WL-FS-CCM equalizer are studied in subection 5.5 through computer simulations.

3.4.3 Simulation results

In this section, we investigate the performances of both WL-BS (i.e., $N = 1$) and $T_s/2$-spaced WL-FS equalizers (i.e., $N = 2$). Specifically, we considered the following equalizers: WL-BS-MMSE, WL-BS-CM, WL-BS-CCM,

WL-FS-MMSE, WL-FS-CM, WL-FS-CCM. For the sake of comparison, we also considered the L-FS-MMSE and L-FS-CM equalizers[1]. All the MMSE equalizers are *non-blind* and are implemented in batch-mode, by assuming perfect knowledge of the channel impulse response and by inverting the appropriate sample correlation matrix, estimated over K symbol intervals; additionally, for each MMSE equalizer, we chose the value of the equalization delay $d \in \{0, 1, \ldots, K-1\}$ assuring the best performance. On the other hand, all the CM *blind* equalizers are adaptively implemented by resorting to the SGD algorithm [5], wherein the step-size is continuously adjusted to achieve fast convergence without compromising stability. More specifically, we set $\mu(k) = 0.01\,\mu_{\max}(k)$, where, according to [39], $\mu_{\max}(k)$ is the maximum value of the step-size that assures SGD stability at iteration k, and can be evaluated in real-time, since it depends only on the equalizer output $y(k)$ and γ_s; moreover, we employed single- and double-spike initialization [5] for baud- and fractionally-spaced CM equalizers, respectively. All the equalizers under comparison jointly elaborate $L_e = 5$ consecutive symbols.

The input stream $s(n)$ is drawn from an OQPSK constellation and the additive noise $w(n)$ is a complex proper Gaussian process. The signal-to-noise ratio (SNR) at the equalizer input is defined as SNR $\stackrel{\triangle}{=} (\sigma_s^2/\sigma_w^2)\|\mathbf{c}\|^2$ and both the symbol and noise sequences are randomly and independently generated at the start of each Monte Carlo run. Since BS and FS equalizers employ different discrete-time channels, we considered for all the receivers the same continuous-time channel $c_a(t)$, which spans $L_c = 3$ symbol periods; more precisely, we started from the $T_s/2$-sampled version of $c_a(t)$, i.e., $c(n) \stackrel{\triangle}{=} c_a(n\,T_s/2)$, for $n \in \{0, 1, \ldots, 2\,L_c - 1\}$, which can be expressed in terms of the two polyphase components $c^{(0)}(k) \stackrel{\triangle}{=} c(2k)$ and $c^{(1)}(k) \stackrel{\triangle}{=} \widetilde{c}(2k+1)$, for $k \in \{0, 1, \ldots, L_c - 1\}$. Thus, we obtain the unique symbol-spaced channel for BS methods as $c(n) = \widetilde{c}^{(0)}(n)$, $n \in \{0, 1, \ldots, L_c - 1\}$. The two channels $\widetilde{c}^{(\ell)}(n)$, for $\ell = 0, 1$, are assigned in terms of their $\mathcal{Z}$-transforms:

$$\widetilde{C}^{(\ell)}(z) = (1 - 0.5\,e^{j\,\theta_{1,\ell}}\,z^{-1})\,(1 - 1.2\,e^{j\,\theta_{2,\ell}}\,z^{-1})\,, \qquad (3.46)$$

where $\theta_{1,0} = 0.5\,\pi + \gamma$, $\theta_{2,0} = \theta_{1,0} + \pi$, $\theta_{1,1} = \theta_{1,0} + \gamma$ and $\theta_{2,1} = \theta_{2,0} + \gamma$, and the angular separation γ is fixed to $0.2\,\pi$ so as to assure the existence of ZF solutions for all the methods under comparison. As performance measure, we evaluated the average bit-error-rate (ABER) and, denoting with q_ℓ, for $\ell \in$

[1] Linear baud-spaced equalizers were not considered since, at symbol spacing T_s, L-FIR-ZF solutions do not exist as we have noted in the section 3.3.

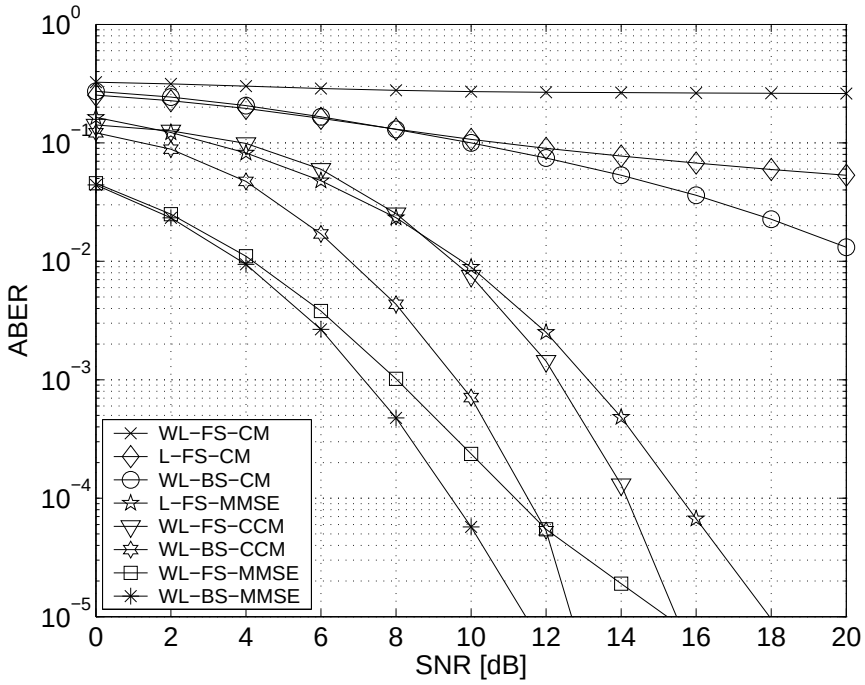

Figure 3.2: ABER versus SNR.

$\{0, 1, \ldots, K-1\}$, the ℓth entry of the combined channel-equalizer impulse response $\widetilde{\mathbf{q}} \stackrel{\triangle}{=} \widetilde{\mathbf{C}}^H \widetilde{\mathbf{f}} \in \mathbb{C}^K$, we also resorted to the residual ISI expressed in dB

$$\text{ISI [dB]} \stackrel{\triangle}{=} 10 \log_{10} \left(\frac{\sum_{\ell=0}^{K-1} |\widetilde{q}_\ell|^2 - \max_\ell |\widetilde{q}_\ell|^2}{\max_\ell |\widetilde{q}_\ell|^2} \right) . \tag{3.47}$$

Note that (3.47) only quantifies the ISI suppression capability of the equalizer and does not take into account noise enhancement at its output. For each of the 10^4 Monte Carlo trials carried out, after estimating the receiver weights on the basis of the given data record of length K, an independent record of 1000 symbols was considered to evaluate the ABER.

In the first experiment, we evaluated the ABER performances of the considered equalizers as a function of the SNR, with $K = 500$ symbols. Results of Fig. 3.2 show that the performances of the L-FS-CM, WL-BS-CM and WL-FS-CM blind equalizers are significantly worse than those of the corresponding non-blind MMSE equalizers. In particular, it is worth noting that both the

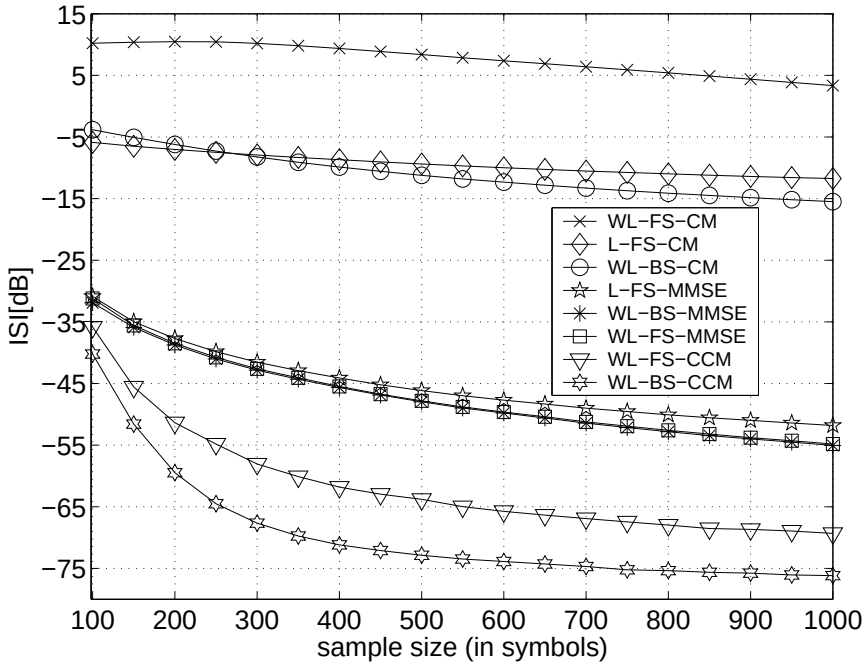

Figure 3.3: ISI versus sample size K (in symbols).

WL-FS-MMSE and WL-BS-MMSE equalizers remarkably outperform the L-FS-MMSE equalizer for all the considered values of the SNR. On the other hand, as it has been previously claimed, the proposed WL-FS-CCM and WL-BS-CCM blind equalizers perform better than their unconstrained WL-BS-CM and WL-FS-CM counterparts, for all the considered values of the SNR. Interestingly, the WL-FS-CCM and WL-BS-CCM equalizers also outperform the L-FS-MMSE one and, as the SNR increases, their ABER curves approach those of their corresponding WL-MMSE equalizers. As a side remark about Fig. 3.2, observe that, for the considered sample size, the ABER performances of the WL-BS-CCM and WL-BS-MMSE equalizers are superior to those of their corresponding WL-FS-CCM and WL-FS-MMSE counterparts. This behavior stems from the fact that, for the WL-FS-CCM and WL-FS-MMSE equalizers, one has to estimate $2\,N\,L_e$ (complex) parameters, whose number is doubled with respect to the number of parameters that must be estimated for the WL-BS-CCM and WL-BS-MMSE equalizers; strictly speaking, reducing the number of parameters to be adapted allows one to reduce the performance

degradation due to the finite sample-size.

In the second experiment, the ISI suppression capabilities of the considered equalizers were studied as a function of the sample size K, with SNR $= 20$ dB. It can be seen from Fig. 3.3 that, due to the presence of undesired global minima, the performances of the WL-FS-CM, L-FS-CM and WL-BS-CM equalizers do not significantly improve as K grows. In contrast, the ISI suppression capabilities of both the WL-FS-CCM and WL-BS-CCM equalizers rapidly improve as K increases. Remarkably, the ISI suppression capabilities of both WL-FS-CCM and WL-BS-CCM equalizers turn out to be better than those of all the MMSE equalizers, for all the considered values of K.

Chapter 4

Equalization Techniques for DS-CDMA Systems

In this chapter we present the general concepts regarding DS-CDMA systems and with reference to Minimum Output Energy Criterion (MOE) we present a theoretical performance analysis of WL multiuser receivers for direct-sequence code-division multiple-access (DS-CDMA) systems, as well as a performance comparison with the conventional linear (L) systems. Receivers based on the minimum output-energy (MOE) criterion are considered, since they offer a good tradeoff between performance and complexity and, moreover, lend to some simplifications in the analysis. After comparing the ideal signal-to-interference-plus-noise-ratio (SINR) performances of the WL-MOE and L-MOE receivers, the chapter presents finite-sample performance results for two typical data-estimated implementations. Specifically, by adopting a first-order perturbative approach, the SINR degradation of the data-estimated WL-MOE receivers is accurately evaluated and compared with that of its linear counterpart. Simulation results are provided to validate and complement the theoretical analysis.

4.1 Introduction

The Code Division Multiple Access (CDMA) is a multiple access technique based on spread-spectrum modulation method. Spread-Spectrum systems have a long story in military and civilian wireless communications and have been developed since about the mid-1950's [40, 41, 42, 20, 19]. Spread Spectrum is a transmission technique in which the transmitted signal exhibits a band-

width in excess of the minimum necessary to send the information. The band spread is accomplished by means of a code which is independent of the data. At the receiver side we need to despreading the transmitted signal to data recovery. For this purpose the received signal is correlated with a synchronized copy of the spreading code. These features distinguish spread-spectrum modulation from other signaling techniques that increase the transmit bandwidth above the minimum required for data transmission, for example frequency modulation and block and convolution coding. An interesting tradeoff arises as to whether, given a specific spreading bandwidth, it is more beneficial to use coding or spread spectrum. The answer depends on the requirements of the system design [19, 43]. In [40] it is shown that there are many reasons for spreading the spectrum, and if properly performed, a multiplicity of benefits can accrue simultaneously. Some of these are: antijamming, antiinterference, low interception probability, multiple user random access communications with selective addressing capability, high resolution ranging, accurate universal timing. Therefore Spread-spectrum can be very useful in solving a wide range of communications problems. The amount of performance improvement that is achieved through the use of spread spectrum is often related to the *processing gain* of the spread-spectrum system. That is, processing gain is often defined as the difference between system performance using spread-spectrum techniques and system performance not using spread-spectrum techniques. Processing gain is approximately the ratio of the spread bandwidth to information rate. The means by which the spectrum is spread is crucial. Several of the techniques are referred to as “direct-sequence” (DS) modulation in which a fast pseudorandomly generated sequence causes phase transition in the carrier containing data, whereas others are called “frequency hopping” (FH) ones, in which the carrier is frequency shifted in a pseudorandom way. Hybrid combinations of these techniques are frequently used. In the sequel, we consider DS-CDMA systems for their valuable properties, among which the narrowband interference and multipath rejecting.

4.2 CDMA Signal Model

In DS-CDMA, each user possess its own code which is used to modulate its data signal.The transmitted signals for all users are superimposed in time and frequency. The performances of this strategy are related to the correlation properties of the used spreading codes [40, 42, 20, 19]. In particular the autocorrelation function of the spreading code determines its multipath rejec-

tion properties. The cross-correlation properties of different spreading codes determine the amount of interference among users. The lower is the cross-correlation, the lower is the interference among users. A code set is orthogonal if the cross-correlation between the spreading codes of all user are equal to zero; in such a case the codes are able to eliminate interferences. A set of spreading code that does not satisfy this cross-correlation property, is called a non-orthogonal code set. Therefore, when orthogonal spreading codes are used, the other-user interference, called multiple-access interference (MAI), does not affect the post-despreading functions at all. Unfortunately, the multipath fading channels distort the signal in such a manner that the orthogonality present at the transmitter is loss at the receiver side. Further, in some applications orthogonal spreading codes are not even used. In these cases the performance is a function of the cross-correlation properties of the codes as well as of the channel properties. These partial correlations will eventually limit the total number of users that can simultaneously access the system.

Let us consider a DS-CDMA channel that is shared by J simultaneously users [44, 45, 25]. Each user is assigned a signature waveform $s_j(t)$ of duration T, where T is the information symbol interval. The signature waveform may be expressed as

$$s_j(t) = \sum_{n=0}^{N-1} c_j(n) p_{T_c}(t - nT_c), \quad 0 \leq t \leq T \tag{4.1}$$

where $\{c_j(n)\}_{n=0}^{N-1}$ is the code sequence associated to jth user consisting of N chips and T_c is the chip interval. Therefore, we can note that Direct-Sequence refers to a specific approach to construct spread-spectrum waveforms in which the normalized chip waveforms $\{p_{Tc}\}$ of duration $T_c = T/N$, are delayed versions of each others. With more details the chip waveform p_{T_c} is orthogonal to any version of itself delayed by an integer multiple of T_c. The signature waveforms are normalized so as to have unit energy. The equivalent lowpass transmitted signal due to the jth user can be expressed as

$$x_j(t) = \sum_{i=-\infty}^{+\infty} b_j(i) s_j(t - iT - \tau_j) \tag{4.2}$$

where $\{b_j(i)\}$ and τ_j denote, respectively, the symbol stream and transmission delay of the jth user with $0 \leq \tau_j < T$. We can note that this model is appropriate when we employ short spreading codes. The signature waveforms

propagate through their respective time invariant multipath channels whose impulse responses are:

$$g_j(t) = \sum_{m=0}^{M_j-1} \xi_{m,j}\delta(t-\tau_{m,j}), \quad \text{with } j \in \{1,2,\ldots,J\} \tag{4.3}$$

where M_j denotes the number of paths, $\{\xi_{m,j}\}$ the complex-valued path gains and $\{\tau_{m,j}\}$ the path delays. We assume, without loss in generality, that the delays related to LOS path are equal to zero for all users, $\tau_{0,j} = 0,\ j \in \{1,2,\ldots,J\}$. In this way the transmission delays $\{\tau_j\}$ must be considered to evaluate the asynchronisms between users.

Therefore, at the receiver, the received signal due to the jth users, recalling (4.2), (4.1) and (4.3), is given by

$$\begin{aligned} u_j(t) = x_j(t) * g_j(t) &= \sum_{m=0}^{M_j-1} \xi_{m,j} x_j(t-\tau_{m,j}) \\ &= \sum_{i=-\infty}^{+\infty} b_j(i) \sum_{m=0}^{M_j-1} \xi_{m,j} s_j(t-iT-\tau_j-\tau_{m,j}) \\ &= \sum_{i=-\infty}^{+\infty} b_j(i) \sum_{m=0}^{M_j-1} \xi_{m,j} \sum_{n=0}^{N-1} c_j(n)\, p_{T_c}(t-iT-nTc-\tau_j-\tau_{m,j}) \\ &= \sum_{i=-\infty}^{+\infty} b_j(i) \sum_{n=0}^{N-1} c_j(n)\, \overline{g}(t-iT-nTc-\tau_j) \end{aligned} \tag{4.4}$$

where $\overline{g}_j(t) = \sum_{m=0}^{M_j-1} \xi_{m,j} p_{T_c}(t-\tau_{m,j})$ is the convolution between chip waveform and the jth complex channel impulse response $g_j(t)$. The total received signal at the receiver is the superposition of the data signals of the J users plus the additive white Gaussian noise, given by

$$u_a(t) = \sum_{j=1}^{J} u_j(t) + n_a(t) \tag{4.5}$$

where $n_a(t)$ is a zero mean complex white Gaussian noise with power spectral density σ^2. The signal model given by (4.4)−(4.5) represents a dispersive asynchronous CDMA channel, which is typical for the uplink channel (i.e., mobile to base station) of a CDMA network. The downlink channel (i.e., base

station to mobile) of a CDMA network is a special case of this model, where the data signals of the J users are synchronous, i.e., $\tau_1 = \tau_2 = \cdots = \tau_J = 0$ and they propagate through a single dispersive channel, i.e., $g_1(t) = g_2(t) = \cdots = g_J(t)$. Since the signal processing required for interference suppression is digitally performed, we resort to a discrete time model obtained by chip matched filtering the received signal and sampling at a chip rate. Therefore at the output of the chip matched filter, the received signal can be expressed as

$$\begin{aligned} r(t) &= u_a(t) * p^*_{T_c}(-t) \\ &= \sum_{j=1}^{J} \sum_{i=-\infty}^{+\infty} b_j(i) \sum_{n=0}^{N-1} c_j(n)\, \overline{g}_j(t - iT - nTc - \tau_j) * p^*_{T_c}(-t) \\ &+ n_a(t) * p^*_{T_c}(-t) \\ &= \sum_{j=1}^{J} \sum_{i=-\infty}^{+\infty} b_j(i) \sum_{n=0}^{N-1} c_j(n)\, f_{a,j}(t - iT - nTc - \tau_j) + v(t) \\ &= \sum_{j=1}^{J} \sum_{i=-\infty}^{+\infty} b_j(i) \varphi_{a,j}(t - iT - \tau_j) + v(t), \end{aligned} \tag{4.6}$$

where $f_{a,j}(t) = \overline{g}_j(t) * p^*_{T_c}(-t)$ is the composite channel impulse response that includes the transmitted, the received filter and the channel impulse response; $\varphi_{a,j}(t) = \sum_{n=0}^{N-1} c_j(n)\, f_{a,j}(t - nTc)$ is the composite received signal waveform of the jth user, encompassing spreading code and channel propagation effects and $v(t)$ is the noise process after the convolution with the chip-matched filter. We assume that the delays τ_j are not necessarily integer multiple of T_c therefore they are constituted by a integer part and by a fractionally part, $\tau_j = (\beta_j + \xi_j)T_c$, with $\beta_j \in [0, 1, \ldots, N)$ and $\xi_j \in [0, 1)$. At the output of the sampler, the resulting discrete-time signal during the lth chip period of the kth symbol interval recalling (4.6), is given by

$$\begin{aligned} r_l(k) &\stackrel{\triangle}{=} r(kT + lT_c) \\ &= \sum_{j=1}^{J} \sum_{i=-\infty}^{+\infty} b_j(i) \sum_{n=0}^{N-1} c_j(n)\, f_{a,j}(kT + lT_c - iT - nTc - (\beta_j + \xi_j)T_c) \\ &+ v(kT + lT_c) \\ &= \sum_{j=1}^{J} \sum_{i=-\infty}^{+\infty} b_j(i) \overline{\varphi}_{l,j}(k - i - \beta_j) + v_l(k) \end{aligned} \tag{4.7}$$

where $\overline{\varphi}_{l,j}(k) = \sum_{n=0}^{N-1} c_j(n)\, f_{l,j}(k-n)$ with $f_{l,j}(k) = f_{a,j}(kT + lT_c - \xi_j T_c)$ and $v_l(k) = v(kT + lT_c)$. If we define in (4.7) $\varphi_{l,j}(k) = \overline{\varphi}_{l,j}(k - \beta_j)$, we can write (4.7) in this way:

$$r_l(k) = \sum_{j=1}^{J} \sum_{i=-\infty}^{+\infty} b_j(i)\psi_{l,j}(k-i) + v_l(k). \tag{4.8}$$

We can collect the N samples of the received signal to obtain the vector model:

$$\mathbf{r}(k) = \sum_{j=1}^{J} \sum_{i=-\infty}^{+\infty} b_j(i)\boldsymbol{\varphi}_j(k-i) + \mathbf{v}(k), \tag{4.9}$$

where we have defined $\mathbf{r}(k) \triangleq [r_0(k), r_1(k), \ldots, r_{N-1}(k)]^T$, $\boldsymbol{\varphi}_j(k) \triangleq [\varphi_{0,j}(k), \varphi_{1,j}(k), \ldots, \varphi_{N-1,j}(k)]^T$ and $\mathbf{v}(k) \triangleq [v_0(k), v_1(k), \ldots, v_{N-1}(k)]^T$.

The model (4.9) can be simplified if we assume that the channel is synchronous and that the intersymbol interference (ISI) can be neglected. In the sequel, we consider a synchronous DS-CDMA system with J users, employing short codes with $1/T_c = N/T$ chip/symbol and transmitting over channels that introduce interchip interference and negligible intersymbol interference [42]. It is noteworthy that synchronous transmissions and negligible intersymbol interference (ISI) are assumed only for the sake of simplicity and the analysis, that we develop in the next sections, can be readily generalized to other scenarios (e.g., asynchronous users and/or channels with ISI). For instance, in a asynchronous system with J_a users, can be described by a synchronous model with $J \leq 2J_a$ equivalent users (see [46] for further details). Therefore (4.9) can written as:

$$\begin{aligned}\mathbf{r}(k) &= \sum_{j=1}^{J} \boldsymbol{\varphi}_j\, b_j(k) + \mathbf{v}(k) = \sum_{j=1}^{J} \alpha_j\, \boldsymbol{\psi}_j\, b_j(k) + \mathbf{v}(k) \\ &= \boldsymbol{\Psi}\mathbf{A}\mathbf{b}(k) + \mathbf{v}(k) = \boldsymbol{\Phi}\, \mathbf{b}(k) + \mathbf{v}(k)\,,\end{aligned} \tag{4.10}$$

where, with reference to the the jth user, we have indicate with $\boldsymbol{\psi}_j \in \mathbb{C}^N$ the *unit-norm* signature (encompassing spreading code and channel propagation effects). As a consequence, $\alpha_j > 0$ is the received amplitude, accounting for transmitted energy and channel propagation loss. Moreover, in (4.10), we have defined $\boldsymbol{\Psi} \triangleq [\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, \ldots, \boldsymbol{\psi}_J] \in \mathbb{C}^{N\times J}$, $\mathbf{A} \triangleq \mathrm{diag}(\alpha_1, \alpha_2, \ldots, \alpha_J) \in$

$\mathbb{R}^{J\times J}$, $\mathbf{\Phi} \stackrel{\triangle}{=} \mathbf{\Psi}\,\mathbf{A} \in \mathbb{C}^{N\times J}$, and $\mathbf{b}(k) \stackrel{\triangle}{=} [b_1(k), b_2(k), \ldots, b_J(k)]^T \in \mathbb{C}^J$. Finally, we recall that $\mathbf{c}_j \stackrel{\triangle}{=} [c_j(0), c_j(1), \ldots, c_j(N-1)]^T \in \mathbb{C}^N$ denotes[1] the spreading code vector of the jth user and $g_j(n)$ is the channel impulse response of length $L_j \ll N$ ($L_j > 1$), with $\mathbf{g}_j \stackrel{\triangle}{=} [g_j(0), g_j(1), \ldots, g_j(L_j-1)]^T \in \mathbb{C}^{L_j}$ being the corresponding *unit-norm* channel vector. Under the assumption that $g_j(n)$ has order $L_j \ll N$, the signature $\boldsymbol{\psi}_j$ in (4.10) can be modeled [15] as

$$\boldsymbol{\psi}_j = \mathbf{G}_j\,\mathbf{c}_j\,, \tag{4.11}$$

where $\mathbf{G}_j \in \mathbb{C}^{N\times N}$ is the Toeplitz lower triangular matrix having $[g_j(0), 0, \ldots, 0]^T$ as first row and $[g_j(0), g_j(1), \ldots, g_j(L_j-1), 0, \ldots, 0]^T$ as first column. It is worth noticing that the signature $\boldsymbol{\psi}_j$ in (4.10) can be modeled also as

$$\boldsymbol{\psi}_j = \mathbf{C}_j\,\mathbf{g}_j\,, \tag{4.12}$$

where $\mathbf{C}_j \in C^{N\times L_j}$ is the Toeplitz matrix having $[c_j(0), 0, \ldots, 0]^T$ as first row and $[c_j(0), c_j(1), \ldots, c_j(N-1)]^T$ as first column. Throughout this chapter, we will rely on these assumptions:

(a1) $\mathbf{b}(k)$ is a binary[2] *real* zero-mean random vector, whose entries are independent and identically distributed (i.i.d.) random variables assuming equiprobable values in $\mathcal{B} = \{-1, 1\}$, with $\mathbf{b}(k_1)$ and $\mathbf{b}(k_2)$ statistically independent for $k_1 \neq k_2$;

(a2) $\mathbf{v}(k)$ is a *complex proper* zero-mean Gaussian random vector, independent of $\mathbf{b}(k)$, having $\mathbf{R}_{\mathbf{vv}} = \sigma_v^2\,\mathbf{I}_N$ and $\mathbf{R}_{\mathbf{vv}^*} = \mathbf{O}_{N\times N}$, with $\mathbf{v}(k_1)$ and $\mathbf{v}(k_2)$ statistically independent of each other for $k_1 \neq k_2$.

4.3 Multiuser Detection

A DSSS receiver that exploits the structure of multiuser interference in signal detection is called a multiuser detector (MUD) [48, 10, 25, 19]. During the last two decades, starting from the seminal works of Verdú [48, 10], a great bulk

[1] The code $\mathbf{c}_j$ accounts also for possible precoding phases, whose role in downlink is discussed in the subsection 4.4.1

[2] This assumptions is not crucial, but simplifies the analysis. Our derivations can be readly extended to the case where the entries of $\mathbf{b}(k)$ assume values in an arbitrary real set, or even when the entries of $\mathbf{b}(k)$ are not real but obey the more general conjugate symmetry [47] property

of research activities has been devoted to multiuser detection (MUD), as an effective way to counteract the multiple-access interference (MAI), which is the predominant source of performance degradation in nonorthogonal DS-CDMA systems. Verdú's solution involves a bank of single-user matched filters followed by a Viterbi algorithm. The complexity of this procedure is exponential in the number of users. The complexity of MUD can be decreased at the expense of optimality . The simplest suboptimum detector is the conventional single-user detector in which the receiver for each user is constituted by a demodulator that correlates the received signal with the signature sequence of the user and passes the correlator output to the detector which makes a decision based on the single correlator output. Thus the conventional detector neglects the presence of the other users of the channel, or, equivalently assumes that the aggregate noise plus interference is white and gaussian. If the signature are orthogonal, the interference from the other users vanishes and the conventional single-user detector is optimum [48, 10, 25, 19]. On the other hand, if one or more of the other signature sequences are not orthogonal to the user signature sequence, the interference from the other users does not vanish and it can became excessive if the power level of the signal of one or more of the other users is sufficiently larger than the power level of the user of interest. This situation is generally called *near-far problem* in multiuser communications and necessities some type of power control for conventional detector. In asynchronous transmission, the conventional detector is more vulnerable to interference from other users because it is not possible to design signature sequences for any pair of users that are orthogonal for all time offsets. Consequently interference from other users is unavoidable with the conventional single-user detection. In such a case, the *near-far problem*, resulting from unequal power in the signals transmitted by the various users is particulary serious. The practical solution generally requires a power adjustment method that is controlled by the receiver via separate communication channel that all users are continuously monitoring. If a separate communication channel is not available, it is possible to employ a suboptimum multiuser detectors. Suboptimal MUDs fall into two broad categories: linear and nonlinear. Linear MUDs apply a linear operator or filter to the output of the matched filter bank. These linear detectors have complexity that is linear in the number of users, a significant complexity improvement over the optimal detector. Among linear MUD techniques, decorrelating receiver [49], the minimum mean-square-error (MMSE) [46] one, and the minimum output-energy (MOE) [15] one, have been investigated in depth, since they offer convenient tradeoffs between performance, complexity, robustness,

amount of *a priori* information, and ease of adaptive implementation. Nonlinear MUDs have somewhat larger complexity than the linear detectors but also much better performance, although not necessarily in all cases, especially with very limited or no coding. The most common nonlinear MUD techniques are multistage detection, decision-feedback detection, and successive interference cancelation. Linear multiuser detectors can be implemented in a decentralized fashion where only the user or users of interest need be demodulated [45]. Therefore generally, they require the only knowledge of the interest user's spreading sequence. In this thesis we focalize our attention on L-MUD techniques. In particulary, we can note that most L-MUD techniques assume that the complex envelope $r(t)$ of the received signal is modeled as a *proper* random process, exploiting hence only the information contained in its statistical autocorrelation function $R_{rr}(t,\tau) \stackrel{\triangle}{=} \mathrm{E}[r(t)\, r^*(t-\tau)]$. When, however, the DS-CDMA signal and/or the disturbance are *improper*, well-established results in detection and estimation theory, as shown in section 2.4, state that linear receivers can be outperformed by *widely-linear* (WL) ones, which jointly elaborate the received signal $r(t)$ and its complex conjugate $r^*(t)$, in order to exploit also the information contained in their statistical *cross-correlation* function $R_{rr^*}(t,\tau) \stackrel{\triangle}{=} \mathrm{E}[r(t)\, r(t-\tau)]$. Motivated from previous observations, in recent years several papers [3, 50, 51, 52] proposed different WL-MUD techniques for DS-CDMA systems with improper signals and/or disturbances, by extending concepts from the classical L-MUD theory. In particular, WL versions of the major L-MUD receivers have been proposed and studied, such as the WL decorrelating receiver [51, 53], the WL-MMSE one [3, 50, 54], the WL-MOE one [52, 53], and the min/max WL-MOE one [55]. In the sequel, we focalize our attention on the minimum output energy criterion (MOE).

4.3.1 Minimum Output Energy Criterion (MOE): Linear and Widely-Linear receivers

The main goal of this section is to derive the WL-MOE receiver as a particular solution of the maximum SINR criterion. We start by reviewing briefly the L-MOE receiver, not only to put the necessary bases for our subsequent derivations, but also to comment on possible inconsistencies concerning the "correct" definition of the SINR to be used for linear receivers, when real symbols are employed.

In order to recover $b_j(k)$ by a linear receiver, it is useful to rewrite (4.10)

as follows:

$$\mathbf{r}(k) = \boldsymbol{\phi}_j\, b_j(k) + \overline{\boldsymbol{\Phi}}_j\, \overline{\mathbf{b}}_j(k) + \mathbf{v}(k) = \boldsymbol{\phi}_j\, b_j(k) + \mathbf{p}_j(k)\,, \tag{4.13}$$

where $\boldsymbol{\phi}_j \in \mathbb{C}^N$ is the jth column of the composite matrix $\boldsymbol{\Phi}$, whereas $\overline{\mathbf{b}}_j(k) \in \mathbb{R}^{J-1}$ denotes the vector that includes all the elements of $\mathbf{b}(k)$ except for the jth entry $b_j(k)$, $\overline{\boldsymbol{\Phi}}_j \in \mathbb{C}^{N\times(J-1)}$ denotes the matrix that includes all the columns of $\boldsymbol{\Phi}$ except for the jth column $\boldsymbol{\phi}_j$, and, finally, $\mathbf{p}_j(k) \stackrel{\triangle}{=} \overline{\boldsymbol{\Phi}}_j\, \overline{\mathbf{b}}_j(k) + \mathbf{v}(k) \in \mathbb{C}^N$ is the interference-plus-noise (disturbance) vector. Accounting for (4.13), the output of a linear receiver can be expressed as

$$y_j(k) = \mathbf{w}_j^H \mathbf{r}(k) = \mathbf{w}_j^H \boldsymbol{\phi}_j\, b_j(k) + \mathbf{w}_j^H \mathbf{p}_j(k)\,. \tag{4.14}$$

The L-MOE receiver [15] is the solution of the following constrained optimization problem:

$$\mathbf{w}_{j,\text{L-MOE}} = \underset{\mathbf{w}_j\in\mathbb{C}^N}{\operatorname{argmin}}\, \mathrm{E}[|y_j(k)|^2] \quad \text{subject to } \mathbf{w}_j^H\, \boldsymbol{\phi}_j = 1\,, \tag{4.15}$$

which can be solved by Lagrange optimization, yielding the two equivalent [3] expressions

$$\mathbf{w}_{j,\text{L-MOE}} = (\boldsymbol{\phi}_j^H \mathbf{R}_{\mathbf{rr}}^{-1} \boldsymbol{\phi}_j)^{-1}\, \mathbf{R}_{\mathbf{rr}}^{-1}\, \boldsymbol{\phi}_j = (\boldsymbol{\phi}_j^H \mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}^{-1} \boldsymbol{\phi}_j)^{-1}\, \mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}^{-1}\, \boldsymbol{\phi}_j\,, \tag{4.16}$$

where the second equality follows by applying the matrix inversion lemma[4] to the autocorrelation matrix $\mathbf{R}_{\mathbf{rr}} = \boldsymbol{\phi}_j \boldsymbol{\phi}_j^H + \mathbf{R}_{\mathbf{p}_j\mathbf{p}_j} \in \mathbb{C}^{N\times N}$.

It can be easily shown that, among all linear receivers, the L-MOE one maximizes the SINR at its output, which, accounting for (4.14), can be defined as

$$\overline{\text{SINR}}(\mathbf{w}_j) \stackrel{\triangle}{=} \frac{\mathrm{E}[|\mathbf{w}_j^H \boldsymbol{\phi}_j\, b_j(k)|^2]}{\mathrm{E}[|\mathbf{w}_j^H \mathbf{p}_j(k)|^2]} = \frac{|\mathbf{w}_j^H \boldsymbol{\phi}_j|^2}{\mathbf{w}_j^H \mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}\, \mathbf{w}_j} = \frac{|(\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}^{1/2}\, \mathbf{w}_j)^H (\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}^{-1/2} \boldsymbol{\phi}_j)|^2}{\|\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}^{1/2}\, \mathbf{w}_j\|^2}\,. \tag{4.17}$$

Indeed, by using the Cauchy-Schwartz's inequality[5], any receiver maximizing (4.17) is given by $\mathbf{w}_{j,\text{max-SINR}} = \gamma_j\, \mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}^{-1}\, \boldsymbol{\phi}_j$, where $\gamma_j \in \mathbb{C} - \{0\}$ is an

[3] The advantage of using $\mathbf{R}_{\mathbf{rr}}$ instead of $\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}$ in (4.16) is that the former can be estimated from received data.

[4] Given the vectors $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$ and the nonsingular matrix $\mathbf{X} \in \mathbb{C}^{n\times n}$, the matrix inversion lemma states that $(\mathbf{X} + \mathbf{x}\,\mathbf{y}^H)^{-1} = \mathbf{X}^{-1} - (1 + \mathbf{y}^H\, \mathbf{X}^{-1}\, \mathbf{x})^{-1} \mathbf{X}^{-1}\, \mathbf{x}\, \mathbf{y}^H\, \mathbf{X}^{-1}$.

[5] Given the vectors $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$, the Cauchy-Schwartz's inequality states that $|\mathbf{x}^H \mathbf{y}|^2 \leq \|\mathbf{x}\|^2 \|\mathbf{y}\|^2$, where the upper bound is achieved by $\mathbf{y} = \gamma\, \mathbf{x}$, with $\gamma \in \mathbb{C}$.

arbitrary (nonnull) complex scalar. Hence, the L-MOE receiver is obtained by setting $\gamma_j = (\boldsymbol{\phi}_j^H \mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}^{-1} \boldsymbol{\phi}_j)^{-1}$, and the maximum value of (4.17) is

$$\overline{\text{SINR}}_{j,\max} \triangleq \overline{\text{SINR}}(\mathbf{w}_{j,\text{L-MOE}}) = \frac{1}{\mathbf{w}_{j,\text{L-MOE}}^H \mathbf{R}_{\mathbf{p}_j\mathbf{p}_j} \mathbf{w}_{j,\text{L-MOE}}} = \boldsymbol{\phi}_j^H \mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}^{-1} \boldsymbol{\phi}_j \,. \tag{4.18}$$

On the other hand, the output of a widely-linear receiver can be expressed as [12, 2] as

$$y_j(k) = \mathbf{f}_{j,1}^H \mathbf{r}(k) + \mathbf{f}_{j,2}^H \mathbf{r}^*(k) = \mathbf{f}_j^H \mathbf{z}(k) \,, \tag{4.19}$$

where $\mathbf{f}_j \triangleq [\mathbf{f}_{j,1}^T, \mathbf{f}_{j,2}^T]^T \in \mathbb{C}^{2N}$ and $\mathbf{z}(k) \triangleq [\mathbf{r}^T(k), \mathbf{r}^H(k)]^T \in \mathbb{C}^{2N}$ is the *augmented* received vector. According to (4.10), vector $\mathbf{z}(k)$ can be expressed as

$$\mathbf{z}(k) = \mathbf{H}\,\mathbf{b}(k) + \mathbf{d}(k) \,, \tag{4.20}$$

with $\mathbf{H} \triangleq [\boldsymbol{\Phi}^T, \boldsymbol{\Phi}^H]^T \in \mathbb{C}^{2N\times J}$ and $\mathbf{d}(k) \triangleq [\mathbf{v}^T(k), \mathbf{v}^H(k)]^T \in \mathbb{C}^{2N}$, where, for **(a2)**, the noise $\mathbf{d}(k)$ is an *improper* Gaussian random vector, with $\mathbf{R}_{\mathbf{dd}} = \sigma_v^2\,\mathbf{I}_{2N}$ and $\mathbf{R}_{\mathbf{dd}^*} = \sigma_v^2\,\mathbf{J}_{2N}$, where

$$\mathbf{J}_{2N} \triangleq \begin{bmatrix} \mathbf{O}_{N\times N} & \mathbf{I}_N \\ \mathbf{I}_N & \mathbf{O}_{N\times N} \end{bmatrix} \in \mathbb{R}^{2N\times 2N} \tag{4.21}$$

is a block permutation matrix [56].

Accounting for (4.20), eq. (4.19) can be equivalently written as

$$y_j(k) = \mathbf{f}_j^H \mathbf{h}_j\, b_j(k) + \mathbf{f}_j^H [\overline{\mathbf{H}}_j\, \overline{\mathbf{b}}_j(k) + \mathbf{d}(k)] = \mathbf{f}_j^H \mathbf{h}_j\, b_j(k) + \mathbf{f}_j^H\, \mathbf{q}_j(k) \,, \tag{4.22}$$

where $\mathbf{h}_j \triangleq [\boldsymbol{\phi}_j^T, \boldsymbol{\phi}_j^H]^T \in \mathbb{C}^{2N}$, with $\boldsymbol{\phi}_j \in \mathbb{C}^N$ being the jth column of the matrix $\boldsymbol{\Phi}$, whereas $\overline{\mathbf{H}}_j \triangleq [\overline{\boldsymbol{\Phi}}_j^T, \overline{\boldsymbol{\Phi}}_j^H]^T \in \mathbb{C}^{2N\times(J-1)}$, with $\overline{\boldsymbol{\Phi}}_j \in \mathbb{C}^{N\times(J-1)}$ denoting the matrix that includes all the columns of $\boldsymbol{\Phi}$ except for the jth column $\boldsymbol{\phi}_j$, $\overline{\mathbf{b}}_j(k) \in \mathbb{R}^{J-1}$ denotes the vector that includes all the elements of $\mathbf{b}(k)$ except for the jth entry $b_j(k)$, and $\mathbf{q}_j(k) \triangleq \overline{\mathbf{H}}_j\, \overline{\mathbf{b}}_j(k) + \mathbf{d}(k) \in \mathbb{C}^{2N}$ is the augmented disturbance (interference-plus-noise) vector.

To establish a general framework encompassing both linear and WL receivers, we refer to the scheme in Fig. 4.1, wherein linear receivers can be obtained by setting $\mathbf{f}_{j,2} = \mathbf{0}_N$ in (4.19), and the Re$[\cdot]$ operation is needed only when $y_j(k)$ is complex, as it happens for linear receivers, or even for WL ones

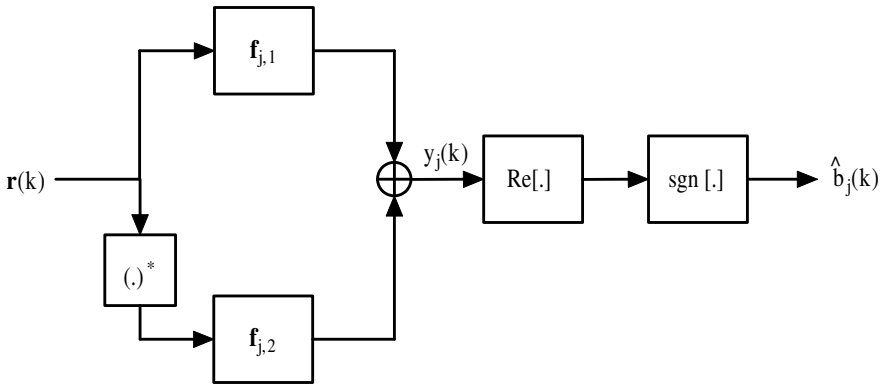

Figure 4.1: The WL processing scheme.

possibly not satisfying the CS constraint($\mathbf{f}_{j,1} = \mathbf{f}_{j,2}^*$ is referred to as *conjugate symmetry* (CS) property). It should be observed that the L-MOE receiver maximizes the $\overline{\text{SINR}}$ given by (4.17), which is evaluated *before* the $\text{Re}[\cdot]$ block. Since, by virtue of **(a1)**, $b_j(k)$ is real-valued, an appropriate performance measure for the jth user is the output SINR (*after* the $\text{Re}[\cdot]$ block) [57, 58] defined as

$$\text{SINR}(\mathbf{f}_j) = \frac{\text{E}^2\{\text{Re}[y_j(k)] \,|\, b_j(k)\}}{\text{Var}\{\text{Re}[y_j(k)] \,|\, b_j(k)\}} \,. \tag{4.23}$$

Indeed, if the disturbance contribution $\mathbf{f}_j^H \, \mathbf{q}_j(k)$ at the receiver output can be approximated as a Gaussian random variable[6], maximizing (4.23) w.r.t $\mathbf{f}_j$ amounts to minimizing the error probability $P_{e,j} \stackrel{\triangle}{=} \Pr\{\hat{b}_j(k) \neq b_j(k)\} \approx Q(\sqrt{\text{SINR}(\mathbf{f}_j)})$, where $Q(x) \stackrel{\triangle}{=} (1/\sqrt{2\pi}) \int_x^{+\infty} e^{-u^2/2} \, du$ denotes the Q function. Definition (4.23) of the SINR is quite general and allows for relatively simple calculations when $\mathbf{h}_j$ and/or $\mathbf{f}_j$ are estimated from data as we will see in the next section. In the particular case where both $\mathbf{f}_j$ and $\mathbf{h}_j$ are perfectly known, it can be shown that (4.23) reduces to

$$\text{SINR}(\mathbf{f}_j) = \frac{\text{Re}^2[\mathbf{f}_j^H \mathbf{h}_j]}{\text{E}\{\text{Re}^2[\mathbf{f}_j^H \mathbf{q}_j(k)]\}} \,, \tag{4.24}$$

which can be employed also when $\mathbf{f}_j$ is estimated from a finite sample-size, under the assumption that the channel is exactly known. Since maximization

[6] When N and J are large enough, this assumption is well-satisfied for maximum-SINR equalizers (see, e.g, [59]).

of (4.24), due to the presence of the Re[·] operator, is not as standard as maximizing (4.17), we discuss it briefly in the following Lemma.

Lemma 4.1 *Any WL receiver* (4.19) *maximizing* (4.24) *can be expressed as*

$$\mathbf{f}_{j,\text{max-SINR}} = \xi_j \, \mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}^{-1} \mathbf{h}_j + \mathbf{f}_{j,\text{a}} \,, \tag{4.25}$$

and $\mathbf{f}_{j,\text{a}}$ *is an arbitrary antisymmetric vector, i.e.,* $\mathbf{f}_{j,\text{a}} \in \mathcal{A} \stackrel{\triangle}{=} \{\mathbf{f} = [\mathbf{f}_1^T, \mathbf{f}_2^T]^T \in \mathbb{C}^{2N} \,|\, \mathbf{f}_1 = -\mathbf{f}_2^* \in \mathbb{C}^N\}$. *The resulting maximum* SINR *is given by*

$$\text{SINR}_{j,\text{max}} \stackrel{\triangle}{=} \mathit{SINR}(\mathbf{f}_{j,\text{max-SINR}}) = \mathbf{h}_j^H \mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}^{-1} \mathbf{h}_j \,. \tag{4.26}$$

Proof. See Appendix B.1.

Note that the maximum SINR solution (4.25) differs from that of the linear case for the fact that the scalar ξ_j must be real and for the presence of the antisymmetric vector $\mathbf{f}_{j,\text{a}}$. Moreover, in Appendix B.1 it is also shown that the value of SINR (4.24) does not depend on $\xi_j \in \mathbb{R} - \{0\}$ and on $\mathbf{f}_{j,\text{a}}$. Hence, we can choose ξ_j such that $\mathbf{f}_{j,\text{max-SINR}}^H \mathbf{h}_j = 1$ and $\mathbf{f}_{j,\text{a}} = \mathbf{0}_{2N}$, which leads to the WL-MOE receiver:

$$\mathbf{f}_{j,\text{WL-MOE}} = (\mathbf{h}_j^H \mathbf{R}_{\mathbf{zz}}^{-1} \mathbf{h}_j)^{-1} \, \mathbf{R}_{\mathbf{zz}}^{-1} \, \mathbf{h}_j = (\mathbf{h}_j^H \mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}^{-1} \mathbf{h}_j)^{-1} \, \mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}^{-1} \mathbf{h}_j \,, \tag{4.27}$$

where the second equality[7] follows by applying the matrix inversion lemma (see footnote 4) to the autocorrelation matrix $\mathbf{R}_{\mathbf{zz}} = \mathbf{h}_j \mathbf{h}_j^H + \mathbf{R}_{\mathbf{q}_j\mathbf{q}_j} \in \mathbb{C}^{2N \times 2N}$. By reasoning as in the proof of Lemma 4.1, it can be shown that (4.27) is obtained equivalently as the unique solution of the following WL-MOE criterion:

$$\mathbf{f}_{j,\text{WL-MOE}} = \underset{\mathbf{f}_j \in \mathbb{C}^{2N}}{\text{argmin}} \, \text{E}\{\text{Re}^2[y_j(k)]\} \;\; \text{subject to} \; \mathbf{f}_j^H \mathbf{h}_j = 1 \,. \tag{4.28}$$

The ideal implementations of the L-MOE and WL-MOE receivers require perfect knowledge of two quantities: the *autocorrelation matrix* (ACM) of the received signal, and the *received signature* (possibly distorted by the channel) of each user to be demodulated. These two quantities can be estimated in practice from a finite number of samples at the receiver. In particular, we can note that when we consider a multipath channel, due to the effects of the unknown

[7] The advantage of using $\mathbf{R}_{\mathbf{zz}}$ instead of $\mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}$ in (4.27) is that the former can be estimated from received data.

channel response, the received signature is a distorted version of the transmitted one, making channel estimation (CE) a necessary step to implement both the L- and WL-MOE receivers. In such a scenario, we will show in the section 4.5 that the performances of the L- and WL-MOE receivers are affected by imperfect ACM estimation and by inaccurate CE.

4.3.2 Blind channel estimation: Subspace method

The conventional method for CE is to periodically transmit training sequences of data that are known *a priori* to the receiver. However, such a scheme might lead to a significant waste of power and bandwidth resources in mobile communication systems, especially when channel conditions require the use of long training sequences and/or frequent repetition of training. Consequently, the past few decades have witnessed a huge number of contributions in the area of *blind* CE approaches, which only exploit the knowledge of the spreading code of the desired user, without requiring any training, and allow one to demodulate the desired transmission without any knowledge of the channels and spreading sequences of the other users. Blind CE approaches relying on second-order statistics (SOS) of the received data are particularly attractive since they require fewer samples than those necessary fo traditional methods based on higher-order statistics [16]. Among existing SOS-based approaches, the subspace CE method first proposed in [60] is one of the most studied blind algorithm for DS-CDMA systems for the following reasons: (i) except for the subspace swap phenomenon, which occurs only for low values of the SNR well below the range of practical interest, it is very robust to noise [61]; (ii) it provides unique channel identification in closed form under mild conditions [62]; (iii) it is a method that not only provides a blind channel estimator but also a robust multiuser detector in the meantime [63]; (iv) it can be optimally combined with training-based approaches (so-called semi-blind methods) [64]; (v) it is amenable of a low-complexity and fast recursive implementation [65]. On the other hand, the main drawbacks of the subspace-based algorithm are the performance degradation when the number of active users is comparable to the code length and the requirement for accurate rank estimation of the ACM of the noise-free received signal. Under the assumption that the transmitted symbols are improper and the noise is proper, the former shortcoming can be overcame by resorting to a generalized subspace-based method, which allows one to enlarge the dimension of the observation space. More precisely, a generalized subspace-based approach exploits the channel information contained in both $R_{rr}(t,\tau)$ and $R_{rr^*}(t,\tau)$, by jointly processing $r(t)$ and its conjugate

version $r^*(t)$. Originally, such an estimation approach was proposed in [37]–[66] to improve channel identification in many application fields, including multicarrier CDMA and single-carrier DS-CDMA systems. To face up to the latter disadvantage, one can use conventional rank estimation techniques as the Akaike information criterion [67] and the minimum description length method [68], or, alternatively, a subspace tracking procedure with successive cancellation techniques [69]. In this subsection we describe, with reference to the L-MOE receiver, the multiuser CE procedure proposed in [61], whereby the impulse response of the desired user is obtained from $R_{rr}(t,\tau)$ by processing the received signal $r(t)$. On the other hand, as regards the WL-MOE receiver, the generalized subspace-based method of [66] is consider.

In the sequel we assume that the following conditions hold:
(c1) when $J \leq N$ (*underloaded systems*), the matrix $\mathbf{\Phi}$ is full-column rank, i.e., rank$(\mathbf{\Phi}) = J$. In section 4.4.1 we will show that, in the downlink case, wherein all the user signals propagate through a common multipath channel, the linear independence of the codes $\mathbf{c}_1$, $\mathbf{c}_2, \ldots, \mathbf{c}_J$ is a necessary and sufficient condition to ensure the rank condition **(c1)**. It is noteworthy that, if $\mathbf{\Phi}$ is full-column rank, the augmented matrix $\mathbf{H}$ is full-column rank, too, i.e., rank$(\mathbf{H}) = J$. In other words, in underloaded environments, condition **(c1)** additionally assures the full-column rank property of $\mathbf{H}$. However, the matrix $\mathbf{H}$ can be full-column rank even when $N < J \leq 2\,N$ (*overloaded systems*), wherein $\mathbf{\Phi}$ is inherently rank-deficient. Thus, in addition to condition **(c1)**, we assume hereinafter that:
(c2) when $N < J \leq 2\,N$, the matrix $\mathbf{H}$ is full-column rank, i.e., rank$(\mathbf{H}) = J$. With reference to the downlink scenario, fulfillment of condition **(c2)** is thoroughly discussed in Theorem 4.1.

Under assumptions **(a1)** and **(a2)**, the autocorrelation matrix of the observation vector $\mathbf{R_{rr}}$ assumes the form

$$\mathbf{R_{rr}} = \mathbf{\Phi}\,\mathbf{\Phi}^H + \sigma_v^2\,\mathbf{I}_N \tag{4.29}$$

. The correlation matrix can also expressed in term of its eigenvector decomposition being a unitarily diagonalizable matrix [56]; therefore there exists a unitary matrix $\mathbf{V}$ such that

$$\mathbf{R_{rr}} = \mathbf{V}\,\mathbf{\Upsilon}\,\mathbf{V}^H, \tag{4.30}$$

where $\mathbf{\Upsilon} \in \mathbb{C}^{N\times N}$ is a diagonal matrix of the eigenvalues $\{\varsigma_j\}_{j=1}^N$ in a nonincreasing order of $\mathbf{R_{rr}}$ and the columns of $\mathbf{V} \in \mathbb{C}^{N\times N}$ are the corresponding

eigenvectors. Moreover, we can recall that the matrix $\mathbf{R_{rr}}$ is positive semidefinite, hermitian matrix, therefore [56] its eigenvalues $\{\varsigma_j\}_{j=1}^{N}$ are all real and non-negative. Recalling the eigenvalue definition,

$$\mathbf{R_{rr}}\,\mathbf{a}_j = \varsigma_j\,\mathbf{a}_j \tag{4.31}$$

and substituting (4.29) in (4.31), we obtain

$$(\mathbf{\Phi}\,\mathbf{\Phi}^H + \sigma_v^2\,\mathbf{I}_N)\,\mathbf{a}_j = \varsigma_j\,\mathbf{a}_j \tag{4.32}$$

. From (4.32) we conclude that

$$\varsigma_j = \mu_j + \sigma_v^2 \quad j \in \{1, 2, \ldots, N\} \tag{4.33}$$

where $\{\mu_j\}$ are the eigenvalues of the matrix $\mathbf{\Phi}\,\mathbf{\Phi}^H$.

Nevertheless, by virtue of conditions **(c1)**, the matrix $\mathbf{\Phi}\,\mathbf{\Phi}^H$ has only J nonzero eigenvalues $\mu_1 \geq \mu_2 \cdots \geq \mu_J > 0$, therefore

$$\varsigma_j = \begin{cases} \mu_j + \sigma_v^2 & \text{if} j \in \{1, 2, \ldots, J\} \\ \sigma_v^2 & \text{if} j \in \{J+1, J+2, \ldots, N\} \end{cases} \tag{4.34}$$

Thus the eigenvalues can be separated into two distinct groups: the signal eigenvalues and the noise eigenvalues, respectively represent by the matrices

$$\mathbf{\Upsilon}_s \stackrel{\triangle}{=} \text{diag}[\mu_1, \mu_2 \cdots, \mu_J] \;\in \mathbb{R}^{J\times J} \tag{4.35}$$

$$\mathbf{\Upsilon}_n \stackrel{\triangle}{=} \text{diag}[\sigma_v, \sigma_v \cdots, \sigma_v] = \sigma_v^2\mathbf{I}_{N-J} \;\in \mathbb{R}^{(N-J)\times(N-J)}. \tag{4.36}$$

Accordingly, the eigenvectors can be separated into the signal and noise eigenvectors. In detail, denote the unit-norm eigenvectors associated with the signal eigenvalues by $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_J$ and denote those corresponding to the noise eigenvalues by $\mathbf{u}_{J+1}, \mathbf{u}_{J+2}, \ldots, \mathbf{u}_N$ we can define the matrixes $\mathbf{V}_s = [\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_J] \in \mathbb{C}^{N\times J}$ and $\mathbf{V}_n = [\mathbf{u}_{J+1}, \mathbf{u}_{J+2}, \ldots, \mathbf{u}_N] \in \mathbb{C}^{N\times(N-J)}$. With these notations, the EVD in (4.30) can be expressed as

$$\mathbf{R_{rr}} = \mathbf{V}_s\mathbf{\Upsilon}_s\mathbf{V}_s^H + \mathbf{V}_n\mathbf{\Upsilon}_n\mathbf{V}_n^H \tag{4.37}$$

It easy to see that $\mathcal{R}(\mathbf{\Phi}) = \mathcal{R}(\mathbf{V}_s)$, as a consequence the columns of the matrix $\mathbf{\Phi}$ and the signal eigenvectors span the same space so-called *signal subspace*. Instead the noise eigenvectors (the columns of $\mathbf{V}_n$) span the so-called *noise subspace* that is the orthogonal complement of the signal subspace. If the noise

subspace is the orthogonal complement of the signal subspace, the columns of $\mathbf{\Phi}$ are orthogonal to any vectors in the noise space, in fact from (4.32), recalling (4.34), we obtain

$$\mathbf{\Phi}\,\mathbf{\Phi}^H\,\mathbf{u}_j = \mathbf{0} \;\text{ if }\; j \in \{J+1, J+2, \ldots, N\}. \tag{4.38}$$

Nevertheless $\mathbf{\Phi}\,\mathbf{\Phi}^H\,\mathbf{u}_j$ is a linear combination (with coefficients equal to $\mathbf{\Phi}^H\,\mathbf{u}_j$) of the columns of the matrix $\mathbf{\Phi}$ that are linear independent because $\mathbf{\Phi}$ is full-column rank (assumption **(c1)**). Therefore any their linear combination is equal to zero if and only if the coefficients are equal to zero:

$$\mathbf{\Phi}^H\,\mathbf{u}_j = \mathbf{u}_j^H\,\mathbf{\Phi} = \mathbf{0} \;\text{ if }\; j \in \{J+1, J+2, \ldots, N\}. \tag{4.39}$$

In this way we have proved that

$$\text{Span}[\mathbf{\Phi}] \;=\; \text{Span}[\mathbf{V}_s] \;\perp\; \text{Span}[\mathbf{V}_n].$$

The blind Subspace method exploits this property, derived from the special structure of $\mathbf{R_{rr}}$, to estimate the channel parameters [61]. The equation (4.39), recalling (4.10) and (4.12), can also be expressed as

$$\mathbf{V}_n^H\,\boldsymbol{\phi}_j = \mathbf{V}_n^H\,\mathbf{C}_j\,\mathbf{g}_j = \mathbf{0}_{N-J}\,, \quad \forall j \in \{1, ..., J\}. \tag{4.40}$$

Assuming that the receiver has the only knowledge of the transmitted signature $\mathbf{c}_j$, the matrix $\mathbf{C}_j$ in (4.40) is known. Eq. (4.40) uniquely characterizes the channel coefficients for each user iff the following condition is satisfied: **(c4)** the null space of $\mathbf{V}_n^H\,\mathbf{C}_j$ has dimension one or, equivalently,[8] $\text{rank}(\mathbf{V}_n^H\,\mathbf{C}_j) = L_j - 1$. A discussion about condition **(c4)** is made in [61]. If condition **(c4)** is satisfied, then an arbitrary unit-norm vector $\mathbf{g}_j^{'} \in \mathbb{C}^{L_j}$ satisfies (4.40) iff $\mathbf{g}_j^{'} = e^{i\,\vartheta_j}\,\mathbf{g}_j$, with $\vartheta_j \in [0, 2\pi)$ and $\forall j \in \{1, 2, \ldots, J\}$. It is noteworthy that fulfillment of condition **(c4)** requires that the number of rows of the matrix $\mathbf{V}_n^H\,\mathbf{C}_j \in \mathbb{C}^{(N-J)\times L_j}$ must be greater than or equal to its number of columns, i.e., $N - J \geq L_j \iff J \leq N - L_j$, and, hence, from the point of view[9] of the jth user, the maximum number $J_{\text{max,L}} = N - L_j$ of users that can be supported by the system is smaller than the number N of users when the

[8] The dimension of the null space of $\mathbf{V}_n^H\,\mathbf{C}_j \in \mathbb{C}^{(N-J)\times L_j}$ is equal to $L_j - \text{rank}(\mathbf{V}_n^H\mathbf{C}_j)$.

[9] In order to meaningfully define the maximum number of users that can be supported by the system, we could consider the worst case, i.e., set L the maximum number of users that can be supported by the system is given by $L_{\max} \triangleq \max_{1\leq j\leq J} L_j$ as the maximum channel length, obtaining thus $J \leq (N - L_{\max})$.

channel is assumed to be perfectly known (see subsection 4.4). When $\mathbf{R_{rr}}$ (and hence $\mathbf{V}_n$) is estimated from a finite sample size, a channel estimate $\widehat{\mathbf{g}}_j$ can be obtained by solving (4.40) in the least-squares sense:

$$\begin{aligned}\widehat{\mathbf{g}}_j &= \underset{\mathbf{x}\in\mathbb{C}^{L_j}}{\operatorname{argmin}} \|\widehat{\mathbf{V}}_n^H\,\mathbf{C}_j\,\mathbf{x}\|^2 \\ &= \underset{\mathbf{x}\in\mathbb{C}^{L_j}}{\operatorname{argmin}} \left(\mathbf{x}^H\,\mathbf{C}_j^H\,\widehat{\mathbf{V}}_n\,\widehat{\mathbf{V}}_n^H\,\mathbf{C}_j\,\mathbf{x}\right), \quad \text{subject to } \|\mathbf{x}\|^2 = 1\,, \end{aligned} \tag{4.41}$$

where the matrix $\widehat{\mathbf{V}}_n \in \mathbb{C}^{N\times(N-J)}$ is the sample estimate of $\mathbf{V}_n$. The solution [56] of (4.41) is the eigenvector associated with the smallest eigenvalue of the matrix $\widehat{\boldsymbol{\mathcal{Q}}}_{j,\mathrm{L}} \stackrel{\triangle}{=} \mathbf{C}_j^H\,\widehat{\mathbf{V}}_n\,\widehat{\mathbf{V}}_n^H\,\mathbf{C}_j \in \mathbb{C}^{L_j\times L_j}$.

We can follow the same analysis also for the WL receiver. In this case under assumptions **(a1)** and **(a2)**, the matrix $\mathbf{R_{zz}}$ assumes the form

$$\mathbf{R_{zz}} = \mathbf{H}\,\mathbf{H}^H + \sigma_v^2\,\mathbf{I}_{2N}. \tag{4.42}$$

Moreover, by virtue of conditions **(c1)** and **(c2)**, the matrix $\mathbf{H}\,\mathbf{H}^H$ has only J nonzero eigenvalues $\lambda_1 \geq \lambda_2 \cdots \geq \lambda_J > 0$. Resorting to the eigenvalue decomposition, the matrix $\mathbf{R_{zz}}$ can be also expressed as

$$\mathbf{R_{zz}} = \mathbf{U}_s\boldsymbol{\Lambda}_s\mathbf{U}_s^H + \mathbf{U}_n\boldsymbol{\Lambda}_n\mathbf{U}_n^H, \tag{4.43}$$

where $\mathbf{U}_s \in \mathbb{C}^{2N\times J}$ collects the eigenvectors associated with the J largest eigenvalues of $\mathbf{R_{zz}}$, whose columns span the *signal subspace*, i.e., the subspace $\mathcal{R}(\mathbf{H})$, $\boldsymbol{\Lambda}_s \stackrel{\triangle}{=} \mathrm{diag}(\lambda_1+\sigma_v^2, \lambda_2+\sigma_v^2, \ldots, \lambda_J+\sigma_v^2) \in \mathbb{R}^{J\times J}$, $\boldsymbol{\Lambda}_n = \sigma_v^2\,\mathbf{I}_{2N-J} \in \mathbb{R}^{2N-J\times 2N-J}$, and, finally, $\mathbf{U}_n \in \mathbb{C}^{2N\times(2N-J)}$ collects the eigenvectors associated with the eigenvalue σ_v^2, whose columns span the *noise subspace*, i.e., the subspace $\mathcal{R}^{\perp}(\mathbf{H})$ in $\mathbb{C}^{2N}$. Also in this case, blind subspace-based CE can be accomplished by exploiting the orthogonality between the signal space $\mathcal{R}(\mathbf{H})$ and the noise subspace $\mathcal{R}^{\perp}(\mathbf{H}) \equiv \mathcal{R}(\mathbf{U}_n)$, obtaining thus, $\forall j \in \{1, ..., J\}$,

$$\mathbf{U}_n^H\,\mathbf{h}_j = \mathbf{0}_{2N-J}\,. \tag{4.44}$$

As regards $\mathbf{h}_j$, we preliminarily observe that, according to (4.10) and (4.12), the jth column $\boldsymbol{\phi}_j$ of the matrix $\boldsymbol{\Phi}$ assumes the form

$$\boldsymbol{\phi}_j = \alpha_j\,\mathbf{C}_j\,\mathbf{g}_j \tag{4.45}$$

and, consequently, one has

$$\mathbf{h}_j = \begin{bmatrix} \boldsymbol{\phi}_j \\ \boldsymbol{\phi}_j^* \end{bmatrix} = \alpha_j \underbrace{\begin{bmatrix} \mathbf{C}_j & \mathbf{O}_{N\times L_j} \\ \mathbf{O}_{N\times L_j} & \mathbf{C}_j^* \end{bmatrix}}_{\boldsymbol{\mathcal{C}}_j \in \mathbb{C}^{2N\times 2L_j}} \begin{bmatrix} \mathbf{g}_j \\ \mathbf{g}_j^* \end{bmatrix}$$

$$= \underbrace{\alpha_j \sqrt{2}}_{\widetilde{\alpha}_j} \boldsymbol{\mathcal{C}}_j \underbrace{\frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{I}_{L_j} & i\,\mathbf{I}_{L_j} \\ \mathbf{I}_{L_j} & -i\,\mathbf{I}_{L_j} \end{bmatrix}}_{\mathbf{T}_j \in \mathbb{C}^{2L_j\times 2L_j}} \underbrace{\begin{bmatrix} \mathbf{g}_{j,\mathrm{R}} \\ \mathbf{g}_{j,\mathrm{I}} \end{bmatrix}}_{\boldsymbol{\varrho}_j \in \mathbb{R}^{2L_j}}$$

$$= \widetilde{\alpha}_j \boldsymbol{\mathcal{C}}_j \mathbf{T}_j \boldsymbol{\varrho}_j \,, \tag{4.46}$$

where $\mathbf{T}_j$ is a *unitary* matrix, i.e., $\mathbf{T}_j \mathbf{T}_j^H = \mathbf{T}_j^H \mathbf{T}_j = \mathbf{I}_{2L_j}$. Substituting (4.46) in (4.44), we obtain:

$$\mathbf{U}_n^H \mathbf{h}_j = \mathbf{U}_n^H \boldsymbol{\mathcal{C}}_j \mathbf{T}_j \boldsymbol{\varrho}_j = \mathbf{0}_{2N-J}, \;\; \forall j \in \{1, ..., J\}. \tag{4.47}$$

The unknown vector $\boldsymbol{\varrho}_j$ can be obtained as the solution of the linear system (4.47), provided that this system *uniquely* characterizes the channel coefficients for each user, i.e., an arbitrary unit-norm vector $\mathbf{g}_j^{'} \in \mathbb{C}^{L_j}$ (with corresponding $\boldsymbol{\varrho}_j^{'} \in \mathbb{R}^{2L_j}$), satisfies (4.47) if and only if (iff) $\mathbf{g}_j^{'} = e^{i\,\psi_j}\,\mathbf{g}_j$, with $\psi_j \in [0, 2\pi)$ and $\forall j \in \{1, 2, \ldots, J\}$. It is clear that (4.47) has a unique solution (up to a scaling factor) iff the following condition is satisfied: **(c3)** the null space of $\mathbf{U}_n^H \boldsymbol{\mathcal{C}}_j \mathbf{T}_j$ has dimension one or, equivalently,[10] $\mathrm{rank}(\mathbf{U}_n^H \boldsymbol{\mathcal{C}}_j) = 2\,L_j - 1$. A reformulation of condition **(c3)** is given in [66]. It can be readily proven that, under **(c3)**, the following two statements are equivalent: (i) the unit-norm vector $\mathbf{g}_j^{'} \in \mathbb{C}^{L_j}$ is a solution of (4.47); (ii) $\mathbf{g}_j^{'} = \pm\mathbf{g}_j$, i.e., $\psi_j = n\,\pi$, with $n \in \mathbb{Z}$. In other words, differently from conventional subspace-based multiuser CE [61], where the estimated channel might differ from the true one by an unknown *rotation* $e^{i\,\psi_j}$, in generalized subspace-based CE based on (4.47) the residual channel ambiguity is limited to a possible *sign inversion*. It is important to observe that condition **(c3)** necessarily imposes that the number of rows of the matrix $\mathbf{U}_n^H \boldsymbol{\mathcal{C}}_j \mathbf{T}_j \in \mathbb{C}^{(2N-J)\times 2L_j}$ be greater than or equal to its number of columns, i.e., $2N - J \geq 2L_j \iff J \leq 2(N - L_j)$. Thereby, it follows that, from the point of view[11] of the jth user, the maximum number

[10] The dimension of the null space of $\mathbf{U}_n^H \boldsymbol{\mathcal{C}}_j \mathbf{T}_j \in \mathbb{C}^{(2N-J)\times 2L_j}$ is equal to $2\,L_j - \mathrm{rank}(\mathbf{U}_n^H \boldsymbol{\mathcal{C}}_j \mathbf{T}_j)$. Moreover, since $\mathbf{T}_j$ is unitary and, hence, nonsingular, it results that $\mathrm{rank}(\mathbf{U}_n^H \boldsymbol{\mathcal{C}}_j \mathbf{T}_j) = \mathrm{rank}(\mathbf{U}_n^H \boldsymbol{\mathcal{C}}_j)$.

[11] Following footnote 9 the maximum number of users that can be supported by the system is given by $J \leq 2(N - L_{\max})$.

$J_{\text{max,WL}} = 2(N - L_j)$ of users supported by the system is smaller than the maximum number $2\,N$ of users when the channel is assumed to be perfectly known (see subsection 4.4). In the following, we assume that condition **(c3)** is satisfied. In practice, however, eq. (4.47) cannot be satisfied exactly when $\mathbf{R_{zz}}$ (and hence $\mathbf{U}_n$) is estimated from a finite sample size. In this case, a channel estimate $\widehat{\boldsymbol{\varrho}}_j \stackrel{\triangle}{=} [\widehat{\mathbf{g}}_{j,\mathrm{R}}^T, (\widehat{\mathbf{g}}_{j,\mathrm{I}}^T]^T$ can still be obtained by solving (4.47) in the least-squares sense, that is, as

$$\begin{aligned}\widehat{\boldsymbol{\varrho}}_j &= \underset{\mathbf{x}\in\mathbb{R}^{2L_j}}{\operatorname{argmin}} \|\widehat{\mathbf{U}}_n^H\, \boldsymbol{\mathcal{C}}_j\, \mathbf{T}_j\, \mathbf{x}\|^2 \\ &= \underset{\mathbf{x}\in\mathbb{R}^{2L_j}}{\operatorname{argmin}} \left(\mathbf{x}^H\, \mathbf{T}_j^H\, \boldsymbol{\mathcal{C}}_j^H\, \widehat{\mathbf{U}}_n\, \widehat{\mathbf{U}}_n^H\, \boldsymbol{\mathcal{C}}_j\, \mathbf{T}_j\, \mathbf{x}\right), \quad \text{subject to } \|\mathbf{x}\|^2 = 1,\end{aligned} \tag{4.48}$$

whose solution [56] is given by the eigenvector associated with the smallest eigenvalue of the matrix $\mathbf{T}_j^H\, \widehat{\boldsymbol{\mathcal{Q}}}_{j,\text{WL}}\, \mathbf{T}_j \in \mathbb{C}^{2L_j \times 2L_j}$, with $\widehat{\boldsymbol{\mathcal{Q}}}_{j,\text{WL}} \stackrel{\triangle}{=} \boldsymbol{\mathcal{C}}_j^H\, \widehat{\mathbf{U}}_n\, \widehat{\mathbf{U}}_n^H\, \boldsymbol{\mathcal{C}}_j \in \mathbb{C}^{2L_j \times 2L_j}$.

4.4 Ideal Performance of L-MOE and WL-MOE receivers

In this section, following the analysis that we have developed in [12], we compare the SINR performances of the *ideal* WL-MOE and the L-MOE receivers, i.e., those receivers whose synthesis is based on perfect knowledge of both the SOS of the received signal and the channel impulse response. Preliminarily, we observe that, recently, with reference to DS-CDMA systems employing BPSK modulation, a few contributions addressing the theoretical performance analysis of WL-MUD techniques appeared in the literature. In [54], the asymptotic (in the number of users) performance analysis of the WL decorrelating and WL-MMSE receivers was carried out, by extending to the WL framework classical analysis tools already developed by Tse and Hanly [70] for L-MUD techniques (a similar study was proposed in [71]). A non-asymptotic performance analysis was instead considered in [72], which provides an algebraic proof that WL-MUD receivers outperform L-MUD ones, and explicitly assesses the expected performance gain in the two-users case. The common conclusion of these studies (see also [73]) is that the performance advantage of WL-MUD receivers over L-MUD ones is twofold: *the input SNR is doubled and the number of effective interferers is halved.* As a consequence, for

a fixed processing gain N, the number of users that can be accommodated by a DS-CDMA system employing WL-MUD is doubled [54, 71, 72] compared to L-MUD. In other words, unlike L-MUD, WL-MUD can be successfully employed not only when the number of users J is smaller than or equal to N (*underloaded system*), but also when $N < J \leq 2\,N$ (*overloaded system*). However, none of the aforementioned papers on WL-MUD carried out a detailed study of the conditions on channel and codes that assure perfect MAI suppression in absence of noise. Thus, in this subsection, following our paper [12], we provide conditions on the spreading codes, which guarantee complete MAI rejection for WL-MOE in both underloaded and overloaded downlink configurations. We will show that in the limiting case of vanishingly small noise, i.e., as $\sigma_v^2 \rightarrow 0$, the performance comparison between the L-MOE and WL-MOE receivers heavily depends on the rank properties of $\mathbf{\Phi}$ and $\mathbf{H}$, respectively.

In order to carry out a meaningful performance comparison between linear and WL receivers, we evaluate for both receivers the SINR *after* the $\mathrm{Re}[\cdot]$ block, given by (4.24). Since the WL-MOE receiver maximizes such a SINR (see Lemma 1), one simply has:

$$\mathrm{SINR}_{j,\text{WL-MOE}} \stackrel{\triangle}{=} \mathrm{SINR}(\mathbf{f}_{j,\text{WL-MOE}}) = \mathbf{h}_j^H\, \mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}^{-1}\, \mathbf{h}_j\,. \tag{4.49}$$

Instead, observe that evaluating the SINR given by (4.24) for the L-MOE receiver leads to a result generally different from (4.18). By observing that the L-MOE receiver can be viewed as a WL receiver with augmented weight vector $\mathbf{f}_{j,\text{L-MOE}} \stackrel{\triangle}{=} [\mathbf{w}_{j,\text{L-MOE}}^T, \mathbf{0}_N^T]^T$, recalling that $\mathbf{w}_{j,\text{L-MOE}}^H\, \boldsymbol{\phi}_j = 1$, and applying the straightforward identity $\mathrm{Re}^2[z] = \frac{1}{2}\{|z|^2 + \mathrm{Re}[z^{\,2}]\}$, $\forall z \in \mathbb{C}$, the SINR (4.24) for the L-MOE receiver can be written as

$$\begin{aligned}\mathrm{SINR}_{j,\text{L-MOE}} &\stackrel{\triangle}{=} \mathrm{SINR}(\mathbf{f}_{j,\text{L-MOE}}) = \frac{1}{\mathrm{E}\{\mathrm{Re}^2[\mathbf{w}_{j,\text{L-MOE}}^H \mathbf{p}_j(k)]\}} \\ &= \frac{2}{\mathbf{w}_{j,\text{L-MOE}}^H \mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}\, \mathbf{w}_{j,\text{L-MOE}} + \mathrm{Re}[\mathbf{w}_{j,\text{L-MOE}}^H \mathbf{R}_{\mathbf{p}_j\mathbf{p}_j^*}\, \mathbf{w}_{j,\text{L-MOE}}^*]}\,. \end{aligned} \tag{4.50}$$

On one hand, since the WL-MOE is a maximum-SINR receiver, it results that $\mathrm{SINR}_{j,\text{L-MOE}} \leq \mathrm{SINR}_{j,\text{WL-MOE}}$. On the other hand, since $\mathrm{Re}^2[z] \leq |z|^2$, $\forall z \in \mathbb{C}$, accounting for (4.18), one has $\mathrm{SINR}_{j,\text{L-MOE}} \geq \overline{\mathrm{SINR}}_{j,\max}$. Overall, we maintain that

$$\mathrm{SINR}_{j,\text{WL-MOE}} \geq \mathrm{SINR}_{j,\text{L-MOE}} \geq \overline{\mathrm{SINR}}_{j,\max}\,. \tag{4.51}$$

Although the first inequality in (4.51) concisely states that the performance of the WL-MOE receiver is not worse than that of its linear counterpart, it does not allow us to quantify the relative performance gain. Indeed, no clear insight on the performance comparison between the WL-MOE and L-MOE receivers can be drawn out from the SINR formulas (4.49) and (4.50). To overcome this conceptual difficulty, we carry out in the next subsection the performance comparison in the high-SNR regime, by deriving the analytical expressions of $\text{SINR}_{j,\text{WL-MOE}}$ and $\text{SINR}_{j,\text{L-MOE}}$ as the noise variance σ_v^2 approaches zero. It should be observed that, more generally, the results reported in Subsection 4.4.1 turn out to be useful in all those situations wherein the DS-CDMA signal dominates the background noise, which is a common occurrence in many practical environments.

4.4.1 Analysis in the High-SNR Regime

The discussion carried out in this subsection is mainly based on some mathematical results whose proofs are reported in Appendix B.2.

As we have recalled in the note 1 in the section 5.2 the code vector $\mathbf{c}_j$ accounts also for possible precoding phases whose role in downlink is discussed in this section. Therefore we modify the system model (4.10) introducing these precoding phases:

$$\mathbf{r}(k) = \sum_{j=1}^{J} \alpha_j \, e^{i\,\theta_j} \boldsymbol{\psi}_j \, b_j(k) + \mathbf{v}(k) = \boldsymbol{\Psi}\mathbf{A}\boldsymbol{\Theta}\,\mathbf{b}(k) + \mathbf{v}(k) = \boldsymbol{\Phi}\,\mathbf{b}(k) + \mathbf{v}(k)\,, \tag{4.52}$$

where, with reference to the jth user, $\theta_j \in [0, 2\,\pi)$ is a precoding phase which is deliberately introduced at the transmitter and whose role will be clear in the sequel. Moreover, in (4.52), we have defined $\boldsymbol{\Theta} \stackrel{\triangle}{=} \text{diag}(e^{i\,\theta_1}, e^{i\,\theta_2}, \ldots, e^{i\,\theta_J}) \in \mathbb{C}^{J\times J}$, and consequently we have redefine with $\boldsymbol{\Phi}$ the matrix $\boldsymbol{\Phi} \stackrel{\triangle}{=} \boldsymbol{\Psi}\,\mathbf{A}\,\boldsymbol{\Theta} \in \mathbb{C}^{N\times J}$. The other parameters in (4.52) are the same of those in (4.10).

First of all, let us start from linear processing. In the Appendix B.2 it is shown that, in the high-SNR regime, the L-MOE receiver is able to achieve perfect MAI suppression for *each* active user, that is, $\lim_{\sigma_v^2\to 0} \text{SINR}_{j,\text{L-MOE}} = \lim_{\sigma_v^2\to 0} \overline{\text{SINR}}_{j,\text{max}} = +\infty$, $\forall j \in \{1, 2, \ldots, J\}$, if and only if (iff) the matrix $\boldsymbol{\Phi}$ is full-column rank, i.e., $\text{rank}(\boldsymbol{\Phi}) = J$. Moreover, in such a case, it results that

$$\lim_{\sigma_v^2\to 0} \frac{\text{SINR}_{j,\text{L-MOE}}}{\overline{\text{SINR}}_{j,\text{max}}} = 2\,, \quad \forall j \in \{1, 2, \ldots, J\}\,, \tag{4.53}$$

which shows that, as intuitively expected, since the Re[·] block in Fig. 4.1 discards one-half of the noise-plus-MAI power in $y_j(k)$, $\text{SINR}_{j,\text{L-MOE}}$ is asymptotically greater than $\overline{\text{SINR}}_{j,\max}$ of exactly 3 dB. Note that this simple result holds only when $\text{rank}(\mathbf{\Phi}) = J$. If the matrix $\mathbf{\Phi}$ is not full-column rank, the L-MOE receiver is unable to perfectly suppress the MAI, even in the absence of noise; in this case, both $\overline{\text{SINR}}_{j,\max}$ and $\text{SINR}_{j,\text{L-MOE}}$ assume finite values, which depend on $\boldsymbol{\phi}_j$ and the eigenstructure of the MAI autocorrelation matrix $\overline{\mathbf{\Phi}}_j\,\overline{\mathbf{\Phi}}_j^H$. Therefore, the assumption $\text{rank}(\mathbf{\Phi}) = J$ is crucial and deserves a brief comment. By virtue of nonsingularity of the diagonal matrices $\mathbf{A}$ and $\mathbf{\Theta}$, it follows that $\text{rank}(\mathbf{\Phi}) = \text{rank}(\mathbf{\Psi}\,\mathbf{A}\,\mathbf{\Theta}) = \text{rank}(\mathbf{\Psi})$. Henceforth, the matrix $\mathbf{\Phi}$ is full-column rank iff the signatures $\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, \ldots, \boldsymbol{\psi}_J$ are linearly independent, a condition which can be fulfilled only if the number of users J is smaller than or equal to the processing gain N (underloaded systems). It is noteworthy that the linear independence of the signatures $\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, \ldots, \boldsymbol{\psi}_J$ depends on both the spreading codes and the channel impulse responses of all the active users. Thus, in general, it is difficult to give easily interpretable conditions assuring that $\mathbf{\Psi}$ is full-column rank. A substantial simplification occurs in the downlink, wherein all the user signals propagate through a common multipath channel, i.e., $g_j(n) = g(n)$, with order $L_j = L_g$, for each user. In this case, the signature $\boldsymbol{\psi}_j$ given by (4.11) becomes $\boldsymbol{\psi}_j = \mathbf{G}\,\mathbf{c}_j$, where the common Toeplitz channel matrix $\mathbf{G} = \mathbf{G}_j$ turns out to be nonsingular under the mild assumption that $g(0) \neq 0$, which is assumed to hold hereinafter. Accounting for this model, the matrix $\mathbf{\Psi}$ becomes

$$\mathbf{\Psi} = \mathbf{G}\,\underbrace{[\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_J]}_{\mathbf{C}\in\mathbb{C}^{N\times J}} = \mathbf{G}\,\mathbf{C}\,, \tag{4.54}$$

which, by virtue of nonsingularity of $\mathbf{G}$, implies that $\text{rank}(\mathbf{\Phi}) = \text{rank}(\mathbf{\Psi}) = \text{rank}(\mathbf{C})$. Consequently, in the downlink scenario, the linear independence of the spreading vectors $\mathbf{c}_1$, $\mathbf{c}_2, \ldots, \mathbf{c}_J$ is a necessary and sufficient condition for assuring the full-column rank property of $\mathbf{\Phi}$ and, hence, allowing the L-MOE receiver to completely reject the MAI in the high-SNR region. Let us focus attention on the performance comparison between the L-MOE and WL-MOE receivers. As a first result, it is shown in Appendix B.2 that, if $\mathbf{\Phi}$ (or, equivalently, $\mathbf{\Psi}$) is full-column rank, then

$$\lim_{\sigma_v^2\to 0} \frac{\text{SINR}_{j,\text{WL-MOE}}}{\text{SINR}_{j,\text{L-MOE}}} = \frac{\|\boldsymbol{\phi}_j\|^2 - \text{Re}[\boldsymbol{\phi}_j^H\overline{\mathbf{\Phi}}_j]\{\text{Re}[\overline{\mathbf{\Phi}}_j^H\overline{\mathbf{\Phi}}_j]\}^{-1}\text{Re}[\overline{\mathbf{\Phi}}_j^H\boldsymbol{\phi}_j]}{\|\boldsymbol{\phi}_j\|^2 - \boldsymbol{\phi}_j^H\overline{\mathbf{\Phi}}_j(\overline{\mathbf{\Phi}}_j^H\overline{\mathbf{\Phi}}_j)^{-1}\overline{\mathbf{\Phi}}_j^H\boldsymbol{\phi}_j}\,, \tag{4.55}$$

which, in addition to (4.51), evidences that, since $\lim_{\sigma_v^2 \to 0} \text{SINR}_{j,\text{L-MOE}} = +\infty$ when $\text{rank}(\mathbf{\Phi}) = J$, the WL-MOE receiver also suppresses the MAI exactly in the high-SNR regime, i.e., $\lim_{\sigma_v^2 \to 0} \text{SINR}_{j,\text{WL-MOE}} = +\infty$, $\forall j \in \{1, 2, \ldots, J\}$. Remarkably, it is apparent from (4.55) that, if

$$\text{Re}[\boldsymbol{\phi}_j^H \overline{\mathbf{\Phi}}_j] \, \{\text{Re}[\overline{\mathbf{\Phi}}_j^H \overline{\mathbf{\Phi}}_j]\}^{-1} \, \text{Re}[\overline{\mathbf{\Phi}}_j^H \boldsymbol{\phi}_j] = \boldsymbol{\phi}_j^H \overline{\mathbf{\Phi}}_j \, (\overline{\mathbf{\Phi}}_j^H \overline{\mathbf{\Phi}}_j)^{-1} \, \overline{\mathbf{\Phi}}_j^H \boldsymbol{\phi}_j \,, \quad (4.56)$$

we have:

$$\lim_{\sigma_v^2 \to 0} \frac{\text{SINR}_{j,\text{WL-MOE}}}{\text{SINR}_{j,\text{L-MOE}}} = 1 \,, \quad (4.57)$$

which renders the L-MOE and WL-MOE receivers perfectly equivalent in terms of SINR, as $\sigma_v^2 \to 0$. In other words, if $\mathbf{\Phi}$ is full-column rank (as may be the case in underloaded systems) and condition (4.56) is fulfilled, WL processing does not improve upon conventional linear processing in the high-SNR region. It is interesting to observe that, for instance, condition (4.56) is trivially satisfied if $\boldsymbol{\phi}_j$ and $\overline{\mathbf{\Phi}}_j$ are real (i.e., matrix $\mathbf{\Phi}$ is real), or when the user signatures are orthogonal[12], i.e., $\boldsymbol{\psi}_{j_1}^H \boldsymbol{\psi}_{j_2} = 0$, $\forall j_1 \neq j_2 \in \{1, 2, \ldots, J\}$, independently of matrices $\mathbf{A}$ and $\mathbf{\Theta}$ [see (4.52)]. To gain further insight about (4.55), we consider the two-users case (i.e., $J = 2$), and, without loss of generality, we assume that the desired user is the first one (i.e., $j = 1$). In this case, eq. (4.55) simplifies to

$$\lim_{\sigma_v^2 \to 0} \frac{\text{SINR}_{1,\text{WL-MOE}}}{\text{SINR}_{1,\text{L-MOE}}} = \frac{1 - |\rho|^2 \cos^2(\Delta\theta - \angle\rho)}{1 - |\rho|^2} \,, \quad (4.58)$$

which suggests that the performance advantage of the WL-MOE receiver over the L-MOE one depends on the magnitude $|\rho|$ and phase $\angle\rho$ of the correlation coefficient $\rho \stackrel{\triangle}{=} \boldsymbol{\psi}_1^H \boldsymbol{\psi}_2$ between the two signatures $\boldsymbol{\psi}_1$ and $\boldsymbol{\psi}_2$, as well as on the phase difference $\Delta\theta \stackrel{\triangle}{=} \theta_1 - \theta_2$. This is in accordance with the results derived in [72] in terms of near-far resistance. Specifically, for a given value of $0 < |\rho| < 1$, the largest performance gap between WL-MOE and L-MOE receivers is obtained when $\Delta\theta - \angle\rho = \pi/2 + h\,\pi$, with $h \in \mathbb{Z}$, whereas the two receivers achieve the same performance when $\Delta\theta - \angle\rho = h\,\pi$, independently of the value of $|\rho|$. On the other hand, for a given value of $\Delta\theta - \angle\rho \neq h\,\pi$, the performance gain of the WL-MOE receiver over the L-MOE one increases

[12] As a matter of fact, if the user signatures are orthogonal, under assumptions **(a1)** and **(a2)**, the single-user detector, which simply matches the received vector $\mathbf{r}(k)$ to $\boldsymbol{\phi}_j$, is indeed the optimal (in the minimum-error-probability sense) receiver.

without bounds, as the magnitude of ρ approaches unity, i.e., the user signatures are maximally correlated.

As a second result, it is evidenced in Appendix B.2 that, contrary to the L-MOE receiver, the WL-MOE one is able to ensure perfect MAI suppression in the high-SNR regime, even when the number of users J exceeds the processing gain N (overloaded systems). Indeed, it is shown that, more generally, $\lim_{\sigma_v^2 \to 0} \mathrm{SINR}_{j,\text{WL-MOE}} = +\infty$, $\forall j \in \{1, 2, \ldots, J\}$, iff $\mathbf{H}$ is full-column rank. If $\mathbf{\Phi}$ is full-column rank (a condition that can hold only when the system is underloaded), then $\mathbf{H}$ is full-column rank, too. However, the matrix $\mathbf{H}$ can be full-column rank even when $N < J \leq 2\,N$, wherein $\mathbf{\Phi}$ is structurally rank-deficient; in this overloaded environment, it results that

$$\lim_{\sigma_v^2 \to 0} \frac{\mathrm{SINR}_{j,\text{WL-MOE}}}{\mathrm{SINR}_{j,\text{L-MOE}}} = +\infty\,, \quad \forall j \in \{1, 2, \ldots, J\}\,. \tag{4.59}$$

In other words, provided that $\mathrm{rank}(\mathbf{H}) = J$, the performance gap between the WL-MOE and L-MOE receivers becomes arbitrarily large for vanishingly small noise, when $N < J \leq 2\,N$. This fact strongly motivates us to provide conditions assuring that $\mathbf{H}$ be full-column rank in overloaded scenarios. To this aim, we provide the following Theorem, by focusing attention directly on the downlink scenario in an effort to give simple and insightful conditions.

Theorem 4.1 *When $N < J \leq 2\,N$, the code matrix can be decomposed as $\mathbf{C} = \mathbf{C}_{left}\,[\mathbf{I}_N, \mathbf{\Pi}]$, where $\mathbf{C}_{left} \stackrel{\triangle}{=} [\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_N] \in \mathbb{C}^{N \times N}$ is non-singular and $\mathbf{\Pi} \in \mathbb{C}^{N \times (J-N)}$ is a tall matrix. In this overloaded scenario, under the assumption that $\mathbf{\Psi}$ exhibits the form given by (4.54), the matrix $\mathbf{H}$ is full-column rank iff $\mathbf{\Pi}^* - (\mathbf{\Theta}_1^2)^*\,\mathbf{\Pi}\,\mathbf{\Theta}_2^2 \in \mathbb{C}^{N \times (J-N)}$ is full-column rank, where $\mathbf{\Theta}_1 \stackrel{\triangle}{=} diag(e^{i\,\theta_1}, e^{i\,\theta_2}, \ldots, e^{i\,\theta_N}) \in \mathbb{C}^{N \times N}$ and $\mathbf{\Theta}_2 \stackrel{\triangle}{=} diag(e^{i\,\theta_{N+1}}, e^{i\,\theta_{N+2}}, \ldots, e^{i\,\theta_J}) \in \mathbb{C}^{(J-N) \times (J-N)}$.*

Proof. See Appendix B.3.

Theorem 4.1 deserves some interesting comments, aimed at clarifying in particular the role of the precoding phases in (4.52), which are at the designer's disposal. First of all, it is apparent that the full-column rank property of $\mathbf{H}$ does not depend on the channel impulse response[13], but depends on both the spreading codes of all the active users and their precoding phases

[13] In the uplink scenario, the full-column rank property of $\mathbf{H}$ and, thus, the performance of the WL-MOE receiver, depends not only on the precoding phases, but also on the channel impulse responses of all the active users.

$\theta_1, \theta_2, \ldots, \theta_J$. To this respect, it is interesting to investigate how such phases influence the full-column rank property of $\mathbf{H}$ in overloaded systems, focusing attention to the case wherein Walsh-Hadamard (WH) spreading codes are employed. To do this, without loss of generality, assume that $\mathbf{c}_{N+j} = \mathbf{c}_j$, for $j \in \{1, 2, \ldots, J-N\}$, and let $\mathbf{C}_{\text{left}}$ denote the common Hadamard matrix of order N. In this case, it is easily verified that $\mathbf{\Pi} = [\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_{J-N}]$, with $\mathbf{e}_j$ denoting the jth column of $\mathbf{I}_N$. Thus, if WH spreading vectors are used, the matrix $\mathbf{\Pi}$ is real-valued (i.e., $\mathbf{\Pi} = \mathbf{\Pi}^*$) and, moreover, one has $(\mathbf{\Theta}_1^2)^* \mathbf{\Pi} = \mathbf{\Pi} (\mathbf{\Theta}_{1,\text{red}}^2)^*$, where $\mathbf{\Theta}_{1,\text{red}} \stackrel{\triangle}{=} \text{diag}(e^{i\theta_1}, e^{i\theta_2}, \ldots, e^{i\theta_{J-N}}) \in \mathbb{C}^{(J-N)\times(J-N)}$. In light of these observations, by additionally remembering that $\mathbf{\Pi}$ is full-column rank, it follows that $\text{rank}[\mathbf{\Pi}^* - (\mathbf{\Theta}_1^2)^* \mathbf{\Pi} \mathbf{\Theta}_2^2] = \text{rank}\{\mathbf{\Pi} [\mathbf{I}_{J-N} - (\mathbf{\Theta}_{1,\text{red}}^2)^* \mathbf{\Theta}_2^2]\} = \text{rank}[\mathbf{I}_{J-N} - (\mathbf{\Theta}_{1,\text{red}}^2)^* \mathbf{\Theta}_2^2]$. Since the matrix $\mathbf{I}_{J-N} - (\mathbf{\Theta}_{1,\text{red}}^2)^* \mathbf{\Theta}_2^2$ is diagonal with diagonal entries $1 - e^{i\,2\,(\theta_{N+j} - \theta_j)}$, $\forall j \in \{1, 2, \ldots, J-N\}$, by virtue of Theorem 4.1, it can be stated that, when $N < J \leq 2\,N$, the augmented matrix $\mathbf{H}$ is full-column rank iff

$$\theta_{N+j} - \theta_j \neq h\,\pi, \;\; \forall j \in \{1, 2, \ldots, J-N\} \text{ and } h \in \mathbb{Z}\,. \tag{4.60}$$

As an immediate implication of (4.60), it is worth pointing out that, if no precoding is performed at the transmitter, i.e., $\theta_1 = \theta_2 = \cdots = \theta_J$, and common WH spreading codes are employed, the WL-MOE receiver is unable to achieve perfect MAI suppression in overloaded systems, even in the absence of noise. Henceforth, in order to allow WL-MUD to successfully work in an overloaded downlink, while employing WH spreading sequences, incorporation of precoding phases is crucial. This is the reason which motivates to introduce the phases $\theta_1, \theta_2, \ldots, \theta_J$ in (4.10). It is worthwhile to observe that condition (4.60) does not uniquely specify the precoding phases and, thus, different choices can be pursued. To corroborate the previous considerations, let us provide a numerical example.

Example 4.1 *:* Consider a DS-CDMA downlink with $\alpha_1 = \alpha_2 = \cdots = \alpha_J = 1$ and processing gain $N = 16$, and without loss of generality, assume that the desired user is the first one (i.e., $j = 1$). The SNR, which is defined as $1/\sigma_v^2$, is set to 15 dB, and the signatures are generated according to (4.54). The system uses unit-norm WH vector codes and operates over a channel of order $L_g = 5$, whose taps $g(0), g(1), \ldots, g(5)$ are modeled as i.i.d. complex proper zero-mean Gaussian random variables, normalized so that $\|\boldsymbol{\psi}_j\|^2 = 1$, $\forall j \in \{1, 2, \ldots, J\}$. Fig. 4.2 reports the ideal SINR performance of the WL-MOE receiver as a function of the number of users J, ranging from an underloaded

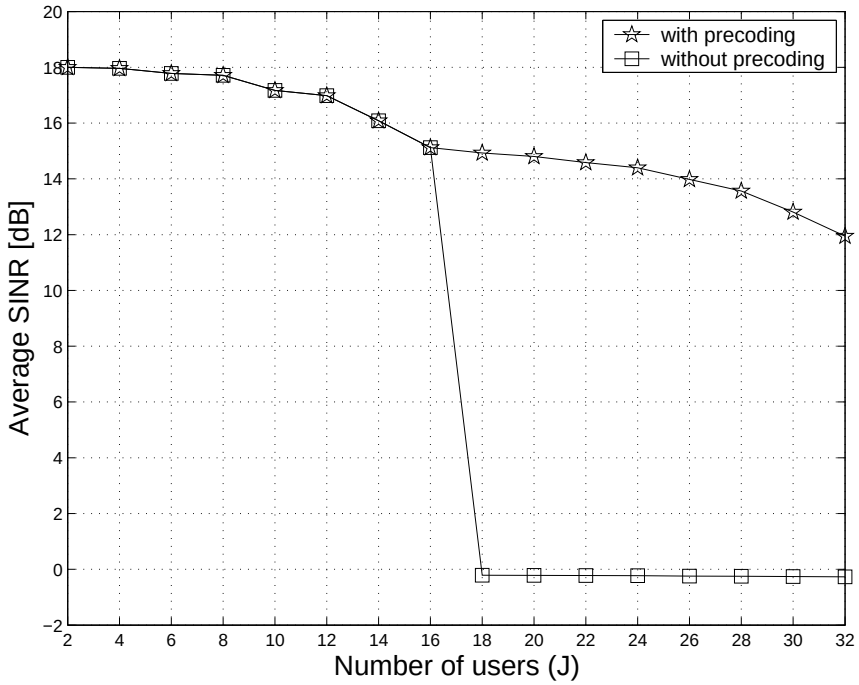

Figure 4.2: Average SINR values of the WL-MOE receiver versus J for different precoding techniques (SNR $= 15$ dB).

($1 < J \leq N$) system to an overloaded ($N < J \leq 2\,N$) one. Specifically, we report $\text{SINR}_{1,\text{WL-MOE}}$ [see (4.49)] in two different situations: in the former one, there is no precoding at the transmitter, i.e., $\theta_1 = \theta_2 = \cdots = \theta_J = 0$ (referred to as "without precoding"); in the latter one, we use a precoding strategy fulfilling (4.60), by setting $\theta_1 = \theta_2 = \cdots = \theta_N = 0$ and $\theta_{N+1} = \theta_{N+2} = \cdots = \theta_{2N} = \pi/4$ (referred to as "with precoding"). The results of Fig. 4.2 are obtained by carrying out 10^4 independent Monte Carlo trials, with each run using only a different channel realization. It can be observed that, if WH spreading sequences are employed and condition (4.60) is not accounted for, the WL-MOE receiver does not work at all, when the system becomes overloaded. In contrast, the proposed precoding strategy allows the WL-MOE receiver to achieve satisfactory performances even when $N < J \leq 2\,N$.

4.5 Finite-sample theoretical performance comparison between L-MOE and WL-MOE receivers

In this section we provide the finite-sample theoretical performance analysis of the L-MOE and WL-MOE receivers and in the known-channel case (see subsection 4.5.1) and in unknown-channel case (see subsection 4.5.3), following our papers [12, 13] and [14] respectively.

As we have note in the subsection 4.3.1, the ideal implementation of the L-MOE and WL-MOE receivers requires perfect knowledge of two quantities: the *autocorrelation matrix* (ACM) of the received signal, and the *received signature* of each user to be demodulated. These two quantities can be estimated in practice from a finite number of samples at the receiver. In partuculary, due to the effects of the unknown channel response, the received signature is a distorted version of the transmitted one, making channel estimation (CE) a necessary step to implement both the L- and WL-MOE receivers. A theoretical performance analysis of the data-aided WL-MMSE and WL-MOE receivers was provided in [74], when the receivers are adaptively implemented by means of the least-mean square (LMS) algorithm, by evaluating the output signal-to-interference-plus-noise ratio (SINR). However, the SINR analysis carried out in [74] considers steady-state performances, i.e., when the sample size is infinite, and, thus, does not allow to evaluate the performance of the receivers as a function of the number of samples. This issue is important from a practical point of view because, especially when short sample-sizes are employed, the data-estimated versions of the WL-MUD receivers exhibit a severe performance degradation with respect to their ideal counterparts, reducing thus the expected performance gain over L-MUD receivers.

To gain more insight about these points, by applying a first-order perturbative approach at the first, we evaluate in the subsection 4.5.1 the performance degradation due to finite-sample ACM estimation in the known-channel case. In particular this analysis is carried out with reference two different data-estimated implementations of the L-MOE and WL-MOE receivers: the SMI (sample matrix inversion) receiver (and L-SMI and WL-SMI), which employs a sample estimate of the data autocorrelation matrix, and the SUB (subspace) receiver (and L-SUB and WL-SUB), which exploits the properties of the eigenvalue decomposition (EVD) of the data autocorrelation matrix to reduce the effects of estimation errors. As we will see in the subsection 4.5.1 we will derive easily interpretable formulas, which allow one to obtain clear insights about the effects of different parameters on performances. Moreover, the results of

the analysis will show that the WL-MOE receiver is more sensitive than its linear counterpart to finite sample-size effects associated to ACM estimation, and it generally requires subspace-based implementation to achieve in practice the performance gains predicted by theory.

Successively, in the subsection 4.5.3, the first-order perturbation analysis developed in subsection 4.5.1 is extended to incorporate the effects due to subspace-based blind CE (see subsection 4.3.2) on the synthesis of the L- and WL-MOE receivers. It is worthwhile to note that, when the desired channel vector has been estimated through the subspace method and, hence, the subspace decomposition of the ACM is already available, it is preferable from a computational viewpoint to implement the L- and WL-SUB receivers rather than their SMI counterparts, since they do not require direct ACM inversion. Notwithstanding this, we have chosen to carry out also the performance analysis of the SMI versions of the receivers with CE since, in this way, an interesting comparison with the SUB versions of the receivers, as well as with the SMI versions when the channel is assumed known, can be established. In particular, we will derive easily interpretable formulas, supported by computer simulation. Moreover, in the subsection (4.5.3) we will show that with reference to subspace-based receivers implementations, for moderate-to-high values of the SNR, errors in estimating the L-SUB-CE and WL-SUB-CE receivers are essentially due to ACM estimation. This is not true for the L-SMI-CE and WL-SMI-CE receivers, for which CE errors undesirably combine with ACM errors (signature mismatch phenomenon). Therefore, we will conclude that when considering finite sample-size implementation, the blind WL-MOE receiver is able to assure a significant performance gain (for low-to-moderate values of the SNR) with respect to its linear counterpart only when it is built by resorting to the more sophisticated subspace-based implementation. In this case, for a given channel length, it allows one to work with an increased number of users J, which makes it a viable choice in heavily-congested DS-CDMA networks.

4.5.1 Finite-sample performances of the L-MOE and WL-MOE receivers with known channel

WL-MOE performance analysis

Let us start from the WL-MOE receiver. Preliminarily, we observe that an equivalent form of the WL-MOE receiver (4.27) can be obtained by exploiting the eigenvalue decomposition (EVD) of $\mathbf{R}_{\mathbf{zz}}$ (4.43). To this end, it is required that the augmented matrix $\mathbf{H}$ is full-column rank (an issue that has been dis-

cussed in subsection 4.4.1 for downlink scenario), which necessarily requires that

$$J \leq 2\,N. \tag{4.61}$$

As a matter of fact, this assumption (4.61) is not required for the WL-SMI receiver and it is necessary only for the WL-SUB one. However, since the WL-MOE receiver is not able to ensure perfect MAI suppression, for each user, when $\mathbf{H}$ is rank-deficient (see subsection 4.4.1) , we maintain the assumption $\text{rank}(\mathbf{H}) = J$ for both the two data-estimated WL receivers. By substituting the EVD of $\mathbf{R_{zz}}$ in (4.27) and exploiting the orthogonality between signal and noise subspaces, one obtains the *subspace-based form* of the WL-MOE receiver as follows

$$\mathbf{f}_{j,\text{WL-MOE}} = (\mathbf{h}_j^H \mathbf{U}_s \mathbf{\Lambda}_s^{-1} \mathbf{U}_s^H \mathbf{h}_j)^{-1} \mathbf{U}_s \mathbf{\Lambda}_s^{-1} \mathbf{U}_s^H\, \mathbf{h}_j, \quad \text{for } J \leq 2\,N. \tag{4.62}$$

Implementation of the WL-MOE receiver defined by (4.27) or (4.62) requires estimation from the received data of $\mathbf{R_{zz}}$ in (4.27) or its EVD in (4.62). Under mild conditions, a consistent estimate $\widehat{\mathbf{R}}_{\mathbf{zz}}$ of $\mathbf{R_{zz}}$ is the sample ACM obtained as

$$\widehat{\mathbf{R}}_{\mathbf{zz}} = \frac{1}{K} \sum_{k=1}^{K} \mathbf{z}(k)\, \mathbf{z}^H(k)\,, \tag{4.63}$$

where K denotes the estimation sample size. Applying the EVD to $\widehat{\mathbf{R}}_{\mathbf{zz}}$, one obtains the decomposition

$$\widehat{\mathbf{R}}_{\mathbf{zz}} = \widehat{\mathbf{U}}_s \widehat{\mathbf{\Lambda}}_s \widehat{\mathbf{U}}_s^H + \widehat{\mathbf{U}}_n \widehat{\mathbf{\Lambda}}_n \widehat{\mathbf{U}}_n^H, \tag{4.64}$$

where the matrices $\widehat{\mathbf{U}}_s \in \mathbb{C}^{2N\times J}$, $\widehat{\mathbf{U}}_n \in \mathbb{C}^{2N\times(2N-J)}$, $\widehat{\mathbf{\Lambda}}_s \in \mathbb{R}^{J\times J}$, and $\widehat{\mathbf{\Lambda}}_n \in \mathbb{R}^{2N\times 2N}$ are estimates of the matrices in (4.43) $\mathbf{U}_s$, $\mathbf{U}_n$, $\mathbf{\Lambda}_s$, and $\mathbf{\Lambda}_n = \sigma_v^2\, \mathbf{I}_{2N}$, respectively. By substituting in (4.27) and (4.62), the sample ACM (4.63) and its EVD (4.64) respectively, the WL-SMI and WL-SUB receivers are given by

$$\mathbf{f}_{j,\text{WL-SMI}} \triangleq (\mathbf{h}_j^H \widehat{\mathbf{R}}_{\mathbf{zz}}^{-1} \mathbf{h}_j)^{-1}\, \widehat{\mathbf{R}}_{\mathbf{zz}}^{-1}\, \mathbf{h}_j\,. \tag{4.65}$$

$$\mathbf{f}_{j,\text{WL-SUB}} \triangleq (\mathbf{h}_j^H\, \widehat{\mathbf{U}}_s\, \widehat{\mathbf{\Lambda}}_s^{-1}\, \widehat{\mathbf{U}}_s^H\, \mathbf{h}_j)^{-1}\, \widehat{\mathbf{U}}_s\, \widehat{\mathbf{\Lambda}}_s^{-1}\, \widehat{\mathbf{U}}_s^H\, \mathbf{h}_j\,. \tag{4.66}$$

It is worth noting that the weight vector $\mathbf{f}_{j,\text{WL-SUB}}$ is not equal to $\mathbf{f}_{j,\text{WL-SMI}}$, since $\widehat{\mathbf{U}}_n^H \mathbf{h}_j \neq \mathbf{0}_{2N-J}$ due to the finite-sample-size effects. This implies that the two receivers WL-SMI and WL-SUB might exhibit different SINR performances.

To carry out the performance analysis for WL-SMI and WL-SUB in an unified framework, let us denote with $\widehat{\mathbf{f}}_j$ *any* data-estimated WL-MOE receiver, i.e, $\widehat{\mathbf{f}}_j = \mathbf{f}_{j,\text{WL-SMI}}$ or $\widehat{\mathbf{f}}_j = \mathbf{f}_{j,\text{WL-SUB}}$, and set $\mathbf{f}_j = \mathbf{f}_{j,\text{WL-MOE}}$ for simplicity, where $\mathbf{f}_{j,\text{WL-MOE}}$ is given by (4.27) or (4.62). Adopting a perturbation perspective, the vector $\widehat{\mathbf{f}}_j$ can be expressed as

$$\widehat{\mathbf{f}}_j = \mathbf{f}_j + \boldsymbol{\delta}\mathbf{f}_j \,, \tag{4.67}$$

where $\boldsymbol{\delta}\mathbf{f}_j$ is a *small* (i.e., $\|\boldsymbol{\delta}\mathbf{f}_j\| \ll 1$) zero-mean perturbation term. Since any data-estimated version of the WL-MOE receiver must satisfy the constraint $\widehat{\mathbf{f}}_j^H \mathbf{h}_j = 1$, it results that $\boldsymbol{\delta}\mathbf{f}_j^H \mathbf{h}_j = 0$, thus the SINR (4.24) for the data-estimated receivers can be written as

$$\text{SINR}(\widehat{\mathbf{f}}_j) = \frac{1}{\text{E}_{\widehat{\mathbf{f}}_j,\mathbf{q}_j}\left\{\text{Re}^2[\widehat{\mathbf{f}}_j^H\, \mathbf{q}_j(k)]\right\}} \,, \tag{4.68}$$

where the symbol $\text{E}_{\widehat{\mathbf{f}}_j,\mathbf{q}_j}[\,\cdot\,]$ denotes *joint* average w.r.t to $\widehat{\mathbf{f}}_j$ and $\mathbf{q}_j(k)$ of the quantity in brackets. A simplifying and reasonable assumption [75] is that $\widehat{\mathbf{f}}_j$ is independent from $\mathbf{q}_j(k)$. In this case, by accounting for the CS property of $\widehat{\mathbf{f}}_j$, substituting (4.67) into (4.68), performing the average w.r.t to $\mathbf{q}_j(k)$, and recalling that, due to assumptions **(a1)** and **(a2)**, the vector $\mathbf{q}_j(k)$ is zero-mean, one has:

$$\text{SINR}(\widehat{\mathbf{f}}_j) = \frac{1}{\mathbf{f}_j^H \mathbf{R}_{\mathbf{q}_j\mathbf{q}_j} \mathbf{f}_j + \text{E}_{\boldsymbol{\delta}\mathbf{f}_j}[\boldsymbol{\delta}\mathbf{f}_j^H \mathbf{R}_{\mathbf{q}_j\mathbf{q}_j} \boldsymbol{\delta}\mathbf{f}_j]} \,, \tag{4.69}$$

where only the average w.r.t to $\boldsymbol{\delta}\mathbf{f}_j$ must be evaluated. To perform this calculation, we need explicit expressions for the perturbation $\boldsymbol{\delta}\mathbf{f}_j$ of the WL-SMI and WL-SUB receivers, which are provided by the following Lemma.

Lemma 4.2 *Let denote with* $\approx$ *first-order equality*[14]*, assume that* $\mathbf{H}$ *is full-column rank and let* $\widehat{\mathbf{R}}_{\mathbf{zz}}$ *be estimated by* (4.63)*. The* first-order *perturbation term of the WL-SMI and WL-SUB receivers can be expressed as*

$$\boldsymbol{\delta}\mathbf{f}_j \approx -\boldsymbol{\Gamma}_{j,WL}\, \widehat{\mathbf{r}}_{\mathbf{q}_j b_j} \,, \tag{4.70}$$

where $\widehat{\mathbf{r}}_{\mathbf{q}_j b_j} \triangleq \frac{1}{K}\sum_{k=0}^{K-1} \mathbf{q}_j(k)\, b_j(k)$ *is the sample estimate of the cross-correlation between the disturbance vector* $\mathbf{q}_j(k)$ *and the desired symbol*

[14] First-order equality means that, as the sample size K approaches infinity, we neglect all the summands that tend to zero faster than the norm of the corresponding perturbation term.

$b_j(k)$, *and*

$$\mathbf{\Gamma}_{j,WL} = \begin{cases} \mathbf{P}_{j,\mathrm{WL}}\,\mathbf{R}^{-1}_{\mathbf{q}_j\mathbf{q}_j} & \text{(WL-SMI)} \\ \mathbf{P}_{j,\mathrm{WL}}\,\mathbf{R}^{-1}_{\mathbf{q}_j\mathbf{q}_j} - \gamma_{j,\mathrm{WL}}\,\mathbf{U}_n\,\mathbf{U}_n^H & \text{(WL-SUB)} \end{cases} \tag{4.71}$$

with $\mathbf{P}_{j,\mathrm{WL}} \triangleq \mathbf{I}_{2N} - (\mathbf{h}_j^H\,\mathbf{R}^{-1}_{\mathbf{q}_j\mathbf{q}_j}\,\mathbf{h}_j)^{-1}\,\mathbf{R}^{-1}_{\mathbf{q}_j\mathbf{q}_j}\,\mathbf{h}_j\,\mathbf{h}_j^H \in \mathbb{C}^{N\times N}$ *denoting an oblique projection matrix [75], and* $\gamma_{j,\mathrm{WL}} \triangleq \sigma_v^{-2} + (\mathbf{h}_j^H\,\mathbf{R}^{-1}_{\mathbf{zz}}\,\mathbf{h}_j)^{-1}\mathbf{h}_j^H\mathbf{U}_s\,\mathbf{\Omega}^{-1}_{WL}\,\mathbf{U}_s^H\mathbf{R}^{-1}_{\mathbf{zz}}\,\mathbf{h}_j$, *where* $\mathbf{\Omega}_{WL} \triangleq \mathbf{\Lambda}_{\mathbf{s}} - \sigma_v^2\,\mathbf{I}_J \in \mathbb{R}^{J\times J}$.

Proof. See Appendix B.4.

It should be observed that Lemma 4.2 provides a compact characterization of the perturbation terms, obtained under the simplifying assumption [75] that the predominant error in estimating $\mathbf{R}_{\mathbf{zz}}$ is due to $\widehat{\mathbf{r}}_{\mathbf{q}_j b_j}$ (see Appendix B.4 for details). This approximation will allows us to obtain simple yet accurate results, which will be validated in subsection 4.5.2. Accounting for Lemma 4.2, the average in (4.69) can be expressed as (we drop the subscript $\boldsymbol{\delta}\mathbf{f}_j$ in $\mathrm{E}_{\boldsymbol{\delta}\mathbf{f}_j}[\cdot]$ for notational simplicity)

$$\begin{aligned} \mathrm{E}[\boldsymbol{\delta}\mathbf{f}_j^H\,\mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}\,\boldsymbol{\delta}\mathbf{f}_j] &= \mathrm{E}[\widehat{\mathbf{r}}^H_{\mathbf{q}_j b_j}\,\mathbf{\Gamma}^H_{j,\mathrm{WL}}\,\mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}\,\mathbf{\Gamma}_{j,\mathrm{WL}}\,\widehat{\mathbf{r}}_{\mathbf{q}_j b_j}] \\ &= \mathrm{trace}\{\mathbf{\Gamma}^H_{j,\mathrm{WL}}\,\mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}\,\mathbf{\Gamma}_{j,\mathrm{WL}}\,\mathrm{E}[\widehat{\mathbf{r}}_{\mathbf{q}_j b_j}\,\widehat{\mathbf{r}}^H_{\mathbf{q}_j b_j}]\}, \end{aligned} \tag{4.72}$$

where, by accounting for assumptions **(a1)** and **(a2)**, one has:

$$\begin{aligned} \mathrm{E}[\widehat{\mathbf{r}}_{\mathbf{q}_j b_j}\,\widehat{\mathbf{r}}^H_{\mathbf{q}_j b_j}] &= \frac{1}{K^2}\sum_{k,h=1}^{K}\mathrm{E}[\mathbf{q}_j(k)\,b_j(k)\,b_j(h)\,\mathbf{q}_j^H(h)] \\ &= \frac{1}{K^2}\sum_{k,h=1}^{K}\mathrm{E}[\mathbf{q}_j(k)\,\mathbf{q}_j^H(h)]\,\mathrm{E}[b_j(k)\,b_j(h)] \\ &= \frac{1}{K^2}\sum_{k,h=1}^{K}\mathrm{E}[\mathbf{q}_j(k)\,\mathbf{q}_j^H(h)]\,\delta_{k-h} \qquad (4.73) \\ &= \frac{1}{K^2}\sum_{k=1}^{K}\mathrm{E}[\mathbf{q}_j(k)\,\mathbf{q}^H(k)] = \frac{1}{K}\,\mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}\,. \qquad (4.74) \end{aligned}$$

By substituting (4.74) in (4.125), the result back in (4.69), and recalling that $\mathbf{f}_j^H\mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}\mathbf{f}_j = \mathrm{SINR}^{-1}_{j,\text{WL-MOE}}$, where $\mathrm{SINR}_{j,\text{WL-MOE}}$ is given by (4.49), we get

$$\text{SINR}(\widehat{\mathbf{f}}_j) = \frac{\text{SINR}_{j,\text{WL-MOE}}}{1 + \frac{\text{trace}(\mathbf{\Gamma}_{j,\text{WL}}^H \mathbf{R}_{\mathbf{q}_j\mathbf{q}_j} \mathbf{\Gamma}_{j,\text{WL}} \mathbf{R}_{\mathbf{q}_j\mathbf{q}_j})}{K} \text{SINR}_{j,\text{WL-MOE}}} . \quad (4.75)$$

The final result is obtained by evaluating the trace$(\cdot)$ term in (4.75), on the basis of the $\mathbf{\Gamma}_{j,\text{WL}}$ expressions given in Lemma 4.2. To do this, it is convenient to consider the SMI and SUB cases separately. With reference to the WL-SMI receiver, since $\mathbf{\Gamma}_{j,\text{WL}} = \mathbf{P}_{j,\text{WL}} \mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}^{-1}$, by using the properties of the trace operator, after some algebraic manipulations, one obtains:

$$\text{trace}(\mathbf{\Gamma}_{j,\text{WL}}^H \mathbf{R}_{\mathbf{q}_j\mathbf{q}_j} \mathbf{\Gamma}_{j,\text{WL}} \mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}) = \frac{\text{trace}(\mathbf{P}_{j,\text{WL}} \mathbf{P}_{j,\text{WL}}^H)}{K} = \frac{2N-1}{K}, \quad (4.76)$$

which can be substituted in (4.75), thus leading to

$$\text{SINR}_{j,\text{WL-SMI}} \stackrel{\triangle}{=} \text{SINR}(\mathbf{f}_{j,\text{WL-SMI}}) = \frac{\text{SINR}_{j,\text{WL-MOE}}}{1 + \frac{2N-1}{K} \text{SINR}_{j,\text{WL-MOE}}} . \quad (4.77)$$

As regards the WL-SUB receiver, since $\mathbf{\Gamma}_{j,\text{WL}} = \mathbf{P}_{j,\text{WL}} \mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}^{-1} - \gamma_{j,\text{WL}} \mathbf{U}_n \mathbf{U}_n^H$, by using again the properties of the trace operator and observing that $\mathbf{U}_n^H \mathbf{h}_j = \mathbf{0}_{2N-J}$, after some algebra, one has:

$$\begin{aligned} &\text{trace}(\mathbf{\Gamma}_{j,\text{WL}}^H \mathbf{R}_{\mathbf{q}_j\mathbf{q}_j} \mathbf{\Gamma}_{j,\text{WL}} \mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}) \\ &\quad = 2N - 1 - (\gamma_{j,\text{WL}}^* \sigma_v^2 + \gamma_{j,\text{WL}} \sigma_v^2 - |\gamma_{j,\text{WL}} \sigma_v^2|^2)(2N - J) \\ &\quad = (J-1) + (2N - J)\,|1 - \gamma_{j,\text{WL}} \sigma_v^2|^2 . \end{aligned} \quad (4.78)$$

After substituting (4.78) into (4.75), one gets:

$$\begin{aligned} \text{SINR}_{j,\text{WL-SUB}} &\stackrel{\triangle}{=} \text{SINR}(\mathbf{f}_{j,\text{WL-SUB}}) \\ &= \frac{\text{SINR}_{j,\text{WL-MOE}}}{1 + \frac{(J-1)+(2N-J)\,|1-\gamma_{j,\text{WL}} \sigma_v^2|^2}{K} \text{SINR}_{j,\text{WL-MOE}}} . \end{aligned} \quad (4.79)$$

The expression (4.79) for the WL-SUB receiver can be further simplified by observing that, for $\sigma_v^2 \to 0$, one has $\gamma_{j,\text{WL}} \sigma_v^2 \to 1$, hence the trace in (4.78) reduces to $J - 1$. By accounting for this observation, for moderate-to-high values of the SNR, eq. (4.79) can be approximatively written as

$$\text{SINR}_{j,\text{WL-SUB}} = \frac{\text{SINR}_{j,\text{WL-MOE}}}{1 + \frac{J-1}{K} \text{SINR}_{j,\text{WL-MOE}}} . \quad (4.80)$$

It is worth noting that, despite of the apparent similarity between (4.77)–(4.80) and the SINR formulas reported in [74, eqs. (14) and (25)], our results are not directly comparable with those of [74]. Indeed, the results of [74] report the SINR performances of the LMS-based adaptive implementation of the WL-MMSE and WL-MOE receivers only for $K \to +\infty$ (steady-state performances); in this latter case the performance penalty paid by the WL-MUD receivers with respect to their ideal counterparts is exclusively due to gradient-noise effects.

L-MOE performance analysis

The finite-sample performance analysis of the L-MOE receivers is now in order. Similarly to the WL-MOE receiver, under condition **(c1)**, the L-MOE one (4.16) can be equivalently represented in subspace-based form by exploiting the eigenvalue decomposition (EVD) of $\mathbf{R_{rr}}$ (4.37). To this end, it is required that the matrix $\mathbf{\Phi}$ is full-column rank (an issue that has been discussed in subsection 4.4.1 for downlink scenario), which necessarily requires that[15]

$$J \leq N. \tag{4.81}$$

By substituting the EVD of $\mathbf{R_{rr}}$ in (4.16) and exploiting the orthogonality between signal and noise subspaces, one obtains the *subspace-based form* of the L-MOE receiver as follows

$$\mathbf{w}_{j,\text{L-SUB}} \stackrel{\triangle}{=} (\boldsymbol{\phi}_j^H \, \mathbf{V}_s \, \boldsymbol{\Upsilon}_s^{-1} \, \mathbf{V}_s^H \, \boldsymbol{\phi}_j)^{-1} \, \mathbf{V}_s \, \boldsymbol{\Upsilon}_s^{-1} \, \mathbf{V}_s^H \, \boldsymbol{\phi}_j \, . \tag{4.82}$$

As in the WL case, implementation of the L-MOE receiver defined by (4.16) or (4.82) requires estimation from the received data of $\mathbf{R_{rr}}$ in (4.16) or its EVD in (4.82). Under mild conditions, a consistent estimate $\widehat{\mathbf{R}}_{\mathbf{rr}}$ of $\mathbf{R_{rr}}$ is the sample ACM obtained as

$$\widehat{\mathbf{R}}_{\mathbf{rr}} = \frac{1}{K} \sum_{k=1}^{K} \mathbf{r}(k) \, \mathbf{r}^H(k) \, . \tag{4.83}$$

where K denotes the estimation sample size. Applying the EVD to $\widehat{\mathbf{R}}_{\mathbf{rr}}$, one obtains the decomposition

$$\widehat{\mathbf{R}}_{\mathbf{rr}} = \widehat{\mathbf{V}}_s \widehat{\boldsymbol{\Upsilon}}_s \widehat{\mathbf{V}}_s^H + \widehat{\mathbf{V}}_n \widehat{\boldsymbol{\Upsilon}}_n \widehat{\mathbf{V}}_n^H \, , \tag{4.84}$$

[15] As we have noted in the WL case, this assumption (4.81) is not required for the L-SMI receiver and it is necessary only for the L-SUB one. However, since the L-MOE receiver is not able to ensure perfect MAI suppression, for each user, when $\mathbf{\Phi}$ is rank-deficient (see subsection 4.4.1) , we maintain the assumption $\text{rank}(\mathbf{\Phi}) = J$ for both the two data-estimated L receivers.

where the matrices $\widehat{\mathbf{V}}_s \in \mathbb{C}^{N\times J}$, $\widehat{\mathbf{V}}_n \in \mathbb{C}^{N\times(N-J)}$, $\widehat{\boldsymbol{\Upsilon}}_s \in \mathbb{R}^{J\times J}$, and $\widehat{\boldsymbol{\Upsilon}}_n \in \mathbb{R}^{N\times N}$ are estimates of the matrices in (4.37) $\mathbf{V}_s$, $\mathbf{V}_n$, $\boldsymbol{\Upsilon}_s$, and $\boldsymbol{\Upsilon}_n = \sigma_v^2\,\mathbf{I}_N$, respectively. By substituting in (4.16) and (4.82), the sample ACM (4.83) and its EVD (4.84) respectively, the L-SMI and L-SUB receivers are given by

$$\mathbf{w}_{j,\text{L-SMI}} \stackrel{\triangle}{=} (\boldsymbol{\phi}_j^H \widehat{\mathbf{R}}_{\mathbf{rr}}^{-1} \boldsymbol{\phi}_j)^{-1}\, \widehat{\mathbf{R}}_{\mathbf{rr}}^{-1}\, \boldsymbol{\phi}_j \tag{4.85}$$

$$\mathbf{w}_{j,\text{L-SUB}} \stackrel{\triangle}{=} (\boldsymbol{\phi}_j^H\, \widehat{\mathbf{V}}_s\, \widehat{\boldsymbol{\Upsilon}}_s^{-1}\, \widehat{\mathbf{V}}_s^H\, \boldsymbol{\phi}_j)^{-1}\, \widehat{\mathbf{V}}_s\, \widehat{\boldsymbol{\Upsilon}}_s^{-1}\, \widehat{\mathbf{V}}_s^H\, \boldsymbol{\phi}_j\,, \tag{4.86}$$

As for Wl receivers, while (4.16) (4.82) are perfectly equivalent, their estimated counterparts (4.85) and (4.86) are different.

In order to carry out the performance analysis of the L-SMI and L-SUB receivers, it should be stressed that, since the relevant SINR is the one after the Re$[\cdot]$ part, one cannot simply apply results available in the literature (e.g., [76]), since they refer to the SINR evaluated before the Re$[\cdot]$ part.

From a unified perspective, let us denote with $\widehat{\mathbf{w}}_j$ *any* data-estimated L-MOE receiver, i.e, $\widehat{\mathbf{w}}_j = \mathbf{w}_{j,\text{L-SMI}}$ or $\widehat{\mathbf{w}}_j = \mathbf{w}_{j,\text{L-SUB}}$, and set $\mathbf{w}_j = \mathbf{w}_{j,\text{L-MOE}}$ for simplicity. Adopting a perturbation approach, the vector $\widehat{\mathbf{w}}_j$ can be expressed as

$$\widehat{\mathbf{w}}_j = \mathbf{w}_j + \boldsymbol{\delta}\mathbf{w}_j\,, \tag{4.87}$$

where $\boldsymbol{\delta}\mathbf{w}_j$ is a *small* (in the Frobenius norm sense) zero-mean additive perturbation. Since any data-estimated version of the L-MOE receiver must satisfy the constraint $\widehat{\mathbf{w}}_j^H \boldsymbol{\phi}_j = 1$, it results that $\boldsymbol{\delta}\mathbf{w}_j^H \boldsymbol{\phi}_j = 0$. Thus, using the identity $\text{Re}^2[z] = \frac{1}{2}\{|z|^2 + \text{Re}[z^2]\}$, $\forall z \in \mathbb{C}$, the SINR (4.24) for data-estimated linear receivers becomes

$$\text{SINR}(\widehat{\mathbf{w}}_j) = \frac{2}{\text{E}_{\widehat{\mathbf{w}}_j,\mathbf{p}_j}\left\{|\widehat{\mathbf{w}}_j^H \mathbf{p}_j(k)|^2\right\} + \text{E}_{\widehat{\mathbf{w}}_j,\mathbf{p}_j}\left\{\text{Re}[(\widehat{\mathbf{w}}_j^H \mathbf{p}_j(k))^2]\right\}}\,. \tag{4.88}$$

Similarly to the WL case, we assume that $\widehat{\mathbf{w}}_j$ is independent from $\mathbf{p}_j(k)$. In this case, by substituting (4.87) into (4.88), performing the average w.r.t to $\mathbf{p}_j(k)$, and recalling that, due to assumptions **(a1)** and **(a2)**, the vector $\mathbf{p}_j(k)$ is zero-mean, one has

$$\begin{aligned}\text{SINR}(\widehat{\mathbf{w}}_j)^{-1} = \frac{1}{2}\Big\{&\mathbf{w}_j^H \mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}\, \mathbf{w}_j + \text{E}_{\boldsymbol{\delta}\mathbf{w}_j}[\boldsymbol{\delta}\mathbf{w}_j^H \mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}\, \boldsymbol{\delta}\mathbf{w}_j] \\ &+\text{Re}[\mathbf{w}_j^H \mathbf{R}_{\mathbf{p}_j\mathbf{p}_j^*}\, \mathbf{w}_j^*] + \text{Re}\{\text{E}_{\boldsymbol{\delta}\mathbf{w}_j}[\boldsymbol{\delta}\mathbf{w}_j^H \mathbf{R}_{\mathbf{p}_j\mathbf{p}_j^*}\, \boldsymbol{\delta}\mathbf{w}_j^*]\}\Big\}\,.\end{aligned} \tag{4.89}$$

The characterization of the perturbation term $\boldsymbol{\delta}\mathbf{w}_j$ is given by the following Lemma.

Lemma 4.3 *Assume that* $\mathbf{\Phi}$ *is full-column rank and let* $\widehat{\mathbf{R}}_{\mathbf{rr}}$ *be estimated by* (4.83). *The* first-order *perturbation term of the L-SMI and L-SUB receivers can be expressed as*

$$\boldsymbol{\delta}\mathbf{w}_j = -\mathbf{\Gamma}_{j,\mathrm{L}}\,\widehat{\mathbf{r}}_{\mathbf{p}_j b_j}\,, \tag{4.90}$$

where $\widehat{\mathbf{r}}_{\mathbf{p}_j b_j} \stackrel{\triangle}{=} \frac{1}{K}\sum_{k=1}^{K}\mathbf{p}_j(k)\,b_j(k)$ *is the sample cross-correlation between the interference and the desired signal, and*

$$\mathbf{\Gamma}_{j,\mathrm{L}} = \begin{cases} \mathbf{P}_{j,\mathrm{L}}\,\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}^{-1}\,, & \text{(L-SMI)} \\ \mathbf{P}_{j,\mathrm{L}}\,\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}^{-1} - \gamma_{j,\mathrm{L}}\mathbf{V}_n\,\mathbf{V}_n^H\,, & \text{(L-SUB)} \end{cases} \tag{4.91}$$

with $\mathbf{P}_{j,\mathrm{L}} \stackrel{\triangle}{=} \mathbf{I}_N - (\boldsymbol{\phi}_j^H\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}^{-1}\boldsymbol{\phi}_j)^{-1}\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}^{-1}\boldsymbol{\phi}_j\,\boldsymbol{\phi}_j^H = \mathbf{I}_N - (\boldsymbol{\phi}_j^H\mathbf{R}_{\mathbf{rr}}^{-1}\boldsymbol{\phi}_j)^{-1}\mathbf{R}_{\mathbf{rr}}^{-1}\boldsymbol{\phi}_j\,\boldsymbol{\phi}_j^H$ *and* $\gamma_{j,\mathrm{L}} \stackrel{\triangle}{=} \sigma_v^{-2} + (\boldsymbol{\phi}_j^H\mathbf{R}_{\mathbf{rr}}^{-1}\boldsymbol{\phi}_j)^{-1}\boldsymbol{\phi}_j^H\mathbf{V}_s\,\mathbf{\Omega}_L^{-1}\mathbf{V}_s^H\mathbf{R}_{\mathbf{rr}}^{-1}\boldsymbol{\phi}_j$, *where* $\mathbf{\Omega}_L \stackrel{\triangle}{=} \mathbf{\Upsilon}_{\mathbf{s}} - \sigma_v^2\,\mathbf{I}_J \in \mathbb{R}^{J\times J}$.

Proof. The proof is omitted since it is similar to that of Lemma 4.2.

By virtue of Lemma 4.3, we are now able to evaluate the averages in (4.89). Specifically, dropping the subscript $\boldsymbol{\delta}\mathbf{w}_j$ in $\mathrm{E}_{\boldsymbol{\delta}\mathbf{w}_j}[\cdot]$ for notational simplicity, we have:

$$\begin{aligned}\mathrm{E}[\boldsymbol{\delta}\mathbf{w}_j^H\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}\,\boldsymbol{\delta}\mathbf{w}_j] &= \mathrm{trace}\{\mathbf{\Gamma}_{j,\mathrm{L}}^H\,\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}\,\mathbf{\Gamma}_{j,\mathrm{L}}\,\mathrm{E}[\widehat{\mathbf{r}}_{\mathbf{p}_j b_j}\,\widehat{\mathbf{r}}_{\mathbf{p}_j b_j}^H]\} \\ &= \frac{1}{K}\,\mathrm{trace}(\mathbf{\Gamma}_{j,\mathrm{L}}^H\,\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}\,\mathbf{\Gamma}_{j,\mathrm{L}}\,\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j})\,,\end{aligned} \tag{4.92}$$

$$\begin{aligned}\mathrm{E}[\boldsymbol{\delta}\mathbf{w}_j^H\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j^*}\,\boldsymbol{\delta}\mathbf{w}_j^*] &= \mathrm{trace}\{\mathbf{\Gamma}_{j,\mathrm{L}}^H\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j^*}\,\mathbf{\Gamma}_{j,\mathrm{L}}^*\,\mathrm{E}[\widehat{\mathbf{r}}_{\mathbf{p}_j b_j}^*\,\widehat{\mathbf{r}}_{\mathbf{p}_j b_j}^H]\} \\ &= \frac{1}{K}\,\mathrm{trace}(\mathbf{\Gamma}_{j,\mathrm{L}}^H\,\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j^*}\,\mathbf{\Gamma}_{j,\mathrm{L}}^*\,\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j^*}^*)\,.\end{aligned} \tag{4.93}$$

By substituting (4.92) and (4.93) into (4.89), and recalling the equation (4.50), that here we report for simplicity $\mathrm{SINR}_{j,\text{L-MOE}}{}^{-1} = (\mathbf{w}_j^H\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}\,\mathbf{w}_j + \mathrm{Re}[\mathbf{w}_j^H\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j^*}\,\mathbf{w}_j^*])/2$, we get:

$$\begin{aligned}\mathrm{SINR}(\widehat{\mathbf{w}}_j)^{-1} = \mathrm{SINR}_{j,\text{L-MOE}}^{-1}\cdot\Big\{1 + \frac{1}{2K}\Big[&\mathrm{trace}(\mathbf{\Gamma}_{j,\mathrm{L}}^H\,\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}\,\mathbf{\Gamma}_{j,\mathrm{L}}\,\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}) \\ &+ \mathrm{Re}[\mathrm{trace}(\mathbf{\Gamma}_{j,\mathrm{L}}^H\,\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j^*}\,\mathbf{\Gamma}_{j,\mathrm{L}}^*\,\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j^*}^*)]\Big]\,\mathrm{SINR}_{j,\text{L-MOE}}\Big\}\,.\end{aligned} \tag{4.94}$$

Along the same lines of the WL case, it can be shown that

$$\text{trace}(\mathbf{\Gamma}_{j,\text{L}}^{H}\,\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}\,\mathbf{\Gamma}_{j,\text{L}}\,\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}) = \begin{cases} N-1\,, & \text{(L-SMI)} \\ J-1+(N-J)|1-\gamma_{j,\text{L}}\,\sigma_v^2|^2 & \text{(L-SUB).} \end{cases} \tag{4.95}$$

On the other hand, evaluation of the term $\text{trace}(\mathbf{\Gamma}_{j,\text{L}}^{H}\,\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j^*}\,\mathbf{\Gamma}_{j,\text{L}}^{*}\,\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j^*}^{*})$ is more complicated and, for its calculation, it is convenient to consider the SMI and SUB cases separately. With reference to the L-SMI receiver, since $\mathbf{\Gamma}_{j,\text{L}} = \mathbf{P}_{j,\text{L}}\,\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}^{-1}$, after simple algebra, one obtains

$$\begin{aligned} \text{trace}(\mathbf{\Gamma}_{j,\text{L}}^{H}\,\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j^*}\,\mathbf{\Gamma}_{j,\text{L}}^{*}\,\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j^*}^{*}) &= \text{trace}[\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}^{-1}\,\mathbf{P}_{j,\text{L}}^{H}\,\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j^*}\,\mathbf{P}_{j,\text{L}}^{*}\,(\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}^{*})^{-1}\,\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j^*}^{*}] \\ &= \text{trace}[\mathbf{P}_{j,\text{L}}\,\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}^{-1}\,\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j^*}\,(\mathbf{P}_{j,\text{L}}\,\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}^{-1}\,\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j^*})^{*}] \\ &= \text{trace}[\,\mathbf{P}_{j,\text{L}}\,\mathbf{R}_{\mathbf{rr}}^{-1}\,\mathbf{R}_{\mathbf{rr}^*}\,(\mathbf{P}_{j,\text{L}}\,\mathbf{R}_{\mathbf{rr}}^{-1}\,\mathbf{R}_{\mathbf{rr}^*})^{*}], \end{aligned} \tag{4.96}$$

where we used the identities $\mathbf{P}_{j,\text{L}}\,\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}^{-1} = \mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}^{-1}\mathbf{P}_{j,\text{L}}^{H}$ and $\mathbf{P}_{j,\text{L}}\,\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}^{-1}\,\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j^*} = \mathbf{P}_{j,\text{L}}\,\mathbf{R}_{\mathbf{rr}}^{-1}\,\mathbf{R}_{\mathbf{rr}^*}$. To obtain a more manageable expression of $\text{trace}(\mathbf{\Gamma}_{j,\text{L}}^{H}\,\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j^*}\,\mathbf{\Gamma}_{j,\text{L}}^{*}\,\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j^*}^{*})$, we consider its asymptotic value as $\sigma_v^2 \to 0$, i.e., in the high-SNR regime. By accounting for the expression of $\mathbf{P}_{j,\text{L}}$ given by Lemma 4.3, recalling that, under assumptions **(a1)** and **(a2)**, $\mathbf{R}_{\mathbf{rr}} = \mathbf{\Phi}\,\mathbf{\Phi}^H + \sigma_v^2\,\mathbf{I}_N$ (see equation 4.29), and $\mathbf{R}_{\mathbf{rr}^*} = \mathbf{\Phi}\,\mathbf{\Phi}^T$, and resorting to the limit formula for the Moore-Penrose inverse [24], one has

$$\begin{aligned} &\lim_{\sigma_v^2\to 0}\,\mathbf{P}_{j,\text{L}}\,\mathbf{R}_{\mathbf{rr}}^{-1}\,\mathbf{R}_{\mathbf{rr}^*} \\ &= \lim_{\sigma_v^2\to 0}\left\{\left[\mathbf{I}_N - \frac{(\mathbf{\Phi}\,\mathbf{\Phi}^H+\sigma_v^2\mathbf{I}_N)^{-1}\mathbf{\Phi}\,\mathbf{1}_j\,\boldsymbol{\phi}_j^H}{\boldsymbol{\phi}_j^H(\mathbf{\Phi}\mathbf{\Phi}^H+\sigma_v^2\mathbf{I})^{-1}\mathbf{\Phi}\,\mathbf{1}_j}\right](\mathbf{\Phi}\,\mathbf{\Phi}^H+\sigma_v^2\,\mathbf{I}_N)^{-1}\mathbf{\Phi}\,\mathbf{\Phi}^T\right\} \\ &= \left[\mathbf{I}_N - \frac{(\mathbf{\Phi}^H)^{\dagger}\mathbf{1}_j\,\boldsymbol{\phi}_j^H}{\boldsymbol{\phi}_j^H(\mathbf{\Phi}^H)^{\dagger}\,\mathbf{1}_j}\right](\mathbf{\Phi}^H)^{\dagger}\mathbf{\Phi}^T = \left[\mathbf{I}_N - (\mathbf{\Phi}^H)^{\dagger}\mathbf{1}_j\,\mathbf{1}_j^T\mathbf{\Phi}^H\right](\mathbf{\Phi}^H)^{\dagger}\mathbf{\Phi}^T \\ &= (\mathbf{\Phi}^H)^{\dagger}\,\mathbf{S}_j\,\mathbf{\Phi}^T \end{aligned} \tag{4.97}$$

where $\mathbf{1}_j \overset{\triangle}{=} [\overbrace{0,\ldots,0}^{j-1},1,0,\ldots,0]^T \in \mathbb{R}^{J\times 1}$ and $\mathbf{S}_j \overset{\triangle}{=} \mathbf{I}_J - \mathbf{1}_j\,\mathbf{1}_j^T \in \mathbb{R}^{J\times J}$. Accounting for (4.97), the asymptotic value of (4.96) is given by

$$\begin{aligned} \lim_{\sigma_v^2\to 0}\text{trace}(\mathbf{\Gamma}_{j,\text{L}}^{H}\,\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j^*}\,\mathbf{\Gamma}_{j,\text{L}}^{*}\,\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j^*}^{*}) &= \text{trace}\left[\mathbf{S}_j\,\mathbf{\Phi}^T(\mathbf{\Phi}^T)^{\dagger}\,\mathbf{S}_j\right] \\ &= \text{trace}\left[\mathbf{\Phi}^T(\mathbf{\Phi}^T)^{\dagger}\,\mathbf{S}_j\right] = J-1\,. \end{aligned} \tag{4.98}$$

As regards the L-SUB receiver, since $\mathbf{\Gamma}_{j,\mathrm{L}} = \mathbf{P}_{j,\mathrm{L}}\,\mathbf{R}^{-1}_{\mathbf{p}_j\mathbf{p}_j} - \gamma_{j,\mathrm{L}}\,\mathbf{V}_n\,\mathbf{V}_n^H$, recalling that $\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j^*} = \overline{\mathbf{\Phi}}_j\,\overline{\mathbf{\Phi}}_j^T$, and observing that $\mathbf{V}_n^H\,\overline{\mathbf{\Phi}}_j = \mathbf{O}_{(N-J)\times(J-1)}$, it follows that

$$\begin{aligned}\mathrm{trace}(&\mathbf{\Gamma}_{j,\mathrm{L}}^H\,\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j^*}\,\mathbf{\Gamma}_{j,\mathrm{L}}^*\,\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j^*}^*)\\ &= \mathrm{trace}[\mathbf{R}^{-1}_{\mathbf{p}_j\mathbf{p}_j}\,\mathbf{P}_{j,\mathrm{L}}^H\,\overline{\mathbf{\Phi}}_j\,\overline{\mathbf{\Phi}}_j^T\,\mathbf{P}_{j,\mathrm{L}}^*\,(\mathbf{R}^*_{\mathbf{p}_j\mathbf{p}_j})^{-1}\,\overline{\mathbf{\Phi}}_j^*\,\overline{\mathbf{\Phi}}_j^H]\\ &= \mathrm{trace}[\mathbf{R}^{-1}_{\mathbf{p}_j\mathbf{p}_j}\,\mathbf{P}_{j,\mathrm{L}}^H\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j^*}\,\mathbf{P}_{j,\mathrm{L}}^*\,(\mathbf{R}^*_{\mathbf{p}_j\mathbf{p}_j})^{-1}\mathbf{R}^*_{\mathbf{p}_j\mathbf{p}_j^*}] \qquad (4.99)\end{aligned}$$

which turns out to be exactly equal to (4.96).

By substituting (4.95) and (4.98) into (4.94), one gets:

$$\mathrm{SINR}_{j,\text{L-SMI}} \triangleq \mathrm{SINR}(\mathbf{w}_{j,\text{L-SMI}}) = \frac{\mathrm{SINR}_{j,\text{L-MOE}}}{1+\frac{N+J-2}{2\,K}\,\mathrm{SINR}_{j,\text{L-MOE}}}\,, \qquad (4.100)$$

$$\mathrm{SINR}_{j,\text{L-SUB}} \triangleq \mathrm{SINR}(\mathbf{w}_{j,\text{L-SUB}}) = \frac{\mathrm{SINR}_{j,\text{L-MOE}}}{1+\frac{2\,(J-1)+(N-J)|1-\gamma_{j,L}\sigma_v^2|}{2\,K}\,\mathrm{SINR}_{j,\text{L-MOE}}}\,. \qquad (4.101)$$

The expression (4.101), as in the WL case, can be further simplified by observing that, for $\sigma_v^2 \to 0$ one has $\gamma_{j,L}\sigma_v^2 \to 1$, hence the trace in (4.95) reduces to $J-1$. By accounting for this observation, for moderate-to-hight values of the SNR, (4.101) can be approximatively written as

$$\mathrm{SINR}_{j,\text{L-SUB}} \triangleq \mathrm{SINR}(\mathbf{w}_{j,\text{L-SUB}}) = \frac{\mathrm{SINR}_{j,\text{L-MOE}}}{1+\frac{J-1}{K}\,\mathrm{SINR}_{j,\text{L-MOE}}}\,. \qquad (4.102)$$

Equations (4.77), (4.80), (4.100) and (4.102) allow one to easily compare the finite-sample performances of WL-MOE and L-MOE receivers. By comparing (4.80) and (4.102) for the subspace receivers, since $\mathrm{SINR}_{j,\text{WL-MOE}} \geq \mathrm{SINR}_{j,\text{L-MOE}}$ by (4.51), it turns out that $\mathrm{SINR}_{j,\text{WL-SUB}} \geq \mathrm{SINR}_{j,\text{L-SUB}}$ for any value of K and for $J \leq N$. A similar conclusion does not hold for the SMI receivers. Indeed, it can be easily proven that, for $J < N$ it results that $\mathrm{SINR}_{j,\text{WL-SMI}} \geq \mathrm{SINR}_{j,\text{L-SMI}}$ only for $K \geq K_{\min}$, where

$$K_{\min} \triangleq \frac{3\,N-J}{2\,(\mathrm{SINR}^{-1}_{j,\text{L-MOE}} - \mathrm{SINR}^{-1}_{j,\text{WL-MOE}})} > 0 \qquad (4.103)$$

is a threshold sample-size. In other words, it can be inferred that, in underloaded scenarios, the WL-SMI receiver assures the expected performance advantage over the L-SMI one only if a sufficient number of samples are processed. This loss of performance is due to the increase of the dimension of the

autocorrelation matrix to be estimated from N to $2\,N$, which entails a diminished estimation accuracy, requiring hence a larger number of data samples for achieving a satisfactory performance, without resorting to subspace concepts.

Another interesting conclusion that can be drawn from (4.77) through (4.102) is that all finite-sample receivers exhibit a SINR saturation effect, i.e., a bit-error-rate (BER) floor, for vanishingly small noise. Indeed, when $\sigma_v^2 \to 0$ and $\mathbf{H}$ is full-column rank ($J \leq 2\,N$), it has been shown in Subsection 4.4.1 that $\text{SINR}_{j,\text{WL-MOE}}$ grows without bound. Thus, accounting for (4.77) and (4.80), we get:

$$\lim_{\sigma_v^2 \to 0} \text{SINR}_{j,\text{WL-SMI}} = \frac{K}{2\,N-1}\,, \quad \lim_{\sigma_v^2 \to 0} \text{SINR}_{j,\text{WL-SUB}} = \frac{K}{J-1}\,, \tag{4.104}$$

which show that, in the high-SNR regime, the performance of the WL-SMI receiver does not depend on the number of users J, whereas the asymptotic value of $\text{SINR}_{j,\text{WL-SUB}}$ is independent of the processing gain N. As regards the linear receivers, if $\sigma_v^2 \to 0$ and $\mathbf{\Phi}$ is full-column rank ($J \leq N$), then $\text{SINR}_{j,\text{L-MOE}} \to +\infty$ (see Subsection 4.4.1), which, accounting for (4.100) and (4.102), implies that

$$\lim_{\sigma_v^2 \to 0} \text{SINR}_{j,\text{L-SMI}} = \frac{2\,K}{N+J-2}\,, \quad \lim_{\sigma_v^2 \to 0} \text{SINR}_{j,\text{L-SUB}} = \frac{K}{J-1}\,. \tag{4.105}$$

It can be seen that, while the WL-SUB and L-SUB receivers exhibit the same asymptotic SINR (for $J \leq N$), the L-SMI receiver for $J < N$ exhibits a better saturation SINR compared with the WL-SMI receiver, for any value of the sample size K. In conclusion, we can state that the advantages of WL receivers could be lost by employing simple estimation methods such as the SMI, whereas it is mandatory to resort to more sophisticated subspace-based methods based on EVD. It is worthwhile to note that in this latter case WL processing incurs an increased computational complexity compared with linear one, due to the increased dimension of the augmented correlation matrix, with respect to the conventional data autocorrelation matrix.

4.5.2 Numerical results

Herein, we present the results of Monte Carlo computer simulations and compare them with the analytical results derived in subsection 4.5.1 [see (4.77), (4.80), (4.100) and (4.102)]. Specifically, in all the experiments, the same simulation setting considered in Example 1 is adopted (downlink scenario and

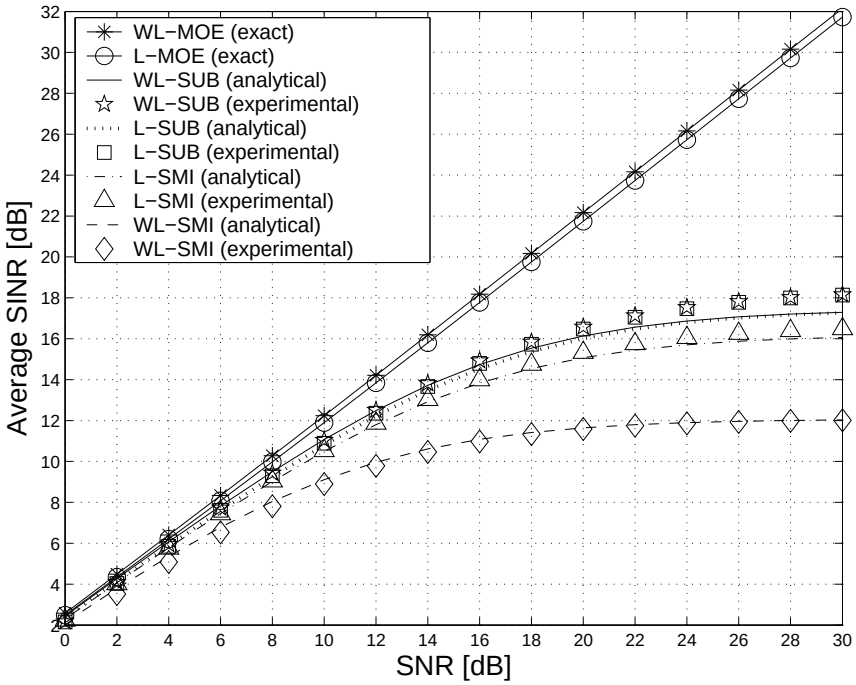

Figure 4.3: Average SINR versus SNR ($J = 10$ users and $K = 500$ symbols).

$N = 16$), with $\theta_1 = \theta_2 = \cdots = \theta_N = 0$ and $\theta_{N+1} = \theta_{N+2} = \cdots = \theta_{2N} = \pi/4$ (we recall that this precoding strategy assures the full-column rank property of the augmented matrix $\mathbf{H}$ in overloaded scenarios). In addition, the symbol vector $\mathbf{b}(k)$ and the additive noise vector $\mathbf{v}(k)$ are generated according to assumptions **(a1)** and **(a2)**. For the sake of comparison, we consider both SMI- and SUB-based data-estimated versions of the L-MOE and WL-MOE receivers (wherein the channel impulse response is assumed to be exactly known), as well as their exact counterparts (wherein, besides the channel impulse response, perfect knowledge of the autocorrelation matrices $\mathbf{R_{rr}}$ and $\mathbf{R_{zz}}$ is assumed). Finally, as performance measure, in addition to the SINR given by (4.24) and averaged over 10^4 Monte Carlo runs, we resort to the average BER at the output of the considered receivers. More specifically, after estimating the receiver weight vectors on the basis of the given data record K, for each run (wherein, besides the channel impulse response, independent sets

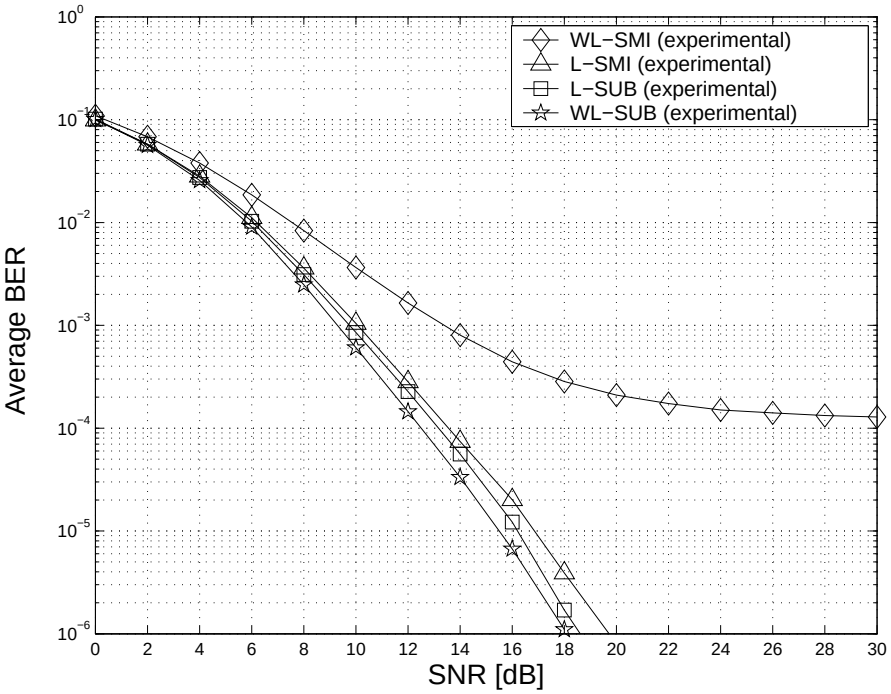

Figure 4.4: Average BER versus SNR ($J = 10$ users and $K = 500$ symbols).

of noise and data sequences are randomly generated), an independent record of $K_{\text{ber}} = 10^3$ symbols is considered to evaluate the BER.

Experiment 4.1 *:*

In this experiment, we evaluate both the (average) SINR and BER performances of the considered receivers as a function of the SNR. The number of users is set equal to $J = 10$ (underloaded system) and the sample size is kept fixed to $K = 500$ symbols. Let us first consider the SINR performances, which are reported in Fig. 4.3. It can be seen that the analytical expressions (4.77), (4.80), (4.100) and (4.102) for the data-estimated linear and WL receivers agree very well with their corresponding simulation results, for all values of the SNR. In particular, in this underloaded scenario, while the L-SUB and WL-SUB receivers perform comparably, the WL-SMI receiver pays a significant performance loss with respect to the L-SMI one. Indeed, in the high-SNR region, the difference between the saturation values of $\text{SINR}_{1,\text{L-SMI}}$ and $\text{SINR}_{1,\text{WL-SMI}}$ is about 4 dB, which is in good agreement with (4.104)

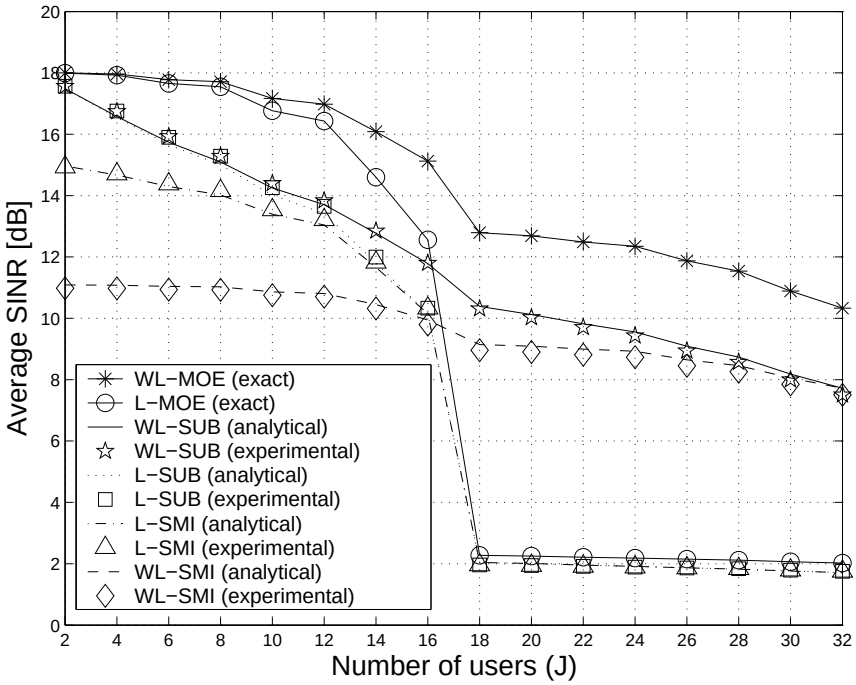

Figure 4.5: Average SINR versus number of users ($K = 500$ symbols and SNR $= 15$ dB).

and (4.105). The unsatisfactory performance of the WL-SMI receiver is also apparent from Fig. 4.4, which depicts the BER curves of the data-estimated receivers under comparison. It is evident that the curves of the WL-SUB, L-SUB and L-SMI receivers go down very quickly as the SNR increases, thus assuring a huge performance gain with respect to the WL-SMI receiver, which instead exhibits a marked BER floor.

Experiment 4.2 *:*

Fig. 4.5 reports the SINR as a function of the number of users J. The SNR is set equal to 15 dB and $K = 500$ symbols are considered.

Besides confirming the very good agreement between analytical and experimental results for all the data-estimated receivers, results of Fig. 4.5 show that the performances of all the linear receivers worsen very quickly when the system tends to be overloaded, i.e., J approaches $N = 16$. Beyond this value, the WL receivers assure a significant performance gain with respect to their

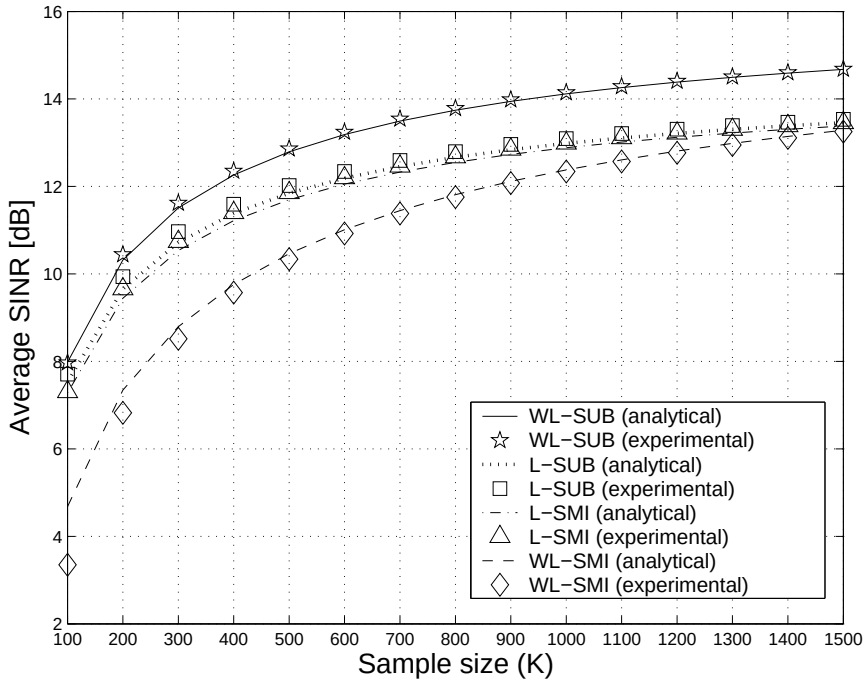

Figure 4.6: Average SINR versus sample size K ($J = 14$ users and SNR $= 15$ dB).

corresponding linear counterparts. Loosely speaking, this indicates the ability of the WL-MOE receiver to accommodate twice the number of users of the L-MOE receiver.

Experiment 4.3 *:* In this last experiment, we report the SINR performances of the considered data-estimated receivers as a function of the sample size K. The SNR is set equal to 15 dB and $J = 14$ users (underloaded system) are considered. It can be observed from Fig. 4.6 that the accuracy of the formulas (4.77), (4.80), (4.100) and (4.102) improves as K increases. Additionally, it is worth observing that the WL-SUB receiver outperforms the L-SUB one, for all the considered values of K. In contrast, the WL-SMI receiver performs worse than its corresponding linear counterpart, by approaching the curve of the L-SMI receiver only when the sample size K is as large as 1500 symbols, which agrees very well with the value $K_{\min} = 1686$ predicted by (4.103).

4.5.3 Performance analysis of WL-MOE and L-MOE receivers with channel estimation

In this subsection, the first-order perturbation analysis carried out in subsection 4.5.1 is extended to incorporate the effects of errors due to subspace-based blind CE on the synthesis of the L- and WL-MOE receivers. As we have noted in the section 4.5, when and the autocorrelation matrix and the channel are estimated from data, as performance measure we consider the SINR defined in (4.23) because it is quite general. In the previous subsections, instead, we have consider the SINR defined in (4.24) because, it can be shown that the SINR (4.23) reduces to the SINR (4.24) and in the ideal case (ACM and channel perfectly known) and when only the ACM is estimated from a finite sample-size, under the assumption that the channel is exactly known.

Performance analysis of WL-MOE with channel estimation

Let us start with WL-MOE receiver. Preliminarily, in the subsection 4.3.2 we have described the subspace-based channel estimation. In particular, with reference to WL-MOE receiver, we have shown that the unknown vector $\boldsymbol{\varrho}_j$ can be obtained as solution of the linear system (4.47), provided that this system uniquely characterizes the channel coefficients for each users. To this end it is required that the condition $(\mathbf{c3})$ is satisfied (see subsection 4.3.2 for details). This condition, from the point of view of the jth user, necessarily imposes that the maximum number of users supported by the system is smaller than the maximum number (see eq. (4.61)) $2\,N$ of users of the known-channel case

$$J_{\text{max},WL} = 2\,(N - L_j) \leq 2\,N. \tag{4.106}$$

Moreover, in the subsection 4.3.2, we have shown that when $\mathbf{R}_{\mathbf{zz}}$ is estimated from a finite sample size, a channel estimate can be obtained as in (4.107), that we report here for completeness:

$$\begin{aligned}\widehat{\boldsymbol{\varrho}}_j &= \underset{\mathbf{x}\in\mathbb{R}^{2L_j}}{\operatorname{argmin}} \|\widehat{\mathbf{U}}_n^H\, \boldsymbol{\mathcal{C}}_j\, \mathbf{T}_j\, \mathbf{x}\|^2 \\ &= \underset{\mathbf{x}\in\mathbb{R}^{2L_j}}{\operatorname{argmin}} \left(\mathbf{x}^H\, \mathbf{T}_j^H\, \boldsymbol{\mathcal{C}}_j^H\, \widehat{\mathbf{U}}_n\, \widehat{\mathbf{U}}_n^H\, \boldsymbol{\mathcal{C}}_j\, \mathbf{T}_j\, \mathbf{x}\right) \text{ , subject to } \|\mathbf{x}\|^2 = 1\,,\end{aligned} \tag{4.107}$$

whose solution [56] is given by the eigenvector associated with the smallest eigenvalue of the matrix $\mathbf{T}_j^H\, \widehat{\boldsymbol{\mathcal{Q}}}_{j,\text{WL}}\, \mathbf{T}_j \in \mathbb{C}^{2L_j\times 2L_j}$, with $\widehat{\boldsymbol{\mathcal{Q}}}_{j,\text{WL}} \triangleq$

$\mathcal{C}_j^H\,\widehat{\mathbf{U}}_n\,\widehat{\mathbf{U}}_n^H\,\mathcal{C}_j \in \mathbb{C}^{2L_j\times 2L_j}$. By substituting (4.107) in (4.46) the resulting estimate of the received signature is

$$\widehat{\mathbf{h}}_j = \widetilde{\alpha}_j\,\mathcal{C}_j\,\mathbf{T}_j\,\widehat{\boldsymbol{\varrho}}_j \tag{4.108}$$

To obtain the WL-SMI and WL-SUB receivers with channel estimation, it is necessary to substitute in (4.65) and (4.66), (4.108):

$$\mathbf{f}_{j,\text{WL-SMI-CE}} \triangleq (\widehat{\mathbf{h}}_j^H\,\widehat{\mathbf{R}}_{\mathbf{zz}}^{-1}\,\widehat{\mathbf{h}}_j)^{-1}\,\widehat{\mathbf{R}}_{\mathbf{zz}}^{-1}\,\widehat{\mathbf{h}}_j, \tag{4.109}$$

$$\mathbf{f}_{j,\text{WL-SUB-CE}} \triangleq (\widehat{\mathbf{h}}_j^H\,\widehat{\mathbf{U}}_s\,\widehat{\boldsymbol{\Lambda}}_s^{-1}\,\widehat{\mathbf{U}}_s^H\,\widehat{\mathbf{h}}_j)^{-1}\,\widehat{\mathbf{U}}_s\,\widehat{\boldsymbol{\Lambda}}_s^{-1}\,\widehat{\mathbf{U}}_s^H\,\widehat{\mathbf{h}}_j. \tag{4.110}$$

A remark is in order about knowledge of the real scalar $\widetilde{\alpha}_j$ and of the sign inversion inherent to channel estimate $\widehat{\boldsymbol{\varrho}}_j$, which are needed to correctly write the estimated signature $\widehat{\mathbf{h}}_j$. These parameters cannot be estimated by means of blind techniques based on second-order statistics; in practice, they can be recovered resorting to automatic gain control and differential modulation or, more robustly, by using a few training symbols. It should be noted, however, that their possible inaccurate knowledge merely introduce a *real* multiplicative factor in the expressions of WL-SMI-CE and WL-SUB-CE receivers, which does not affect SINR calculation based on (4.23). Therefore, to simplify matters, we assume in the sequel that they are are known exactly.

As in the subsection 4.5.1, we adopt a first-order perturbative approach [76, 77] to model all estimation errors. In the following, in order to carry out the analysis in an unified framework, we denote with $\widehat{\mathbf{f}}_j$ *any* data-estimated WL-MOE receiver, i.e, $\widehat{\mathbf{f}}_j = \mathbf{f}_{j,\text{WL-SMI-CE}}$ or $\widehat{\mathbf{f}}_j = \mathbf{f}_{j,\text{WL-SUB-CE}}$, and set $\mathbf{f}_j = \mathbf{f}_{j,\text{WL-MOE}}$, where $\mathbf{f}_{j,\text{WL-MOE}}$ is the ideal WL-MOE receiver given by (4.27) or (4.62). When $\widehat{\mathbf{f}}_j$ is employed, accounting for (4.22), it can be shown that (4.23) can be expanded as

$$\text{SINR}(\widehat{\mathbf{f}}_j) = \frac{\text{E}^2\{\text{Re}[\widehat{\mathbf{f}}_j^H\,\mathbf{h}_j]\}}{\text{E}\{\text{Re}^2[\widehat{\mathbf{f}}_j^H\,\mathbf{q}_j(k)]\} + \text{Var}\{\text{Re}[\widehat{\mathbf{f}}_j^H\,\mathbf{h}_j]\}}\,. \tag{4.111}$$

Since $\widehat{\mathbf{f}}_j$, $\mathbf{h}_j$ and $\mathbf{q}_j(k)$ exhibit the CS property, the real parts in (4.111) can be omitted, thus yielding

$$\text{SINR}(\widehat{\mathbf{f}}_j) = \frac{\text{E}^2[\widehat{\mathbf{f}}_j^H\,\mathbf{h}_j]}{\text{E}\{|\widehat{\mathbf{f}}_j^H\,\mathbf{q}_j(k)|^2\} + \text{Var}[\widehat{\mathbf{f}}_j^H\,\mathbf{h}_j]}\,. \tag{4.112}$$

Following the perturbative approach, the vectors $\widehat{\mathbf{f}}_j$ and $\widehat{\mathbf{h}}_j$ are expressed as $\widehat{\mathbf{f}}_j = \mathbf{f}_j + \boldsymbol{\delta}\mathbf{f}_j$ and $\widehat{\mathbf{h}}_j = \mathbf{h}_j + \boldsymbol{\delta}\mathbf{h}_j$, respectively, where $\boldsymbol{\delta}\mathbf{f}_j$ and $\boldsymbol{\delta}\mathbf{h}_j$ are *small* (i.e., $\|\boldsymbol{\delta}\mathbf{f}_j\| \ll 1$ and $\|\boldsymbol{\delta}\mathbf{h}_j\| \ll 1$) and *zero-mean* CS perturbation terms. Thus, we have $\widehat{\mathbf{f}}_j^H \mathbf{h}_j = \mathbf{f}_j^H \mathbf{h}_j + \boldsymbol{\delta}\mathbf{f}_j^H \mathbf{h}_j = 1 + \boldsymbol{\delta}\mathbf{f}_j^H \mathbf{h}_j$, since, from (4.28), $\mathbf{f}_j^H \mathbf{h}_j = 1$. The term $\mathrm{Var}[\widehat{\mathbf{f}}_j^H \mathbf{h}_j]$ in (4.112) is equal to

$$\mathrm{Var}[\widehat{\mathbf{f}}_j^H \mathbf{h}_j] = \mathrm{E}_{\boldsymbol{\delta}\mathbf{f}_j}[(\widehat{\mathbf{f}}_j^H \mathbf{h}_j)^2] - \mathrm{E}^2_{\boldsymbol{\delta}\mathbf{f}_j}[\widehat{\mathbf{f}}_j^H \mathbf{h}_j]. \tag{4.113}$$

Moreover, denoting with $\mathrm{E}_{\boldsymbol{\delta}\mathbf{f}_j}[\cdot]$ the average with respect to (w.r.t.) $\boldsymbol{\delta}\mathbf{f}_j$, since $\boldsymbol{\delta}\mathbf{f}_j$ is zero-mean and $\boldsymbol{\delta}\mathbf{f}_j^H \mathbf{h}_j$ is a real-valued scalar, it turns out that $\mathrm{E}_{\boldsymbol{\delta}\mathbf{f}_j}[\widehat{\mathbf{f}}_j^H \mathbf{h}_j] = 1$ (see subsection 4.5.1) and that

$$\begin{aligned}\mathrm{E}_{\boldsymbol{\delta}\mathbf{f}_j}[(\widehat{\mathbf{f}}_j^H \mathbf{h}_j)^2] &= \mathrm{E}_{\boldsymbol{\delta}\mathbf{f}_j}[1 + 2\boldsymbol{\delta}\mathbf{f}_j^H \mathbf{h}_j + (\boldsymbol{\delta}\mathbf{f}_j^H \mathbf{h}_j)^2] \\ &= 1 + \mathrm{E}_{\boldsymbol{\delta}\mathbf{f}_j}[|\boldsymbol{\delta}\mathbf{f}_j^H \mathbf{h}_j|^2] = 1 + \mathrm{E}_{\boldsymbol{\delta}\mathbf{f}_j}[\boldsymbol{\delta}\mathbf{f}_j^H \mathbf{h}_j \mathbf{h}_j^H \boldsymbol{\delta}\mathbf{f}_j]\end{aligned} \tag{4.114}$$

Therefore, (4.113) can be simplify

$$\mathrm{Var}[\widehat{\mathbf{f}}_j^H \mathbf{h}_j] = \mathrm{E}_{\boldsymbol{\delta}\mathbf{f}_j}[\boldsymbol{\delta}\mathbf{f}_j^H \mathbf{h}_j \mathbf{h}_j^H \boldsymbol{\delta}\mathbf{f}_j]. \tag{4.115}$$

By substituting it in (4.112) leads to:

$$\mathrm{SINR}(\widehat{\mathbf{f}}_j) = \frac{1}{\mathrm{E}_{\widehat{\mathbf{f}}_j \mathbf{q}_j}[\widehat{\mathbf{f}}_j^H \mathbf{q}_j(k)\mathbf{q}_j^H(k)\widehat{\mathbf{f}}_j] + \mathrm{E}_{\boldsymbol{\delta}\mathbf{f}_j}[\boldsymbol{\delta}\mathbf{f}_j^H \mathbf{h}_j \mathbf{h}_j^H \boldsymbol{\delta}\mathbf{f}_j]}, \tag{4.116}$$

where $\mathrm{E}_{\widehat{\mathbf{f}}_j \mathbf{q}_j}[\cdot]$ denotes joint average w.r.t. $\widehat{\mathbf{f}}_j$ and $\mathbf{q}_j(k)$. Under the simplifying and reasonable assumption [75] that $\widehat{\mathbf{f}}_j$ is independent of $\mathbf{q}_j(k)$, $\mathrm{E}_{\widehat{\mathbf{f}}_j \mathbf{q}_j}[\widehat{\mathbf{f}}_j^H \mathbf{q}_j(k)\mathbf{q}_j^H(k)\widehat{\mathbf{f}}_j] = \mathrm{E}_{\widehat{\mathbf{f}}_j}\{\widehat{\mathbf{f}}_j^H \mathrm{E}_{\mathbf{q}_j}[\mathbf{q}_j(k)\mathbf{q}_j^H(k)]\widehat{\mathbf{f}}_j\} = \mathrm{E}_{\widehat{\mathbf{f}}_j}[\widehat{\mathbf{f}}_j^H \mathbf{R}_{\mathbf{q}_j\mathbf{q}_j} \widehat{\mathbf{f}}_j]$, which, accounting for $\widehat{\mathbf{f}}_j = \mathbf{f}_j + \boldsymbol{\delta}\mathbf{f}_j$ and $\mathrm{E}[\boldsymbol{\delta}\mathbf{f}_j] = 0$ leads to

$$\mathrm{E}_{\widehat{\mathbf{f}}_j \mathbf{q}_j}[\widehat{\mathbf{f}}_j^H \mathbf{q}_j(k)\mathbf{q}_j^H(k)\widehat{\mathbf{f}}_j] = \mathrm{E}_{\widehat{\mathbf{f}}_j}[\widehat{\mathbf{f}}_j^H \mathbf{R}_{\mathbf{q}_j\mathbf{q}_j} \widehat{\mathbf{f}}_j] = \mathbf{f}_j^H \mathbf{R}_{\mathbf{q}_j\mathbf{q}_j} \mathbf{f}_j + \mathrm{E}_{\boldsymbol{\delta}\mathbf{f}_j}[\boldsymbol{\delta}\mathbf{f}_j^H \mathbf{R}_{\mathbf{q}_j\mathbf{q}_j} \boldsymbol{\delta}\mathbf{f}_j]. \tag{4.117}$$

By substituting (4.117) into (4.116), one has

$$\mathrm{SINR}(\widehat{\mathbf{f}}_j) = \frac{1}{\mathbf{f}_j^H \mathbf{R}_{\mathbf{q}_j\mathbf{q}_j} \mathbf{f}_j + \mathrm{E}_{\boldsymbol{\delta}\mathbf{f}_j}[\boldsymbol{\delta}\mathbf{f}_j^H \mathbf{R}_{\mathbf{q}_j\mathbf{q}_j} \boldsymbol{\delta}\mathbf{f}_j] + \mathrm{E}_{\boldsymbol{\delta}\mathbf{f}_j}[\boldsymbol{\delta}\mathbf{f}_j^H \mathbf{h}_j \mathbf{h}_j^H \boldsymbol{\delta}\mathbf{f}_j]}. \tag{4.118}$$

Since $\mathrm{E}_{\boldsymbol{\delta f}_j}[\boldsymbol{\delta f}_j^H \mathbf{R}_{\mathbf{q}_j\mathbf{q}_j} \boldsymbol{\delta f}_j] + \mathrm{E}_{\boldsymbol{\delta f}_j}[\boldsymbol{\delta f}_j^H \mathbf{h}_j \mathbf{h}_j^H \boldsymbol{\delta f}_j] = \mathrm{E}_{\boldsymbol{\delta f}_j}[\boldsymbol{\delta f}_j^H \mathbf{R}_{\mathbf{zz}} \boldsymbol{\delta f}_j]$, noting also that, according to (4.49), $\mathbf{f}_j^H \mathbf{R}_{\mathbf{q}_j\mathbf{q}_j} \mathbf{f}_j = (\mathrm{SINR}_{j,\text{WL-MOE}})^{-1}$ we obtain the compact expression

$$\mathrm{SINR}(\widehat{\mathbf{f}}_j) = \frac{\mathrm{SINR}_{j,\text{WL-MOE}}}{1 + \mathrm{SINR}_{j,\text{WL-MOE}}\, \mathrm{E}_{\boldsymbol{\delta f}_j}[\boldsymbol{\delta f}_j^H\, \mathbf{R}_{\mathbf{zz}}\, \boldsymbol{\delta f}_j]}\,, \tag{4.119}$$

where only the average w.r.t. $\boldsymbol{\delta f}_j$ is left to be evaluated. To proceed further, explicit expressions for the perturbation $\boldsymbol{\delta f}_j$ are needed for both the WL-SMI-CE and WL-SUB-CE receivers.

Lemma 4.4 *Let* $\approx$ *denote* first-order equality, *the first-order perturbation term of the WL-SMI-CE and WL-SUB-CE receivers can be expressed as*

$$\boldsymbol{\delta f}_j \approx \underbrace{-\boldsymbol{\Gamma}_{j,WL}\, \widehat{\mathbf{r}}_{\mathbf{q}_j b_j}}_{\boldsymbol{\delta f}_j^{(1)}} + \underbrace{\boldsymbol{\Delta}_{j,WL} \boldsymbol{\delta h}_j}_{\boldsymbol{\delta f}_j^{(2)}} = \boldsymbol{\delta f}_j^{(1)} + \boldsymbol{\delta f}_j^{(2)} \tag{4.120}$$

where $\widehat{\mathbf{r}}_{\mathbf{q}_j b_j} \triangleq \frac{1}{K}\sum_{k=0}^{K-1} \mathbf{q}_j(k)\, b_j(k) \in \mathbb{C}^{2N}$ *is the sample estimate of the cross-correlation between the disturbance vector* $\mathbf{q}_j(k)$ *and the desired symbol* $b_j(k)$*, whereas*

$$\boldsymbol{\Gamma}_{j,WL} \triangleq \begin{cases} \mathbf{P}_{j,\text{WL}}\, \mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}^{-1}, & \text{(WL-SMI-CE)} \\ \mathbf{P}_{j,\text{WL}}\, \mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}^{-1} - \gamma_{j,\text{WL}}\, \mathbf{U}_n \mathbf{U}_n^H, & \text{(WL-SUB-CE)} \end{cases} \tag{4.121}$$

$$\boldsymbol{\Delta}_{j,WL} \triangleq \begin{cases} (\mathbf{h}_j^H \mathbf{R}_{\mathbf{zz}}^{-1} \mathbf{h}_j)^{-1} \mathbf{R}_{\mathbf{zz}}^{-1} - 2\, \mathbf{f}_j\, \mathbf{f}_j^H, & \text{(WL-SMI-CE)} \\ (\mathbf{h}_j^H \mathbf{R}_{\mathbf{zz}}^{-1} \mathbf{h}_j)^{-1} \mathbf{U}_s \boldsymbol{\Lambda}_s^{-1} \mathbf{U}_s^H - 2\, \mathbf{f}_j\, \mathbf{f}_j^H, & \text{(WL-SUB-CE)} \end{cases} \tag{4.122}$$

with $\mathbf{P}_{j,\text{WL}} \triangleq \mathbf{I}_{2N} - (\mathbf{h}_j^H\, \mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}^{-1}\, \mathbf{h}_j)^{-1}\, \mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}^{-1}\, \mathbf{h}_j\, \mathbf{h}_j^H = \mathbf{I}_{2N} - \mathbf{f}_j\, \mathbf{h}_j^H \in \mathbb{C}^{2N\times 2N}$ *denoting an oblique projection matrix [75] and* $\gamma_{j,\text{WL}} \triangleq \sigma_v^{-2} + (\mathbf{h}_j^H\, \mathbf{R}_{\mathbf{zz}}^{-1}\, \mathbf{h}_j)^{-1} \mathbf{h}_j^H \mathbf{U}_s\, \boldsymbol{\Omega}_{WL}^{-1}\, \mathbf{U}_s^H \mathbf{R}_{\mathbf{zz}}^{-1}\, \mathbf{h}_j > 0$*, while the diagonal matrix* $\boldsymbol{\Omega}_{WL} \triangleq diag(\lambda_1, \lambda_2, \ldots, \lambda_J) \in \mathbb{R}^{J\times J}$ *collects the nonzero eigenvalues of* $\mathbf{H}\,\mathbf{H}^H$.

Proof: See Appendix B.5.

It should be noted that the quantities $\mathbf{P}_{j,\text{WL}}$, $\gamma_{j,\text{WL}}$ and $\boldsymbol{\Omega}_{\text{WL}}$ are already defined in Lemma 4.2, nevertheless they are here report for completeness.

Moreover we emphasize that $\boldsymbol{\delta}\mathbf{f}_j^{(1)}$ and $\boldsymbol{\delta}\mathbf{f}_j^{(2)}$ represent the perturbations due to estimation of $\mathbf{R}_{\mathbf{zz}}$ and $\mathbf{h}_j$, respectively; indeed, a comparison shows that the expression of $\boldsymbol{\delta}\mathbf{f}_j^{(1)}$ is the same as that reported in Lemma 4.2. In order to characterize the perturbation term $\boldsymbol{\delta}\mathbf{f}_j^{(2)}$, it is necessary to evaluate the perturbation $\boldsymbol{\delta}\mathbf{h}_j$ associated to the subspace-based CE procedure given by (4.107).

Lemma 4.5 *Given the estimate* $\widehat{\mathbf{h}}_j = \widetilde{\alpha}_j \, \boldsymbol{\mathcal{C}}_j \, \mathbf{T}_j \, \widehat{\boldsymbol{\varrho}}_j$ *of the signature* $\mathbf{h}_j$*, where the channel estimate* $\widehat{\boldsymbol{\varrho}}_j$ *is the solution of* (4.107)*, the perturbation* $\boldsymbol{\delta}\mathbf{h}_j$ *can be expressed as*

$$\boldsymbol{\delta}\mathbf{h}_j \approx \boldsymbol{\Pi}_{j,WL} \, \widehat{\mathbf{r}}_{\mathbf{q}_j b_j} \, , \tag{4.123}$$

where $\boldsymbol{\Pi}_{j,WL} \stackrel{\triangle}{=} (\mathbf{h}_j^H \, \mathbf{U}_s \, \boldsymbol{\Omega}_{WL}^{-1} \, \mathbf{U}_s^H \, \mathbf{h}_j) \, \boldsymbol{\mathcal{C}}_j \, \boldsymbol{\mathcal{Q}}_{j,WL}^{\dagger} \, \boldsymbol{\mathcal{C}}_j^H \, \mathbf{U}_n \, \mathbf{U}_n^H \in \mathbb{C}^{2N \times 2N}$*, with* $\boldsymbol{\Omega}_{WL}$ *and* $\widehat{\mathbf{r}}_{\mathbf{q}_j b_j}$ *defined in Lemma 4.4, and* $\boldsymbol{\mathcal{Q}}_{j,WL} \stackrel{\triangle}{=} \boldsymbol{\mathcal{C}}_j^H \, \mathbf{U}_n \, \mathbf{U}_n^H \, \boldsymbol{\mathcal{C}}_j \in \mathbb{C}^{2L_j \times 2L_j}$.

Proof: See Appendix B.6.

Accounting for (4.123) and Lemma 4.4, the overall perturbation of the WL-SMI-CE and WL-SUB-CE weight vectors can be expressed as a *linear* function of $\widehat{\mathbf{r}}_{\mathbf{q}_j b_j}$, as summarized by the following Lemma:

Lemma 4.6 *The first-order overall perturbation term* $\boldsymbol{\delta}\mathbf{f}_j = \boldsymbol{\delta}\mathbf{f}_j^{(1)} + \boldsymbol{\delta}\mathbf{f}_j^{(2)}$ *of the WL-SMI-CE and WL-SUB-CE receivers can be expressed in a unified manner as*

$$\boldsymbol{\delta}\mathbf{f}_j \approx \boldsymbol{\Sigma}_{j,WL} \, \widehat{\mathbf{r}}_{\mathbf{q}_j b_j} \, , \tag{4.124}$$

where $\boldsymbol{\Sigma}_{j,WL} \stackrel{\triangle}{=} -\boldsymbol{\Gamma}_{j,WL} + \boldsymbol{\Delta}_{j,WL} \, \boldsymbol{\Pi}_{j,WL} \in \mathbb{C}^{2N \times 2N}$*, with* $\boldsymbol{\Gamma}_{j,WL} \in \mathbb{C}^{2N \times 2N}$*,* $\boldsymbol{\Delta}_{j,WL} \in \mathbb{C}^{2N \times 2N}$ *and* $\widehat{\mathbf{r}}_{\mathbf{q}_j b_j}$ *given by Lemma 4.4, whereas* $\boldsymbol{\Pi}_{j,WL} \in \mathbb{C}^{2N \times 2N}$ *has been defined in Lemma 4.5.*

It should be observed that Lemma 4.6 provides a compact characterization of the overall perturbation $\boldsymbol{\delta}\mathbf{f}_j$, which is obtained under the simplifying assumption [75] that the error in estimating $\mathbf{R}_{\mathbf{zz}}$ is mainly due to the term $\widehat{\mathbf{r}}_{\mathbf{q}_j b_j}$. Equipped with such a nice result, we are now in the position to evaluate the average $\mathrm{E}_{\boldsymbol{\delta}\mathbf{f}_j}[\boldsymbol{\delta}\mathbf{f}_j^H \mathbf{R}_{\mathbf{zz}} \, \boldsymbol{\delta}\mathbf{f}_j]$ at the denominator of (4.119). Dropping the subscript $\boldsymbol{\delta}\mathbf{f}_j$ in $\mathrm{E}_{\boldsymbol{\delta}\mathbf{f}_j}[\cdot]$ for notational simplicity, by accounting for (4.124) and

using the trace identity, we have

$$\begin{aligned}\mathrm{E}[\boldsymbol{\delta}\mathbf{f}_j^H \mathbf{R}_{\mathbf{zz}}\, \boldsymbol{\delta}\mathbf{f}_j] &= \mathrm{E}[\widehat{\mathbf{r}}_{\mathbf{q}_j b_j}^H\, \boldsymbol{\Sigma}_{j,\mathrm{WL}}^H\, \mathbf{R}_{\mathbf{zz}}\, \boldsymbol{\Sigma}_{j,\mathrm{WL}}\, \widehat{\mathbf{r}}_{\mathbf{q}_j b_j}] \\ &= \mathrm{trace}\left\{\boldsymbol{\Sigma}_{j,\mathrm{WL}}^H\, \mathbf{R}_{\mathbf{zz}}\, \boldsymbol{\Sigma}_{j,\mathrm{WL}}\, \mathrm{E}[\widehat{\mathbf{r}}_{\mathbf{q}_j b_j}\, \widehat{\mathbf{r}}_{\mathbf{q}_j b_j}^H]\right\} \\ &= \frac{1}{K}\mathrm{trace}\left\{\boldsymbol{\Sigma}_{j,\mathrm{WL}}^H\, \mathbf{R}_{\mathbf{zz}}\, \boldsymbol{\Sigma}_{j,\mathrm{WL}}\, \mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}\right\}, \end{aligned} \tag{4.125}$$

where, moreover, we have used the equality (4.74) $\mathrm{E}[\widehat{\mathbf{r}}_{\mathbf{q}_j b_j}\, \widehat{\mathbf{r}}_{\mathbf{q}_j b_j}^H] = \frac{1}{K}\, \mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}$. Therefore, by substituting (4.125) in (4.119), we get

$$\mathrm{SINR}(\widehat{\mathbf{f}}_j) = \frac{\mathrm{SINR}_{j,\text{WL-MOE}}}{1 + \frac{\mathrm{trace}(\boldsymbol{\Sigma}_{j,\mathrm{WL}}^H\, \mathbf{R}_{\mathbf{zz}}\, \boldsymbol{\Sigma}_{j,\mathrm{WL}}\, \mathbf{R}_{\mathbf{q}_j\mathbf{q}_j})}{K}\, \mathrm{SINR}_{j,\text{WL-MOE}}}\,. \tag{4.126}$$

The final result is obtained by evaluating the trace term in (4.126), on the basis of the different expressions for $\boldsymbol{\Sigma}_{j,\mathrm{WL}}$ given by Lemmas 4.4–4.6. In order to do this, it is convenient to consider the SMI and SUB cases separately. With reference to the WL-SMI-CE receiver, it is shown in Appendix B.7 that

$$\begin{aligned}\mathrm{trace}(\boldsymbol{\Sigma}_{j,\mathrm{WL}}^H\, \mathbf{R}_{\mathbf{zz}}\, \boldsymbol{\Sigma}_{j,\mathrm{WL}}\, \mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}) &= (2\,N-1) - 2\,\zeta_{j,\mathrm{WL}}(2\,L_j-1) \\ &\quad + \zeta_{j,\mathrm{WL}}^2\, \sigma_v^2\, \mathrm{trace}(\mathbf{R}_{\mathbf{zz}}^{-1}\, \boldsymbol{\mathcal{C}}_j\, \boldsymbol{\mathcal{Q}}_{j,\mathrm{WL}}^{\dagger}\, \boldsymbol{\mathcal{C}}_j^H), \end{aligned} \tag{4.127}$$

where $\zeta_{j,\mathrm{WL}} \stackrel{\triangle}{=} (\mathbf{h}_j^H\, \mathbf{R}_{\mathbf{zz}}^{-1}\, \mathbf{h}_j)^{-1}\, \mathbf{h}_j^H\, \mathbf{U}_s\, \boldsymbol{\Omega}_{\mathrm{WL}}^{-1}\, \mathbf{U}_s^H\, \mathbf{h}_j > 0$. Instead, as regards the WL-SUB-CE receiver, it is shown in Appendix B.7 that

$$\begin{aligned}\mathrm{trace}(\boldsymbol{\Sigma}_{j,\mathrm{WL}}^H\, \mathbf{R}_{\mathbf{zz}}\, \boldsymbol{\Sigma}_{j,\mathrm{WL}}\, \mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}) &= (J-1) + (2\,N-J)|1-\gamma_{j,\mathrm{WL}}\, \sigma_v^2|^2 \\ &\quad - \zeta_{j,\mathrm{WL}}^2(2\,L_j-1) + \zeta_{j,\mathrm{WL}}^2\, \sigma_v^2\, \mathrm{trace}(\mathbf{R}_{\mathbf{zz}}^{-1}\, \boldsymbol{\mathcal{C}}_j\, \boldsymbol{\mathcal{Q}}_{j,\mathrm{WL}}^{\dagger}\, \boldsymbol{\mathcal{C}}_j^H)\,. \end{aligned} \tag{4.128}$$

The trace expressions (4.127) and (4.128) are still too complicated to allow for a simple discussion, but they can be considerably simplified in the high-SNR region, i.e., by studying their behavior as $\sigma_v^2 \to 0$. Let us first examine the trace term, which is present in both (4.127) and (4.128). One has

$$\begin{aligned}\sigma_v^2\, \mathrm{trace}(\mathbf{R}_{\mathbf{zz}}^{-1}\, \boldsymbol{\mathcal{C}}_j\, \boldsymbol{\mathcal{Q}}_{j,\mathrm{WL}}^{\dagger}\, \boldsymbol{\mathcal{C}}_j^H) &= \sigma_v^2\, \mathrm{trace}[(\mathbf{U}_s\boldsymbol{\Lambda}_s^{-1}\mathbf{U}_s^H + \sigma_v^{-2}\mathbf{U}_n\mathbf{U}_n^H) \\ &\quad \boldsymbol{\mathcal{C}}_j\, \boldsymbol{\mathcal{Q}}_{j,\mathrm{WL}}^{\dagger}\, \boldsymbol{\mathcal{C}}_j^H] = \sigma_v^2\, \mathrm{trace}(\mathbf{U}_s\boldsymbol{\Lambda}_s^{-1}\mathbf{U}_s^H\, \boldsymbol{\mathcal{C}}_j\, \boldsymbol{\mathcal{Q}}_{j,\mathrm{WL}}^{\dagger}\, \boldsymbol{\mathcal{C}}_j^H) \\ &\quad + \mathrm{trace}(\mathbf{U}_n\mathbf{U}_n^H\, \boldsymbol{\mathcal{C}}_j\, \boldsymbol{\mathcal{Q}}_{j,\mathrm{WL}}^{\dagger}\, \boldsymbol{\mathcal{C}}_j^H). \end{aligned} \tag{4.129}$$

Therefore, for $\sigma_v^2 \to 0$, observing that $\mathbf{\Lambda}_s^{-1} \to \mathbf{\Omega}_{\text{WL}}^{-1}$ and using also the trace properties, one has

$$\lim_{\sigma_v^2 \to 0} \sigma_v^2 \, \text{trace}(\mathbf{R}_{\mathbf{zz}}^{-1} \, \boldsymbol{\mathcal{C}}_j \, \boldsymbol{\mathcal{Q}}_{j,\text{WL}}^{\dagger} \, \boldsymbol{\mathcal{C}}_j^H) = \text{trace}(\underbrace{\boldsymbol{\mathcal{Q}}_{j,\text{WL}}^{\dagger} \, \boldsymbol{\mathcal{C}}_j^H \, \mathbf{U}_n \, \mathbf{U}_n^H \, \boldsymbol{\mathcal{C}}_j}_{\boldsymbol{\mathcal{Q}}_{j,\text{WL}}}) = 2L_j - 1, \tag{4.130}$$

where we refer to Appendix B.7 for a formal proof of the result $\text{trace}(\boldsymbol{\mathcal{Q}}_{j,\text{WL}}^{\dagger} \boldsymbol{\mathcal{Q}}_{j,\text{WL}}) = 2L_j - 1$. In addition, as $\sigma_v^2 \to 0$, it can be easily checked that $\gamma_{j,\text{WL}} \, \sigma_v^2 \to 1$ and $\zeta_{j,\text{WL}} \to 1$. Consequently, accounting for (4.126)–(4.128) and (4.130), the SINR behavior in the high-SNR region of the WL-SMI-CE and WL-SUB-CE receivers is (approximately) governed by

$$\text{SINR}_{j,\text{WL-SMI-CE}} \stackrel{\triangle}{=} \text{SINR}(\mathbf{f}_{j,\text{WL-SMI-CE}}) = \frac{\text{SINR}_{j,\text{WL-MOE}}}{1 + \frac{2(N-L_j)}{K} \, \text{SINR}_{j,\text{WL-MOE}}}, \tag{4.131}$$

$$\text{SINR}_{j,\text{WL-SUB-CE}} \stackrel{\triangle}{=} \text{SINR}(\mathbf{f}_{j,\text{WL-SUB-CE}}) = \frac{\text{SINR}_{j,\text{WL-MOE}}}{1 + \frac{J-1}{K} \, \text{SINR}_{j,\text{WL-MOE}}}, \tag{4.132}$$

which are directly comparable to (4.77) and (4.80). Our simulation results show that (4.131) and (4.132) accurately predict the SINR performances of the WL-SMI-CE and WL-SUB-CE receivers not only in the high-SNR regime, but also for moderate values of the SNR, wherein many systems of practical interest are envisioned to operate. A first exam of the obtained expression shows that, for $K \to +\infty$, both receivers attain the maximum SINR equal to $\text{SINR}_{j,\text{WL-MOE}}$. A more interesting comparison is between (4.131)–(4.132) and the corresponding ones (4.77)–(4.80) derived in subsection 4.5.1 in the known-channel case.

For the WL-SUB receiver, such a comparison shows that the SINR when the channel is estimated is the same as that obtained when the channel is known, namely, for moderate-to-high values of the SNR, the WL-SUB-CE receiver (approximately) pays about no penalty w.r.t. its counterpart employing the exact channel. Such a result indirectly shows the reliability of the considered subspace-based CE procedure, which simultaneously exploits the channel information contained in both $\mathbf{R}_{\mathbf{rr}}$ and $\mathbf{R}_{\mathbf{rr}^*}$ by jointly processing the received vector $\mathbf{r}(k)$ and its conjugate version $\mathbf{r}^*(k)$.

Surprisingly enough, the SINR of the WL-SMI-CE turns out to be even better than that of the corresponding WL-SMI receiver with known channel:

as a matter of fact, this phenomenon is well-known in the array processing literature (see e.g. [78, 79, 80]), where it is sometimes referred to as *signature mismatch*, and its effects vanish only when $K \to +\infty$. Indeed, the noise-subspace estimated from the sample ACM $\widehat{\mathbf{R}}_{\mathbf{zz}}$ is not orthogonal to the true signature $\mathbf{h}_j$, i.e., $\widehat{\mathbf{U}}_n \, \mathbf{h}_j \neq \mathbf{0}_{2N-J}$; therefore, in the high-SNR regime, the noise subspace eigenvalues are significantly smaller than the signal subspace eigenvalues and, consequently, when the WL-SMI receiver (4.65) is implemented, the component of $\mathbf{h}_j$ in the estimated noise subspace is greatly amplified when evaluating $\widehat{\mathbf{R}}_{\mathbf{zz}}^{-1} \, \mathbf{h}_j$, thus introducing measurement noise at the output of the receiver, at the expense of the desired signal component. Instead, the signature mismatch effects are alleviated in the WL-SUB receiver (4.66) (and, similarly, in its estimated version) and the WL-SMI-CE receiver (4.109), since, in the former case, the known signature $\mathbf{h}_j$ is preliminarily projected on the signal subspace of the sample ACM, whereas, in the latter one, the estimated signature $\widehat{\mathbf{h}}_j$ is orthogonal to the estimated noise subspace by construction. A clear geometric interpretation of this phenomenon is reported in [79, Fig. 17]. As a by-product, eqs. (4.77), (4.80), (4.131) and (4.132) provide the SINR assessment of the signature mismatch problem, thereby showing the simplicity and insightfulness of our SINR formulas. For a finite sample-size K, indeed, accounting for (4.77) and (4.131), the SINR degradation due to signature mismatch in the high-SNR region is given by

$$\lim_{\sigma_v^2 \to 0} \frac{\text{SINR}_{j,\text{WL-SMI}}}{\text{SINR}_{j,\text{WL-SMI-CE}}} = \frac{2(N - L_j)}{2\,N - 1} < 1 \,, \tag{4.133}$$

which increases with the channel length. Another interesting conclusion that can be drawn from (4.131) and (4.132) is that, not differently from the case where the channel is known (see subsection 4.5.1), the data-estimated receivers exhibit a SINR saturation effect, for vanishingly small noise. Indeed, when $\sigma_v^2 \to 0$ and $\mathbf{H}$ is full-column rank ($J \leq 2\,N$), it has been shown in subsection 4.4.1 that $\text{SINR}_{j,\text{WL-MOE}}$ grows without bound. Thus, for $J \leq 2(N - L_j) < 2N$ (this inequality is due to (4.106)) , accounting for (4.131) and (4.132), we get

$$\lim_{\sigma_v^2 \to 0} \text{SINR}_{j,\text{WL-SMI-CE}} = \frac{K}{2(N - L_j)} \,, \tag{4.134}$$

$$\lim_{\sigma_v^2 \to 0} \text{SINR}_{j,\text{WL-SUB-CE}} = \frac{K}{J - 1} \,, \tag{4.135}$$

which show that, in the high-SNR regime, the performance of the WL-SMI-CE receiver does not depend on the number of users J, but it depends on the

processing gain N as well as on the channel length L_j of user j, whereas the performance of the WL-SUB-CE receiver is independent of both the processing gain N and the channel length L_j, while depending on the number of users J.

Performance analysis of L-MOE with channel estimation

As done in subsection 4.5.1 in the case of known channel, it is interesting to compare the SINR performances of the data-estimated WL-MOE receivers with CE based on (4.107) against the data-estimated L-MOE receivers with CE based on (4.41).

Preliminarily, it is worthwhile to note that, although a similar perturbative performance analysis was addressed in [81]–[76] for the blind L-MMSE (minimum mean-square error) receiver, the analysis carried out in this subsection for the L-MOE receiver with blind CE allows a more direct and fruitful comparison with the WL-MOE one and, moreover, leads to more easily interpretable results (although slightly less accurate) than those obtained in [81]–[76]. Moreover, we observe that the problem considered in this subsection exhibits interesting analogies with a well-studied topic in array processing, since the L-MOE-based multiuser detector is mathematically equivalent to the linear minimum variance (L-MV) beamformer [78], where in the latter the role of the received signature is played by the array steering vector (SV). Finite-sample performance analysis of the L-MV beamformer was carried out in [78, 75, 82] for the SMI version, and in [80] for the subspace-based implementation (so called projection method). Specifically, in [75] only the effects of ACM estimation were considered, whereas in [78, 80] the effects of ACM estimation and SV perturbation were *separately* studied, and a complete analysis of the *joint* effects of ACM estimation and SV perturbation was carried out only in [82]. However, the latter analysis does not explicitly account for the situation wherein the SV is blindly estimated from the received data and, consequently, the SV perturbation depends in its turn on the accuracy in ACM estimation, which is exactly the case of the subspace-based CE algorithms considered herein.

It is now in order, to turn to the L-MOE receiver with CE. We have shown that the unknown vector $\mathbf{g}_j$ can be obtained as solution of the linear system (4.40), provided that this system uniquely characterizes the channel coefficients for each users. To this end it is required that the condition $(\mathbf{c4})$ is satisfied (see subsection 4.3.2 for details). This condition, from the point of view of the jth user, necessarily imposes that the maximum number of users supported

by the system is smaller than the maximum number (see eq. (4.81)) N of users of the known-channel case

$$J_{\max,L} = (N - L_j) \leq N. \tag{4.136}$$

Observe that the maximum number of allowable users for the linear case is exactly one-half of the corresponding number for the WL case. In practice, when $J \leq N - L_j$ both blind L and WL receivers can be utilized, whereas for $N - L_j < J \leq 2(N - L_j)$ only the blind WL receivers can work (note that the above limitations are mainly due to the considered blind channel identification procedure). Moreover, in the subsection 4.3.2, we have shown that when $\mathbf{R}_{\mathbf{rr}}$ is estimated from a finite sample size, a channel estimate can be obtained as in (4.41), that we report here for completeness:

$$\begin{aligned}\widehat{\mathbf{g}}_j &= \underset{\mathbf{x}\in\mathbb{C}^{L_j}}{\operatorname{argmin}} \|\widehat{\mathbf{V}}_n^H \mathbf{C}_j \mathbf{x}\|^2 \\ &= \underset{\mathbf{x}\in\mathbb{C}^{L_j}}{\operatorname{argmin}} \left(\mathbf{x}^H \mathbf{C}_j^H \widehat{\mathbf{V}}_n \widehat{\mathbf{V}}_n^H \mathbf{C}_j \mathbf{x}\right), \quad \text{subject to } \|\mathbf{x}\|^2 = 1,\end{aligned} \tag{4.137}$$

whose solution [56] is given by the eigenvector associated with the smallest eigenvalue of the matrix associated with the smallest eigenvalue of the matrix $\widehat{\boldsymbol{\mathcal{Q}}}_{j,\mathrm{L}} \triangleq \mathbf{C}_j^H \widehat{\mathbf{V}}_n \widehat{\mathbf{V}}_n^H \mathbf{C}_j \in \mathbb{C}^{L_j \times L_j}$. By substituting (4.137) in (4.12) the resulting estimate of the received signature[16] is

$$\widehat{\boldsymbol{\phi}}_j = \alpha_j \mathbf{C}_j \widehat{\mathbf{g}}_j \tag{4.138}$$

To obtain the L-SMI and L-SUB receivers with channel estimation, it is necessary to substitute (4.138) in (4.85) and in (4.86):

$$\mathbf{w}_{j,\text{L-SMI-CE}} \triangleq (\widehat{\boldsymbol{\phi}}_j^H \widehat{\mathbf{R}}_{\mathbf{rr}}^{-1} \widehat{\boldsymbol{\phi}}_j)^{-1} \widehat{\mathbf{R}}_{\mathbf{rr}}^{-1} \widehat{\boldsymbol{\phi}}_j\,, \qquad \text{(L-SMI-CE)} \tag{4.139}$$

$$\mathbf{w}_{j,\text{L-SUB-CE}} \triangleq (\widehat{\boldsymbol{\phi}}_j^H \widehat{\mathbf{V}}_s \widehat{\boldsymbol{\Upsilon}}_s^{-1} \widehat{\mathbf{V}}_s^H \widehat{\boldsymbol{\phi}}_j)^{-1} \widehat{\mathbf{V}}_s \widehat{\boldsymbol{\Upsilon}}_s^{-1} \widehat{\mathbf{V}}_s^H \widehat{\boldsymbol{\phi}}_j\,. \qquad \text{(L-SUB-CE)} \tag{4.140}$$

As for WL receivers, while (4.16) and (4.82) are perfectly equivalent, their finite-sample counterparts given by (4.85)–(4.86) and (4.139)–(4.140) are different. The performance analysis of the L-SMI-CE and L-SUB-CE receivers is

[16] Along the same considerations done in the WL-case (see 4.5.3) we assume that both α_j and ϑ_j are known exactly (where ϑ_j is the phase ambiguity in the subspace-based channel estimation (see subsection 4.3.2))

complicated from the fact that, again, the SINR (4.111) must be evaluated but, differently from the WL ones, linear receivers do not exhibit the CS property, since the L-MOE receiver can be viewed as a WL receiver with augmented weight vector $\mathbf{f}_{j,\text{L-MOE}} \stackrel{\triangle}{=} [\mathbf{w}_{j,\text{L-MOE}}^T, \mathbf{0}_N^T]^T$. Such an analysis is similar in principle to the one carried out in [81]–[76] (as we have emphasized in the introduction to the subsection), but the approach adopted here leads to more easily interpretable results, which are directly comparable with those obtained in the WL case, at the cost of a minimal loss in accuracy. Also in the L-case, we adopt a first-order perturbative approach [76, 77] to model all estimation errors.

In the following, in order to carry out the analysis in an unified framework, we denote with $\widehat{\mathbf{w}}_j$ *any* data-estimated L-MOE receiver, i.e, $\widehat{\mathbf{w}}_j = \mathbf{w}_{j,\text{L-SMI-CE}}$ or $\widehat{\mathbf{w}}_j = \mathbf{w}_{j,\text{L-SUB-CE}}$, and set $\mathbf{w}_j = \mathbf{w}_{j,\text{L-MOE}}$ for simplicity. When a linear data-estimated receiver $\widehat{\mathbf{w}}_j$ is employed [i.e., $\mathbf{f}_{j,1} = \widehat{\mathbf{w}}_j$ and $\mathbf{f}_{j,2} = \mathbf{0}_N$ in (4.22)], eq. (4.23) assumes the form

$$\text{SINR}(\widehat{\mathbf{w}}_j) = \frac{\text{E}^2\{\text{Re}[\widehat{\mathbf{w}}_j^H \boldsymbol{\phi}_j]\}}{\text{E}\{\text{Re}^2[\widehat{\mathbf{w}}_j^H \mathbf{p}_j(k)]\} + \text{Var}\{\text{Re}[\widehat{\mathbf{w}}_j^H \boldsymbol{\phi}_j]\}} \,. \tag{4.141}$$

It is important to observe that, differently from the WL case, the real parts in (4.141) cannot be omitted, since $\widehat{\mathbf{w}}_j^H \boldsymbol{\phi}_j$ and $\widehat{\mathbf{w}}_j^H \mathbf{p}_j(k)$ are in general complex-valued random variables. This fact significantly complicates the analysis with respect to the WL case. Following the perturbative approach, the vectors $\widehat{\mathbf{w}}_j$ and $\widehat{\boldsymbol{\phi}}_j$ are expressed as $\widehat{\mathbf{w}}_j = \mathbf{w}_j + \boldsymbol{\delta}\mathbf{w}_j$ and $\widehat{\boldsymbol{\phi}}_j = \boldsymbol{\phi}_j + \boldsymbol{\delta}\boldsymbol{\phi}_j$, respectively, where $\boldsymbol{\delta}\mathbf{w}_j$ and $\boldsymbol{\delta}\boldsymbol{\phi}_j$ are *small* (i.e., $\|\boldsymbol{\delta}\mathbf{w}_j\| \ll 1$ and $\|\boldsymbol{\delta}\boldsymbol{\phi}_j\| \ll 1$), *zero-mean* perturbation terms. Let $\text{E}_{\boldsymbol{\delta}\mathbf{w}_j}[\cdot]$ be the average w.r.t. $\boldsymbol{\delta}\mathbf{w}_j$. The variance term in (4.141) is equal, for definition, to

$$\text{Var}\{\text{Re}[\widehat{\mathbf{w}}_j^H \boldsymbol{\phi}_j]\} = \text{E}_{\boldsymbol{\delta}\mathbf{w}_j}[\text{Re}^2\{\widehat{\mathbf{w}}_j^H \boldsymbol{\phi}_j\}] - \text{E}^2_{\boldsymbol{\delta}\mathbf{w}_j}[\text{Re}\{\widehat{\mathbf{w}}_j^H \boldsymbol{\phi}_j\}]. \tag{4.142}$$

Using the equality $\text{Re}^2(z) = \frac{1}{2}[|z|^2 + \text{Re}(z^2)]$, $\forall z \in \mathbb{C}$, the term $\text{E}_{\boldsymbol{\delta}\mathbf{w}_j}[\text{Re}^2\{\widehat{\mathbf{w}}_j^H \boldsymbol{\phi}_j\}]$ can be expressed as

$$2\,\text{E}_{\boldsymbol{\delta}\mathbf{w}_j}[\text{Re}^2\{\widehat{\mathbf{w}}_j^H \boldsymbol{\phi}_j\}] = \text{E}_{\boldsymbol{\delta}\mathbf{w}_j}[|\widehat{\mathbf{w}}_j^H \boldsymbol{\phi}_j|^2] + \text{E}_{\boldsymbol{\delta}\mathbf{w}_j}[\text{Re}\{(\widehat{\mathbf{w}}_j^H \boldsymbol{\phi}_j)^2\}] \tag{4.143}$$

with $\text{E}_{\boldsymbol{\delta}\mathbf{w}_j}[|\widehat{\mathbf{w}}_j^H \boldsymbol{\phi}_j|^2] = 1 + \text{E}_{\boldsymbol{\delta}\mathbf{w}_j}[\boldsymbol{\delta}\mathbf{w}_j^H \boldsymbol{\phi}_j\, \boldsymbol{\phi}_j^H\, \boldsymbol{\delta}\mathbf{w}_j]$ and $\text{E}_{\boldsymbol{\delta}\mathbf{w}_j}[\text{Re}\{(\widehat{\mathbf{w}}_j^H \boldsymbol{\phi}_j)^2\}] = \text{Re}\{\text{E}_{\boldsymbol{\delta}\mathbf{w}_j}[(\widehat{\mathbf{w}}_j^H \boldsymbol{\phi}_j)^2]\} = 1 + \text{Re}\{\text{E}_{\boldsymbol{\delta}\mathbf{w}_j}[\boldsymbol{\delta}\mathbf{w}_j^H \boldsymbol{\phi}_j\, \boldsymbol{\phi}_j^T\, \boldsymbol{\delta}\mathbf{w}_j^*]\}$, since $\mathbf{w}_j^H \boldsymbol{\phi}_j = 1$ by (4.16) and $\boldsymbol{\delta}\mathbf{w}_j$ is zero-mean by assumption. Moreover, $\text{E}_{\boldsymbol{\delta}\mathbf{w}_j}[\text{Re}\{\widehat{\mathbf{w}}_j^H \boldsymbol{\phi}_j\}] = \text{Re}\{\text{E}_{\boldsymbol{\delta}\mathbf{w}_j}[\widehat{\mathbf{w}}_j^H \boldsymbol{\phi}_j]\} = 1$

(see subsection 4.5.1). Consequently, by substituting these relations in (4.142), it follows that

$$\mathrm{Var}\{\mathrm{Re}[\widehat{\mathbf{w}}_j^H \boldsymbol{\phi}_j]\} = \frac{1}{2}\,\mathrm{E}_{\boldsymbol{\delta}\mathbf{w}_j}[\boldsymbol{\delta}\mathbf{w}_j^H \boldsymbol{\phi}_j\, \boldsymbol{\phi}_j^H\, \boldsymbol{\delta}\mathbf{w}_j] + \frac{1}{2}\,\mathrm{Re}\{\mathrm{E}_{\boldsymbol{\delta}\mathbf{w}_j}[\boldsymbol{\delta}\mathbf{w}_j^H \boldsymbol{\phi}_j\, \boldsymbol{\phi}_j^T\, \boldsymbol{\delta}\mathbf{w}_j^*]\} \tag{4.144}$$

Similarly to the WL case, we assume that the weight vector $\widehat{\mathbf{w}}_j$ is independent from the data vector $\mathbf{p}_j(k)$. Let $\mathrm{E}_{\widehat{\mathbf{w}}_j \mathbf{p}_j}[\cdot]$ denote the joint average w.r.t. $\widehat{\mathbf{w}}_j$ and $\mathbf{p}_j(k)$. Using again the identity $\mathrm{Re}^2(z) = \frac{1}{2}[|z|^2 + \mathrm{Re}(z^2)]$, $\forall z \in \mathbb{C}$, performing the average w.r.t to $\mathbf{p}_j(k)$, and recalling that, due to assumptions **(a1)** and **(a2)**, the vector $\mathbf{p}_j(k)$ is zero-mean, the term $\mathrm{E}_{\widehat{\mathbf{w}}_j,\mathbf{p}_j}\{\mathrm{Re}^2[\widehat{\mathbf{w}}_j^H \mathbf{p}_j(k)]\}$ in (4.141) can be expressed as

$$\begin{aligned}\mathrm{E}_{\widehat{\mathbf{w}}_j,\mathbf{p}_j}\{\mathrm{Re}^2[\widehat{\mathbf{w}}_j^H \mathbf{p}_j(k)]\} &= \frac{\mathrm{E}_{\widehat{\mathbf{w}}_j,\mathbf{p}_j}\{|\widehat{\mathbf{w}}_j^H \mathbf{p}_j(k)|^2\} + \mathrm{E}_{\widehat{\mathbf{w}}_j,\mathbf{p}_j}\{\mathrm{Re}[(\widehat{\mathbf{w}}_j^H \mathbf{p}_j(k))^2]\}}{2} \\ &= \frac{\mathbf{w}_j^H \mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}\, \mathbf{w}_j + \mathrm{Re}(\mathbf{w}_j^H \mathbf{R}_{\mathbf{p}_j\mathbf{p}_j^*}\, \mathbf{w}_j^*)}{2} \\ &+ \frac{\mathrm{E}_{\boldsymbol{\delta}\mathbf{w}_j}\{\boldsymbol{\delta}\mathbf{w}_j^H \mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}\, \boldsymbol{\delta}\mathbf{w}_j\} + \mathrm{Re}\{\mathrm{E}_{\boldsymbol{\delta}\mathbf{w}_j}[\boldsymbol{\delta}\mathbf{w}_j^H \mathbf{R}_{\mathbf{p}_j\mathbf{p}_j^*}\, \boldsymbol{\delta}\mathbf{w}_j^*]\}}{2}\end{aligned} \tag{4.145}$$

Noticing that $\mathbf{R}_{\mathbf{rr}} = \boldsymbol{\phi}_j\, \boldsymbol{\phi}_j^H + \mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}$ and $\mathbf{R}_{\mathbf{rr}^*} = \boldsymbol{\phi}_j\, \boldsymbol{\phi}_j^T + \mathbf{R}_{\mathbf{p}_j\mathbf{p}_j^*}$, recalling the equation (4.50), and substituting (4.144) and (4.145) in (4.141), we get

$$\mathrm{SINR}(\widehat{\mathbf{w}}_j) = \frac{\mathrm{SINR}_{j,\text{L-MOE}}}{1 + \frac{\mathrm{SINR}_{j,\text{L-MOE}}\{\mathrm{trace}(\mathbf{R}_{\mathbf{rr}}\, \mathbf{R}_{\boldsymbol{\delta}\mathbf{w}_j\boldsymbol{\delta}\mathbf{w}_j}) + \mathrm{Re}[\mathrm{trace}(\mathbf{R}_{\mathbf{rr}^*}\, \mathbf{R}^*_{\boldsymbol{\delta}\mathbf{w}_j\boldsymbol{\delta}\mathbf{w}_j^*})]\}}{2}}, \tag{4.146}$$

where only the average w.r.t. $\boldsymbol{\delta}\mathbf{w}_j$ is left to evaluated. To proceed further, explicit expressions for the perturbation $\boldsymbol{\delta}\mathbf{w}_j$ are needed for both the L-SMI-CE and L-SUB-CE receivers. The following Lemma gives a first-order characterization of the perturbation vector $\boldsymbol{\delta}\mathbf{w}_j$:

Lemma 4.7 *Let $\approx$ denote first-order equality, the first-order perturbation term of the L-SMI-CE and L-SUB-CE receivers can be expressed as*

$$\boldsymbol{\delta}\mathbf{w}_j \approx \underbrace{-\boldsymbol{\Gamma}_{j,L}\, \widehat{\mathbf{r}}_{\mathbf{p}_j b_j}}_{\boldsymbol{\delta}\mathbf{w}_j^{(1)}} + \underbrace{\boldsymbol{\Delta}_{j,L}^{(1)}\, \boldsymbol{\delta}\boldsymbol{\phi}_j + \boldsymbol{\Delta}_{j,L}^{(2)}\, \boldsymbol{\delta}\boldsymbol{\phi}_j^*}_{\boldsymbol{\delta}\mathbf{w}_j^{(2)}} = \boldsymbol{\delta}\mathbf{w}_j^{(1)} + \boldsymbol{\delta}\mathbf{w}_j^{(2)} \tag{4.147}$$

where the random vector $\widehat{\mathbf{r}}_{\mathbf{p}_j b_j} \overset{\triangle}{=} \frac{1}{K}\sum_{k=1}^{K} \mathbf{p}_j(k)\, b_j(k) \in \mathbb{C}^N$ is the sample estimate of the cross-correlation between the disturbance vector $\mathbf{p}_j(k)$ and the

desired symbol $b_j(k)$*, whereas*

$$\mathbf{\Gamma}_{j,L} \triangleq \begin{cases} \mathbf{P}_{j,\mathrm{L}}\,\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}^{-1}\,, & \text{(L-SMI-CE)} \\ \mathbf{P}_{j,\mathrm{L}}\,\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}^{-1} - \gamma_{j,\mathrm{L}}\,\mathbf{V}_n\mathbf{V}_n^H\,, & \text{(L-SUB-CE)} \end{cases} \tag{4.148}$$

$$\mathbf{\Delta}_{j,L}^{(1)} \triangleq \begin{cases} (\boldsymbol{\phi}_j^H\mathbf{R}_{\mathbf{rr}}^{-1}\,\boldsymbol{\phi}_j)^{-1}\,\mathbf{R}_{\mathbf{rr}}^{-1}\,\mathbf{P}_{j,\mathrm{L}}^H\,, & \text{(L-SMI-CE)} \\ (\boldsymbol{\phi}_j^H\mathbf{R}_{\mathbf{rr}}^{-1}\,\boldsymbol{\phi}_j)^{-1}\,\mathbf{V}_s\,\mathbf{\Upsilon}_s^{-1}\,\mathbf{V}_s^H\,\mathbf{P}_{j,\mathrm{L}}^H\,, & \text{(L-SUB-CE)} \end{cases} \tag{4.149}$$

and $\mathbf{\Delta}_{j,L}^{(2)} \triangleq -\mathbf{w}_j\,\mathbf{w}_j^T$*, with* $\mathbf{P}_{j,\mathrm{L}} \triangleq \mathbf{I}_N - (\boldsymbol{\phi}_j^H\,\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}^{-1}\,\boldsymbol{\phi}_j)^{-1}\,\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}^{-1}\,\boldsymbol{\phi}_j\,\boldsymbol{\phi}_j^H = \mathbf{I}_N - \mathbf{w}_j\,\boldsymbol{\phi}_j^H \in \mathbb{C}^{N\times N}$ *being an oblique projection matrix [75] and* $\gamma_{j,\mathrm{L}} \triangleq \sigma_v^{-2} + (\boldsymbol{\phi}_j^H\,\mathbf{R}_{\mathbf{rr}}^{-1}\,\boldsymbol{\phi}_j)^{-1}\boldsymbol{\phi}_j^H\mathbf{V}_s\,\mathbf{\Omega}_L^{-1}\,\mathbf{V}_s^H\mathbf{R}_{\mathbf{rr}}^{-1}\,\boldsymbol{\phi}_j > 0$*, while the diagonal matrix* $\mathbf{\Omega}_L \triangleq diag(\mu_1, \mu_2, \ldots, \mu_J) \in \mathbb{R}^{J\times J}$ *collects the nonzero eigenvalues of* $\mathbf{\Phi}\,\mathbf{\Phi}^H$*.*

Proof: The proof can be conducted along the same lines of Appendices B.5 with the additional complication that, contrary to $\mathbf{f}_j^H$ and $\boldsymbol{\delta}\mathbf{h}_j$, $\mathbf{w}_j^H$ and $\widehat{\boldsymbol{\phi}}_j$ do not exhibit the CS property.

It should be noted that the quantities $\mathbf{P}_{j,\mathrm{L}}$, $\gamma_{j,\mathrm{L}}$ and $\mathbf{\Omega}_{\mathrm{L}}$ are already defined in Lemma 4.3, nevertheless they are here report for completeness. Moreover we emphasize that $\boldsymbol{\delta}\mathbf{w}_j^{(1)}$ and $\boldsymbol{\delta}\mathbf{w}_j^{(2)}$ represent the perturbations due to estimation of $\mathbf{R}_{\mathbf{rr}}$ and $\boldsymbol{\phi}_j$, respectively; indeed, a comparison shows that the expression of $\boldsymbol{\delta}\mathbf{w}_j^{(1)}$ is the same as that reported in Lemma 4.3. In order to characterize the perturbation term $\boldsymbol{\delta}\mathbf{w}_j^{(2)}$, it is necessary to evaluate the perturbation $\boldsymbol{\delta}\boldsymbol{\phi}_j$ associated to the subspace-based CE procedure given by (4.137).

Lemma 4.8 *Given the estimate* $\widehat{\boldsymbol{\phi}}_j = \boldsymbol{\phi}_j + \boldsymbol{\delta}\boldsymbol{\phi}_j = \alpha_j\,\mathbf{C}_j\,\widehat{\mathbf{g}}_j$ *of the signature* $\boldsymbol{\phi}_j$*, where the channel estimate* $\widehat{\mathbf{g}}_j$ *is the solution of* (4.137)*, the perturbation* $\boldsymbol{\delta}\boldsymbol{\phi}_j$ *can be expressed as*

$$\boldsymbol{\delta}\boldsymbol{\phi}_j \approx \mathbf{\Pi}_{j,L}\,\widehat{\mathbf{r}}_{\mathbf{p}_j b_j}, \tag{4.150}$$

where $\mathbf{\Pi}_{j,L} \triangleq (\boldsymbol{\phi}_j^H\,\mathbf{V}_s\,\mathbf{\Omega}_L^{-1}\,\mathbf{V}_s^H\,\boldsymbol{\phi}_j)\,\mathbf{C}_j\,\boldsymbol{\mathcal{Q}}_{j,L}^{\dagger}\,\mathbf{C}_j^H\,\mathbf{V}_n\,\mathbf{V}_n^H \in \mathbb{C}^{N\times N}\,\mathbf{U}_s^H\,\mathbf{h}_j)\,\boldsymbol{\mathcal{C}}_j\,\boldsymbol{\mathcal{Q}}_{j,WL}^{\dagger}\,\boldsymbol{\mathcal{C}}_j^H\,\mathbf{U}_n\,\mathbf{U}_n^H \in \mathbb{C}^{2N\times 2N}$*, with* $\mathbf{\Omega}_L$ *and* $\widehat{\mathbf{r}}_{\mathbf{p}_j b_j}$ *defined in Lemma 4.7, and* $\boldsymbol{\mathcal{Q}}_{j,\mathrm{L}} \triangleq \mathbf{C}_j^H\,\mathbf{V}_n\,\mathbf{V}_n^H\,\mathbf{C}_j \in \mathbb{C}^{L_j\times L_j}$*.*

Proof: The Proof can be conducted along the same lines of Appendix B.6.

Accounting for (4.150) and Lemma 4.7, the overall perturbation of the L-SMI-CE and L-SUB-CE weight vectors can be expressed similarly to the WL case, as a *linear* function of $\widehat{\mathbf{r}}_{\mathbf{q}_j b_j} = [\widehat{\mathbf{r}}^T_{\mathbf{p}_j b_j}, \widehat{\mathbf{r}}^H_{\mathbf{p}_j b_j}]^T$, as summarized by the following Lemma:

Lemma 4.9 *The first-order overall perturbation term $\boldsymbol{\delta}\mathbf{w}_j = \boldsymbol{\delta}\mathbf{w}^{(1)}_j + \boldsymbol{\delta}\mathbf{w}^{(2)}_j$ of the L-SMI-CE and L-SUB-CE receivers can be expressed in a unified manner as*

$$\boldsymbol{\delta}\mathbf{w}_j \approx \boldsymbol{\Sigma}_{j,L}\, \widehat{\mathbf{r}}_{\mathbf{q}_j b_j}, \tag{4.151}$$

where $\boldsymbol{\Sigma}_{j,L} \triangleq [\boldsymbol{\Sigma}^{(1)}_{j,L}, \boldsymbol{\Sigma}^{(2)}_{j,L}]$, with $\boldsymbol{\Sigma}^{(1)}_{j,L} \triangleq -\boldsymbol{\Gamma}_{j,L} + \boldsymbol{\Delta}^{(1)}_{j,L}\, \boldsymbol{\Pi}_{j,L} \in \mathbb{C}^{N\times N}$ and $\boldsymbol{\Sigma}^{(2)}_{j,L} \triangleq \boldsymbol{\Delta}^{(2)}_{j,L}\, \boldsymbol{\Pi}^*_{j,L} \in \mathbb{C}^{N\times N}$. $\boldsymbol{\Pi}_{j,L} \in \mathbb{C}^{N\times N}$ and $\boldsymbol{\Delta}_{j,L} \in \mathbb{C}^{N\times N}$ are given by Lemma 4.7 and $\boldsymbol{\Pi}_{j,L} \in \mathbb{C}^{N\times N}$ has been defined in Lemma 4.8.

It should be observed that Lemma 4.9 provides a compact characterization of the overall perturbation $\boldsymbol{\delta}\mathbf{w}_j$, which is obtained under the simplifying assumption [75] that the error in estimating $\mathbf{R}_{\mathbf{rr}}$ is mainly due to the term $\widehat{\mathbf{r}}_{\mathbf{p}_j b_j}$. In particular by substituting eq. (4.151) in (4.146), the SINR assumes the form

$$\mathrm{SINR}(\widehat{\mathbf{w}}_j) = \frac{\mathrm{SINR}_{j,\text{L-MOE}}}{1+\frac{\mathrm{SINR}_{j,\text{L-MOE}}\{\mathrm{trace}(\mathbf{R}_{\mathbf{rr}}\, \boldsymbol{\Sigma}_{j,\mathrm{L}}\, \mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}\, \boldsymbol{\Sigma}^H_{j,\mathrm{L}}) + \mathrm{Re}[\mathrm{trace}(\mathbf{R}_{\mathbf{rr}^*}\, \boldsymbol{\Sigma}^*_{j,\mathrm{L}}\, \mathbf{J}\, \mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}\, \boldsymbol{\Sigma}^H_{j,\mathrm{L}})]\}}{2K}}, \tag{4.152}$$

where, by virtue of assumptions **(a1)** and **(a2)**, we have used (see 4.5.1 for details) the fact that $\mathrm{E}[\widehat{\mathbf{r}}_{\mathbf{q}_j b_j}\, \widehat{\mathbf{r}}^H_{\mathbf{q}_j b_j}] = \frac{1}{K}\, \mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}$ and $\mathrm{E}[\widehat{\mathbf{r}}_{\mathbf{q}_j b_j}\, \widehat{\mathbf{r}}^T_{\mathbf{q}_j b_j}] = \frac{1}{K}\, \mathbf{J}\, \mathbf{R}^*_{\mathbf{q}_j\mathbf{q}_j}$, with $\mathbf{J} \triangleq \big[[\mathbf{O}_{N\times N}, \mathbf{I}_N]^T, [\mathbf{I}_N, \mathbf{O}_{N\times N}]^T\big]^T$. The matrix $\mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}$ has a particular block structure where the lower-right block $\mathbf{R}^*_{\mathbf{p}_j\mathbf{p}_j}$ is the conjugate of the upper-left one $\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}$, and the lower-left block $\mathbf{R}^*_{\mathbf{p}_j\mathbf{p}^*_j}$ is the conjugate of the upper-right one $\mathbf{R}_{\mathbf{p}_j\mathbf{p}^*_j}$. Moreover, since $\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j} = \overline{\boldsymbol{\Phi}}_j\, \overline{\boldsymbol{\Phi}}^H_j + \sigma^2_v\, \mathbf{I}_N$, $\mathbf{R}_{\mathbf{p}_j\mathbf{p}^*_j} = \overline{\boldsymbol{\Phi}}_j\, \overline{\boldsymbol{\Phi}}^T_j$ and $\mathbf{V}^H_n\, \overline{\boldsymbol{\Phi}}_j = \mathbf{O}_{(N-J)\times(J-1)}$, one has $\boldsymbol{\Pi}^*_{j,\mathrm{L}}\, \mathbf{R}^*_{\mathbf{p}_j\mathbf{p}^*_j} = \mathbf{O}_{N\times N}$ and $\boldsymbol{\Pi}^*_{j,\mathrm{L}}\, \mathbf{R}^*_{\mathbf{p}_j\mathbf{p}_j} = \sigma^2_v\, \boldsymbol{\Pi}^*_{j,\mathrm{L}}$ and, hence, $\boldsymbol{\Sigma}^{(2)}_{j,\mathrm{L}}\, \mathbf{R}^*_{\mathbf{p}_j\mathbf{p}^*_j} = \mathbf{R}_{\mathbf{p}_j\mathbf{p}^*_j}\, (\boldsymbol{\Sigma}^{(2)}_{j,\mathrm{L}})^H = \mathbf{O}_{N\times N}$ and $\boldsymbol{\Sigma}^{(2)}_{j,\mathrm{L}}\, \mathbf{R}^*_{\mathbf{p}_j\mathbf{p}_j} = \sigma^2_v\, \boldsymbol{\Sigma}^{(2)}_{j,\mathrm{L}}$. By exploiting the block structure of $\mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}$ and $\boldsymbol{\Sigma}_{j,\mathrm{L}}$, it follows that

$$\begin{aligned}&\mathrm{trace}(\mathbf{R}_{\mathbf{rr}}\, \boldsymbol{\Sigma}_{j,\mathrm{L}}\, \mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}\, \boldsymbol{\Sigma}^H_{j,\mathrm{L}})\\ &= \mathrm{trace}[(\boldsymbol{\Sigma}^{(1)}_{j,\mathrm{L}})^H \mathbf{R}_{\mathbf{rr}}\, \boldsymbol{\Sigma}^{(1)}_{j,\mathrm{L}}\, \mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}] + \sigma^2_v\, \mathrm{trace}[(\boldsymbol{\Sigma}^{(2)}_{j,\mathrm{L}})^H \mathbf{R}_{\mathbf{rr}}\, \boldsymbol{\Sigma}^{(2)}_{j,\mathrm{L}}].\end{aligned} \tag{4.153}$$

Using the expressions of $\mathbf{\Delta}_{j,\mathrm{L}}^{(2)}$ and $\mathbf{\Pi}_{j,\mathrm{L}}$, after simple manipulations, we get $\mathrm{trace}[(\mathbf{\Sigma}_{j,\mathrm{L}}^{(2)})^H \mathbf{R}_{\mathbf{rr}}\, \mathbf{\Sigma}_{j,\mathrm{L}}^{(2)}] = \zeta_{j,\mathrm{L}}^2\, \mathrm{trace}[\mathbf{R}_{\mathbf{rr}}^{-1}\, (\boldsymbol{\phi}_j\, \mathbf{w}_j^H)\, (\mathbf{C}_j\, \boldsymbol{\mathcal{Q}}_{j,\mathrm{L}}^{\dagger}\, \mathbf{C}_j^H)]$, where $\zeta_{j,\mathrm{L}} \stackrel{\triangle}{=} (\boldsymbol{\phi}_j^H\, \mathbf{R}_{\mathbf{rr}}^{-1}\, \boldsymbol{\phi}_j)^{-1}\, \boldsymbol{\phi}_j^H\, \mathbf{V}_s\, \mathbf{\Omega}_{\mathrm{L}}^{-1}\, \mathbf{V}_s^H\, \boldsymbol{\phi}_j > 0$. Therefore, proceeding similarly to the WL case (see Appendix B.7) and accounting for (4.95), it can be verified that, with reference to the L-SMI-CE receiver, the first trace term in (4.152) is given by

$$\begin{aligned}&\mathrm{trace}(\mathbf{R}_{\mathbf{rr}}\, \mathbf{\Sigma}_{j,\mathrm{L}}\, \mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}\, \mathbf{\Sigma}_{j,\mathrm{L}}^H)\\ &= (N-1) - 2\,\zeta_{j,\mathrm{L}}(L_j-1) + \zeta_{j,\mathrm{L}}^2\, \sigma_v^2\, \mathrm{trace}(\mathbf{R}_{\mathbf{rr}}^{-1}\, \mathbf{C}_j\, \boldsymbol{\mathcal{Q}}_{j,\mathrm{L}}^{\dagger}\, \mathbf{C}_j^H)\,, \end{aligned} \tag{4.154}$$

whereas, for the L-SUB-CE receiver, one obtains

$$\begin{aligned}\mathrm{trace}(\mathbf{R}_{\mathbf{rr}}\, \mathbf{\Sigma}_{j,\mathrm{L}}\, \mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}\, \mathbf{\Sigma}_{j,\mathrm{L}}^H) &= (J-1) + (N-J)|1-\gamma_{j,\mathrm{L}}\, \sigma_v^2|^2\\ &\quad - \zeta_{j,\mathrm{L}}^2(L_j-1) + \zeta_{j,\mathrm{L}}^2\, \sigma_v^2\, \mathrm{trace}(\mathbf{R}_{\mathbf{rr}}^{-1}\, \mathbf{C}_j\, \boldsymbol{\mathcal{Q}}_{j,\mathrm{L}}^{\dagger}\, \mathbf{C}_j^H)\,. \end{aligned} \tag{4.155}$$

As regards the other trace term in (4.152), proceeding as done for the first one, it can be shown that

$$\begin{aligned}\mathrm{trace}(\mathbf{R}_{\mathbf{rr}^*}\, \mathbf{\Sigma}_{j,\mathrm{L}}^*\, \mathbf{J}\, \mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}\, \mathbf{\Sigma}_{j,\mathrm{L}}^H) &= \mathrm{trace}[(\mathbf{\Sigma}_{j,\mathrm{L}}^{(1)})^H\, \mathbf{R}_{\mathbf{rr}^*}\, (\mathbf{\Sigma}_{j,\mathrm{L}}^{(1)})^*\, \mathbf{R}_{\mathbf{p}_j\mathbf{p}_j^*}^*]\\ &\quad + 2\,\mathrm{trace}[(\mathbf{\Sigma}_{j,\mathrm{L}}^{(2)})^H \mathbf{R}_{\mathbf{rr}^*}\, (\mathbf{\Sigma}_{j,\mathrm{L}}^{(1)})^*\, \mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}^*]. \end{aligned} \tag{4.156}$$

Since, in addition to $\mathbf{\Pi}_{j,\mathrm{L}}^*\, \mathbf{R}_{\mathbf{p}_j\mathbf{p}_j^*}^* = \mathbf{O}_{N\times N}$, the fact that $\mathbf{V}_n^H\, \overline{\mathbf{\Phi}}_j = \mathbf{O}_{(N-J)\times(J-1)}$ also implies that $(\mathbf{\Sigma}_{j,\mathrm{L}}^{(1)})^*\, \mathbf{R}_{\mathbf{p}_j\mathbf{p}_j^*}^* = -\mathbf{\Gamma}_{j,\mathrm{L}}\, \mathbf{R}_{\mathbf{p}_j\mathbf{p}_j^*}^*$, by resorting to the properties of the trace operator, one has

$$\begin{aligned}&\mathrm{trace}[(\mathbf{\Sigma}_{j,\mathrm{L}}^{(1)})^H\, \mathbf{R}_{\mathbf{rr}^*}\, (\mathbf{\Sigma}_{j,\mathrm{L}}^{(1)})^*\, \mathbf{R}_{\mathbf{p}_j\mathbf{p}_j^*}^*]\\ &= \mathrm{trace}(\mathbf{\Gamma}_{j,\mathrm{L}}^H\, \mathbf{R}_{\mathbf{rr}^*}\, \mathbf{\Gamma}_{j,\mathrm{L}}^*\, \mathbf{R}_{\mathbf{p}_j\mathbf{p}_j^*}^*) - \mathrm{trace}(\mathbf{\Pi}_{j,\mathrm{L}}^H\, \mathbf{\Delta}_{j,\mathrm{L}}^{(1)}\, \mathbf{R}_{\mathbf{rr}^*}\, \mathbf{\Gamma}_{j,\mathrm{L}}^*\, \mathbf{R}_{\mathbf{p}_j\mathbf{p}_j^*}^*)\\ &= \mathrm{trace}(\mathbf{\Gamma}_{j,\mathrm{L}}^H\, \mathbf{R}_{\mathbf{p}_j\mathbf{p}_j^*}\, \mathbf{\Gamma}_{j,\mathrm{L}}^*\, \mathbf{R}_{\mathbf{p}_j\mathbf{p}_j^*}^*) - \mathrm{trace}(\mathbf{\Delta}_{j,\mathrm{L}}^{(1)}\, \mathbf{R}_{\mathbf{rr}^*}\, \mathbf{\Gamma}_{j,\mathrm{L}}^*\, \mathbf{R}_{\mathbf{p}_j\mathbf{p}_j^*}^*\, \mathbf{\Pi}_{j,\mathrm{L}}^H)\\ &= \mathrm{trace}(\mathbf{\Gamma}_{j,\mathrm{L}}^H\, \mathbf{R}_{\mathbf{p}_j\mathbf{p}_j^*}\, \mathbf{\Gamma}_{j,\mathrm{L}}^*\, \mathbf{R}_{\mathbf{p}_j\mathbf{p}_j^*}^*)\,, \end{aligned} \tag{4.157}$$

where it is verified that $\mathbf{\Gamma}_{j,\mathrm{L}}^H\, \mathbf{R}_{\mathbf{rr}^*} = \mathbf{\Gamma}_{j,\mathrm{L}}^H\, \mathbf{R}_{\mathbf{p}_j\mathbf{p}_j^*}$ and $\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j^*}^*\, \mathbf{\Pi}_{j,\mathrm{L}}^H = \mathbf{O}_{N\times N}$. Moreover, observing that $\mathbf{R}_{\mathbf{rr}^*} = \mathbf{\Phi}_j\, \mathbf{\Phi}_j^T$ is symmetric, substituting in the second summand of (4.156), the expression of $\mathbf{\Sigma}_{j,\mathrm{L}}^{(1)}$, $\mathbf{\Sigma}_{j,\mathrm{L}}^{(2)}$ and $\mathbf{w}_j$ [see (4.16)],

one obtains

$$\mathrm{Re}\{\mathrm{trace}[(\mathbf{\Sigma}_{j,\mathrm{L}}^{(2)})^H \mathbf{R}_{\mathbf{r}\mathbf{r}^*} (\mathbf{\Sigma}_{j,\mathrm{L}}^{(1)})^* \mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}^*]\} = -\frac{\sigma_v^2 \zeta_{j,\mathrm{L}}^2}{\boldsymbol{\phi}_j^H \mathbf{R}_{\mathbf{r}\mathbf{r}}^{-1} \boldsymbol{\phi}_j} \mathrm{Re}[\boldsymbol{\phi}_j^H \mathbf{R}_{\mathbf{r}\mathbf{r}}^{-1} \mathbf{C}_j \boldsymbol{\mathcal{Q}}_{j,\mathrm{L}}^{\dagger} \mathbf{C}_j^H \mathbf{P}_{j,\mathrm{L}} \mathbf{R}_{\mathbf{r}\mathbf{r}}^{-1} \mathbf{R}_{\mathbf{r}\mathbf{r}^*} (\mathbf{R}_{\mathbf{r}\mathbf{r}}^{-1})^* \boldsymbol{\phi}_j^*], \quad (4.158)$$

where we have also observed that, with reference to both L-SMI-CE and L-SUB-CE receivers, $\mathbf{\Pi}_{j,\mathrm{L}} \mathbf{R}_{\mathbf{p}_j\mathbf{p}_j} \mathbf{\Gamma}_{j,\mathrm{L}}^H \mathbf{R}_{\mathbf{r}\mathbf{r}^*} = \mathbf{\Pi}_{j,\mathrm{L}} \mathbf{R}_{\mathbf{p}_j\mathbf{p}_j} \mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}^{-1} \mathbf{P}_{j,\mathrm{L}}^H \mathbf{R}_{\mathbf{r}\mathbf{r}^*} = \mathbf{\Pi}_{j,\mathrm{L}} \mathbf{P}_{j,\mathrm{L}}^H \mathbf{R}_{\mathbf{r}\mathbf{r}^*} = \mathbf{\Pi}_{j,\mathrm{L}} \mathbf{R}_{\mathbf{r}\mathbf{r}^*} = \mathbf{O}_{N\times N}$, since $\mathbf{\Pi}_{j,\mathrm{L}} \mathbf{P}_{j,\mathrm{L}}^H = \mathbf{\Pi}_{j,\mathrm{L}}$ and moreover, that $\mathbf{\Pi}_{j,\mathrm{L}} \mathbf{R}_{\mathbf{p}_j\mathbf{p}_j} \mathbf{\Pi}_{j,\mathrm{L}}^H = \sigma_v^2 (\boldsymbol{\phi}_j^H \mathbf{V}_s \mathbf{\Omega}_{\mathrm{L}}^{-1} \mathbf{V}_s^H \boldsymbol{\phi}_j)^2 \mathbf{C}_j \boldsymbol{\mathcal{Q}}_{j,\mathrm{L}}^{\dagger} \mathbf{C}_j^H$. The evaluation of the trace terms at the last hand of (4.157) and (4.158) are complicated and, to obtain manageable expressions, it is convenient to consider their asymptotic values as $\sigma_v^2 \to 0$. Using the limit formula for the generalized inverse [24], one gets

$$\lim_{\sigma_v^2\to 0} \boldsymbol{\phi}_j^H \mathbf{R}_{\mathbf{r}\mathbf{r}}^{-1} \mathbf{C}_j \boldsymbol{\mathcal{Q}}_{j,\mathrm{L}}^{+} \mathbf{C}_j^H \mathbf{P}_{j,\mathrm{L}} \mathbf{R}_{\mathbf{r}\mathbf{r}}^{-1} \mathbf{R}_{\mathbf{r}\mathbf{r}^*} (\mathbf{R}_{\mathbf{r}\mathbf{r}}^{-1})^* \boldsymbol{\phi}_j^* = \mathbf{1}_j^T \mathbf{\Phi}^{\dagger} \mathbf{C}_j \boldsymbol{\mathcal{Q}}_{j,\mathrm{L}}^{+} \mathbf{C}_j^H (\mathbf{\Phi}^H)^{\dagger} (\mathbf{I}_J - \mathbf{1}_j \mathbf{1}_j^T) \mathbf{1}_j = 0\,, \quad (4.159)$$

with $\mathbf{1}_j \triangleq [\overbrace{0,\ldots,0}^{j-1},1,0,\ldots,0]^T \in \mathbb{R}^{J\times 1}$, where we have observed that $\mathbf{\Phi}^{\dagger}\mathbf{\Phi} = \mathbf{I}_J$. Similarly, it can be verified that $\lim_{\sigma_v^2\to 0} (\boldsymbol{\phi}_j^H \mathbf{R}_{\mathbf{r}\mathbf{r}}^{-1} \boldsymbol{\phi}_j)^{-1} = 1$ and $\lim_{\sigma_v^2\to 0} \zeta_{j,\mathrm{L}} = 1$. Henceforth, noticing that the trace term at the last hand of (4.157) has been evaluated in subsection 4.5.1 (see eq. (4.98)), accounting for (4.159), it can be shown that, with reference to both L-SMI-CE and L-SUB-CE receivers, the real part of the trace term in (4.156) is given by

$$\lim_{\sigma_v^2\to 0} \mathrm{Re}[\mathrm{trace}(\mathbf{R}_{\mathbf{r}\mathbf{r}^*} \mathbf{\Sigma}_{j,\mathrm{L}}^* \mathbf{J} \mathbf{R}_{\mathbf{q}_j\mathbf{q}_j} \mathbf{\Sigma}_{j,\mathrm{L}}^H)] = \lim_{\sigma_v^2\to 0} \mathrm{Re}[\mathrm{trace}(\mathbf{\Gamma}_{j,\mathrm{L}}^H \mathbf{R}_{\mathbf{p}_j\mathbf{p}_j^*} \mathbf{\Gamma}_{j,\mathrm{L}}^* \mathbf{R}_{\mathbf{p}_j\mathbf{p}_j^*}^*)] = J - 1\,. \quad (4.160)$$

Finally, for $\sigma_v^2 \to 0$, as in the WL case (see subsection 4.5.3), one has $\sigma_v^2 \,\mathrm{trace}(\mathbf{R}_{\mathbf{r}\mathbf{r}}^{-1} \mathbf{C}_j \boldsymbol{\mathcal{Q}}_{j,\mathrm{L}}^{\dagger} \mathbf{C}_j^H) \to L_j - 1$, $\gamma_{j,\mathrm{L}} \sigma_v^2 \to 1$ and $\zeta_{j,\mathrm{L}} \to 1$. Thus, it follows from (4.154) and (4.155) that

$$\lim_{\sigma_v^2\to 0} \mathrm{trace}(\mathbf{R}_{\mathbf{r}\mathbf{r}} \mathbf{\Sigma}_{j,\mathrm{L}} \mathbf{R}_{\mathbf{p}_j\mathbf{p}_j} \mathbf{\Sigma}_{j,\mathrm{L}}^H) = \begin{cases} N - L_j\,, & \text{(L-SMI-CE)} \\ J - 1\,. & \text{(L-SUB-CE)} \end{cases} \quad (4.161)$$

By substituting (4.160) and (4.161) into (4.152), we obtain that, in the high-SNR regime, the output SINR of the L-SMI-CE and L-SUB-CE receivers can

be approximately written as

$$\text{SINR}_{j,\text{L-SMI-CE}} \triangleq \text{SINR}(\mathbf{w}_{j,\text{L-SMI-CE}}) = \frac{\text{SINR}_{j,\text{L-MOE}}}{1+\frac{N+J-L_j-1}{2K}\text{SINR}_{j,\text{L-MOE}}}, \tag{4.162}$$

$$\text{SINR}_{j,\text{L-SUB-CE}} \triangleq \text{SINR}(\mathbf{w}_{j,\text{L-SUB-CE}}) = \frac{\text{SINR}_{j,\text{L-MOE}}}{1+\frac{J-1}{K}\text{SINR}_{j,\text{L-MOE}}}. \tag{4.163}$$

Our simulation results show that the SINR performances of the L-SMI-CE and L-SUB-CE receivers are accurately described by (4.162) and (4.163) even when the SNR assumes moderate values. Due to the similarity between the SINR expressions obtained for L and WL receivers, most observations regarding the comparison between receivers with or without CE apply also in this case. Summarizing, the SINR of the L-SUB-CE receiver turns out to be (approximately) equal to that of the L-SUB one (4.102). Moreover, due to the mentioned signature mismatch problem [79], the SINR of the L-SMI receiver with known channel (4.100) is worse than that of the corresponding L-SMI-CE receiver: indeed, for a finite sample size K, in the high-SNR regime, it results that

$$\lim_{\sigma_v^2\to 0}\frac{\text{SINR}_{j,\text{L-SMI}}}{\text{SINR}_{j,\text{L-SMI-CE}}} = \frac{N+J-L_j-1}{N+J-2} < 1\,. \tag{4.164}$$

Additionally, similarly to the the WL case, the data-estimated linear receivers exhibit a SINR saturation effect, for $\sigma_v^2 \to 0$. In this case, if $\mathbf{\Phi}$ is full-column rank ($J \leq N$), it is readily verified that $\text{SINR}_{j,\text{L-MOE}} \to +\infty$. Henceforth, for $J \leq N - L_j < N$, accounting for (4.162) and (4.163), one obtains

$$\lim_{\sigma_v^2\to 0}\text{SINR}_{j,\text{L-SMI-CE}} = \frac{2K}{N+J-L_j-1}\,, \tag{4.165}$$

$$\lim_{\sigma_v^2\to 0}\text{SINR}_{j,\text{L-SUB-CE}} = \frac{K}{J-1}\,, \tag{4.166}$$

which show that, in the high-SNR regime, the performance of the L-SMI-CE receiver depends on the processing gain N and the number of users J, as well as on the channel length L_j of the jth user, whereas the performance of the L-SUB-CE receiver is independent of both the processing gain N and the channel length L_j, while depending on the number of users J. At this point, we are able to establish a direct comparison between blind L- and WL-MOE data-estimated receivers, focusing our attention on the case $J \leq J_{\text{max,L}} = N - L_j$, wherein both blind L- and WL-MOE receivers can work [note indeed that

the WL-MOE can accommodate up to $J_{\text{max,WL}} = 2(N - L_j)$ users (see eq. (4.106))]. By comparing (4.132) and (4.163) for the subspace-based receivers, it turns out that $\text{SINR}_{j,\text{WL-SUB-CE}} \geq \text{SINR}_{j,\text{L-SUB-CE}}$ for any value of K. Instead, for the SMI-based receivers [see (4.131) and (4.162)], it results that $\text{SINR}_{j,\text{WL-SMI-CE}} \geq \text{SINR}_{j,\text{L-SMI-CE}}$ only when $K \geq K_{\text{min}}$, where

$$K_{\text{min}} \triangleq \frac{3(N - L_j) - J + 1}{2(\text{SINR}_{j,\text{L-MOE}}^{-1} - \text{SINR}_{j,\text{WL-MOE}}^{-1})} \tag{4.167}$$

is a threshold sample size, that is, similarly to the known channel case, the WL-SMI-CE receiver assures a performance advantage only by processing a sufficient number of samples.[17] Finally, for $J \leq N - L_j$, as regards the comparison between the saturation SINR (i.e., the SINR for $\sigma_v^2 \to 0$) of blind linear and WL receivers, it can be observed from (4.165) and (4.134) that the value for the L-SMI-CE receiver is better than the corresponding value for WL-SMI-CE, whereas, according to (4.166) and (4.135), the saturation SINRs for the subspace-based receivers are exactly coincident.

4.5.4 Numerical results

In this section, Monte Carlo simulations are presented, aimed at validating and extending our performance analysis. We consider a DS-CDMA system with $\alpha_1 = \alpha_2 = \cdots = \alpha_J = 1$ and $N = 16$. The J users employ unit-norm (i.e., $\|\mathbf{c}_j\| = 1$) random signatures $\mathbf{c}_j$, whose entries are i.i.d. random variables assuming equiprobable values in the complex set $\{\pm 1/\sqrt{2N}, \pm i/\sqrt{2N}\}$, with $\mathbf{c}_{j_1}$ and $\mathbf{c}_{j_2}$ statistically independent of each other for $j_1 \neq j_2 \in \{1, 2, \ldots, J\}$. The channel lengths are $L_j = 5$, $\forall j \in \{1, 2, \ldots, J\}$, i.e., they are equal for all the users, and, as in [66, 76], the entries of the unit-norm channel vectors $\mathbf{g}_j$ are randomly and independently drawn with equal power from a zero-mean complex circular (or proper) Gaussian process. The symbol and noise sequences are generated according to assumptions **(a1)** and **(a2)**, and the SNR is defined as $1/\sigma_v^2$. In each simulation, we carry out 10^4 independent Monte Carlo runs, with each run employing a different set of spreading sequences, channel vectors, symbol sequences and noise. In all simulations, we assume that the users have identical powers, i.e. there is perfect power control, and, without loss of generality, that the desired user is the first one, i.e., $j = 1$. Note that, in the considered scenario, the maximum number of users that can be accommodated

[17] A comparison with eq. (4.103) shows that, in the estimated-channel case, the value of K_{min} is slightly lower.

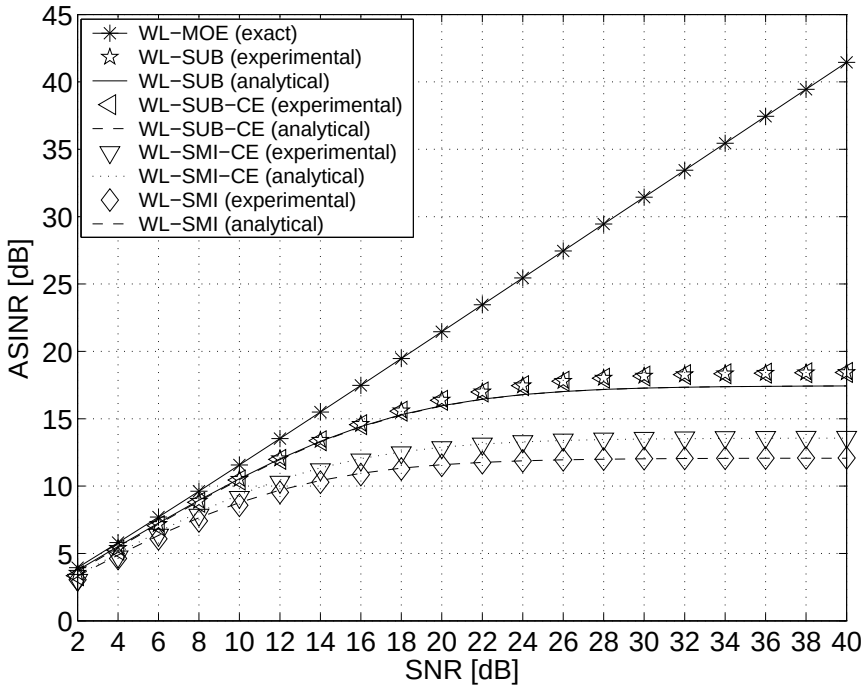

Figure 4.7: ASINR versus SNR for WL-MOE receivers ($J = 10$ users and $K = 500$ symbols).

by the blind receivers with CE is equal to $J_{\text{max,L}} = 11$ for the L-MOE receivers and $J_{\text{max,WL}} = 22$ for the WL-MOE receivers. To extensively compare WL-MOE and L-MOE receivers, we assume that J satisfy the first, more stringent condition, exception made for the second experiment, where we evaluate the performances as a function of J.

Experiment 4.4 *:* in this experiment, we evaluate the average SINR (ASINR) as a function of SNR for the WL-MOE (Fig. 4.7) and L-MOE (Fig. 4.8) receivers (both with and without CE), for $J = 10$ users and a sample size equal to $K = 500$ symbols. For the sake of comparison, we also report the ASINR of the exact (i.e., data-independent) WL-MOE and L-MOE receivers given by (4.27) and (4.16), respectively. All the curves show a good agreement between simulation and analytical results. Looking in detail at Fig. 4.7, the simulation results confirm the theoretical prediction that the two subspace versions of the WL-MOE receiver (with or without CE) exhibit practically

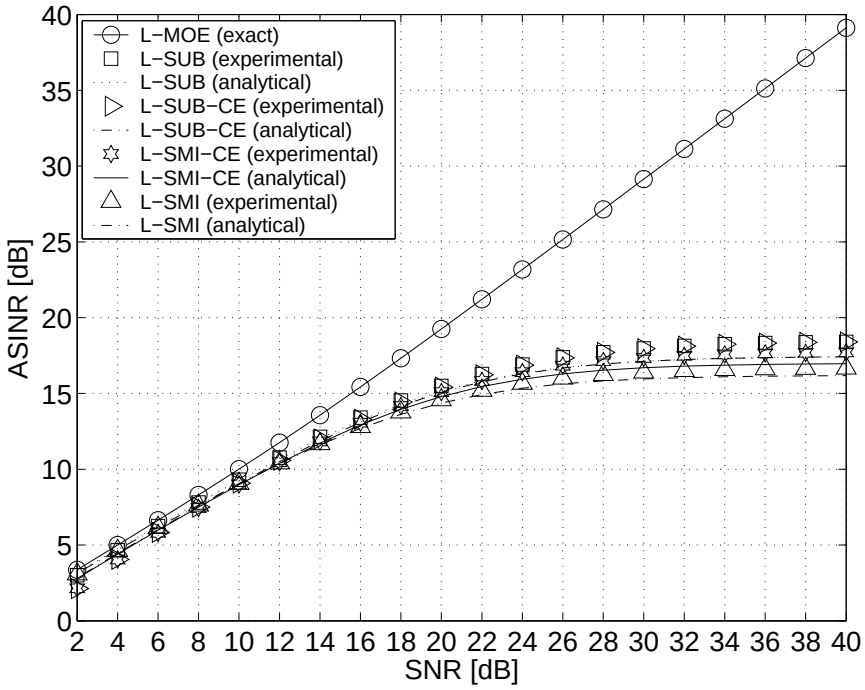

Figure 4.8: ASINR versus SNR for L-MOE receivers ($J = 10$ users and $K = 500$ symbols).

the same performances, whereas the WL-SMI-CE receiver performs slightly better than the WL-SMI one (with known channel), since the latter is penalized by the signature mismatch phenomenon; in particular, the asymptotic (for SNR $\rightarrow +\infty$) difference between the ASINR curves of the WL-SMI-CE and WL-SMI receivers is about 1.5 dB, which is in good agreement with the value theoretically predicted by (4.133). Similar considerations apply to Fig. 4.8, where the asymptotic gain of the L-SMI-CE receiver over the L-SMI one (with known channel) is about 1 dB, as correctly predicted by (4.164). As regards the comparison between WL-MOE and L-MOE receivers, results of Figs. 4.7 and 4.8 allow us to extend an important conclusion done in the known-channel case , relative to the underloaded case (i.e., $J \leq N$): although the exact WL-MOE receiver generally exhibits a SINR gain over the L-MOE one also when $J \leq J_{\max,\mathrm{L}}$, in practice, due to SINR saturation effects, the subspace implementations of the WL-MOE and L-MOE receivers exhibit the same perfor-

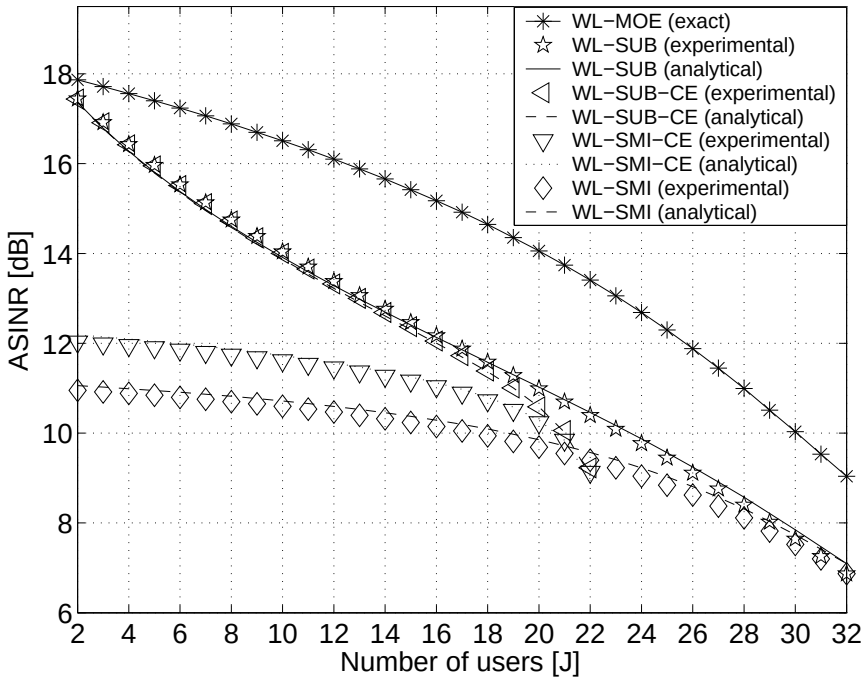

Figure 4.9: ASINR versus number of users J for WL-MOE receivers (SNR = 15 dB and $K = 500$ symbols).

mances, whereas the L-SMI receivers (both with and without CE) outperform their WL-SMI counterparts.

Experiment 4.5 *:* in this experiment, we evaluate the ASINR as a function of the number of users J for the WL-MOE (Fig. 4.9) and L-MOE (Fig. 4.10) receivers (both with and without CE), for a sample size equal to $K = 500$ symbols and SNR = 15 dB. Since the subspace-based CE procedure poses a strict limit of $J_{\text{max,WL}} = 22$ users for the WL-MOE receivers and $J_{\text{max,L}} = 11$ for the L-MOE receivers with CE, the performances of the receivers with CE are not reported (i.e., the corresponding curves are truncated) for values of J exceeding these limits. Besides confirming again a good agreement between simulation and analytical results, the curves for the WL-MOE receivers (Fig. 4.9) show that the performance advantage of the WL-SUB receiver over the WL-SMI one (both with and without CE) progressively decreases as J increases, becoming negligible in correspondence of about $J = 20$ users for the

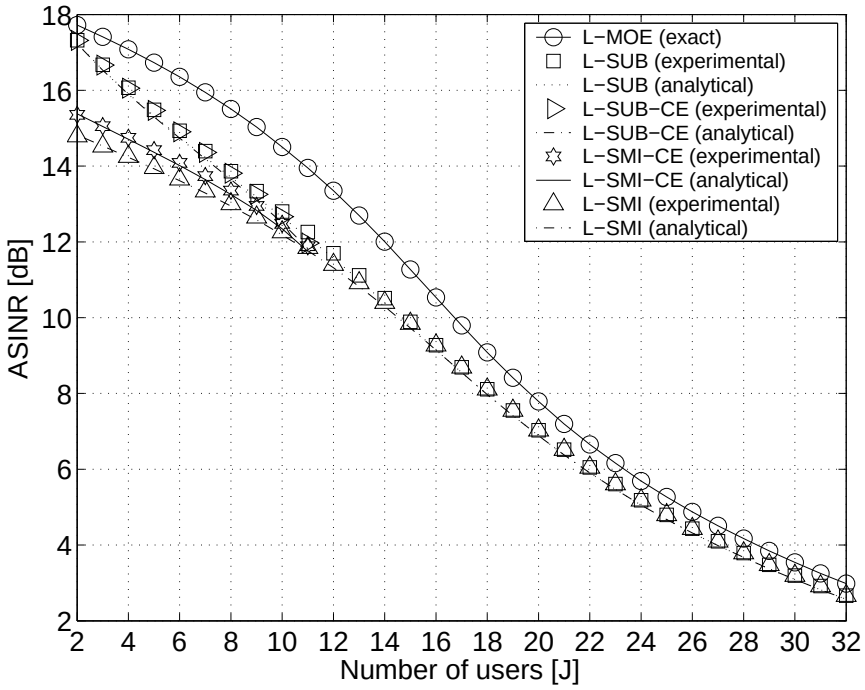

Figure 4.10: ASINR versus number of users J for L-MOE receivers (SNR = 15 dB and $K = 500$ symbols).

receivers with CE, and $J = 30$ users for the receivers with known channel. It is worthwhile to observe, moreover, that when J approaches the upper limit $J_{\text{max,WL}} = 22$ for CE, the performances of the WL-MOE receivers with CE degrade rapidly, suffering from a clear threshold effect. Similar considerations apply to Fig. 4.10, where, however, the ASINR curves of the L-MOE receivers are more closely spaced and the performance advantage of the L-SUB receiver over the L-SMI one becomes negligible in correspondence of about $J = 10$ users for the receivers with CE, and $J = 14$ users for the receivers with known channel. A careful comparison between the performances of WL-MOE and L-MOE receivers shows again that the largest advantage in using WL-MOE receivers is obtained in the "overloaded" region, i.e., when $11 \leq J \leq 22$ for the receivers with CE (where the L-MOE receivers cannot operate at all), and when $16 \leq J \leq 32$ for the receivers with known channel (where the L-MOE receivers, although capable of operating, exhibit poor performances).

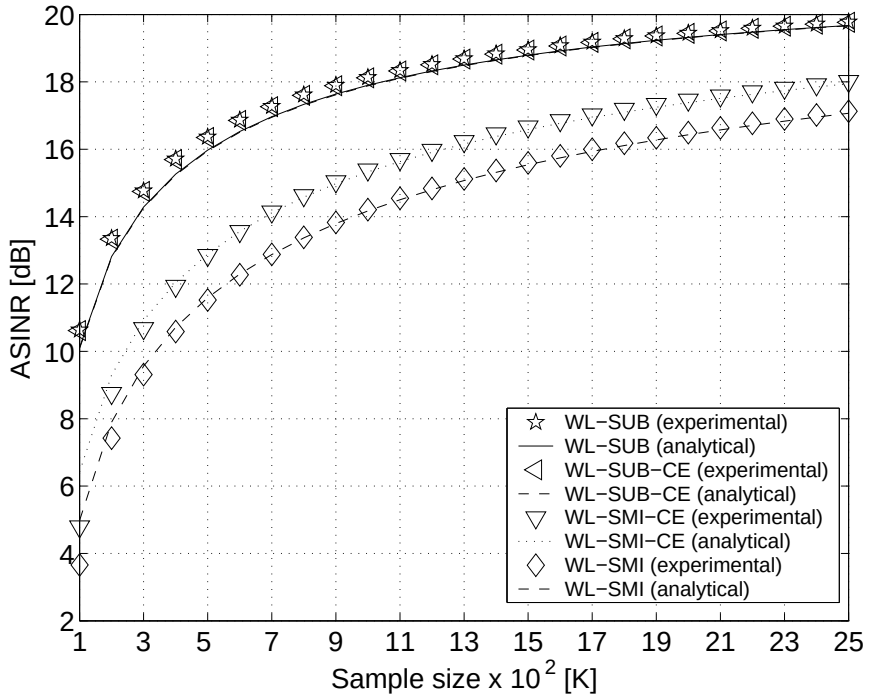

Figure 4.11: ASINR versus sample size K for WL-MOE receivers ($J = 10$ users and SNR = 20 dB).

Experiment 4.6 *:* in this last experiment, we report the ASINR as a function of the sample size K for the WL-MOE (Fig. 4.11) and L-MOE (Fig. 4.12) receivers (both with and without CE), for $J = 10$ users and SNR = 20 dB. The ASINR values of the exact (i.e, data-independent) WL-MOE and L-MOE receivers, in this scenario, are equal to 21.5 and 19.2 dB, respectively, and obviously do not depend on K. The simulation and analytical results are again in good agreement, and, as expected, the accuracy of the formulas (4.131)–(4.132) and (4.162)–(4.163) improves as K increases. In particular, Fig. 4.11 shows that the two versions of the WL-SUB receivers (with or without CE) exhibit almost the same performances, outperforming the WL-SMI-CE receiver by about 2 dB, and the WL-SMI one (with known channel) by about 3 dB, for all considered values of K. Instead, the ASINR curves of the L-MOE receivers (see Fig. 4.12) are more closely spaced, exhibiting only marginal differences in performances between the various receivers. By comparing Figs. 4.11

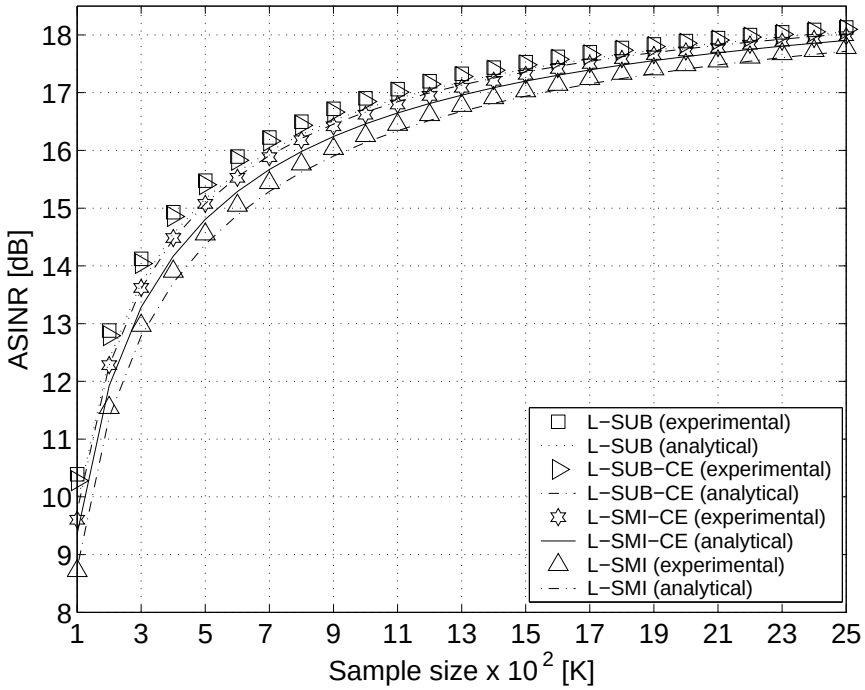

Figure 4.12: ASINR versus sample size K for L-MOE receivers ($J = 10$ users and SNR = 20 dB).

and 4.12, it can be seen that the two WL-SUB receivers (with or without CE) outperform the corresponding L-SUB ones, for all the considered values of K. In contrast, the WL-SMI receiver (with known channel) again performs worse than its linear counterpart for all values of K (in this case the threshold sample size (4.103) is $K_{\min} = 3844$, thus larger than the maximum value of $K = 2500$ considered in the simulations), whereas the performances of the WL-SMI-CE receiver approaches those of the L-SMI-CE one for K approaching 2500, which agrees very well with the value $K_{\min} = 2428$ predicted by (4.167).

Chapter 5

Equalization Techniques for MC-CDMA Systems

This chapter focuses on multiuser detection for downlink MC-CDMA systems, employing cyclic-prefixed (CP) or zero-padded (ZP) transmission techniques. For both systems, we consider the linear and widely linear FIR receiving structures, showing that if the number of users does not exceed a given threshold and their codes are appropriately designed, L-FIR and WL-FIR universal zero-forcing (ZF) multiuser detectors can be synthesized. Thus in the absence of noise, it is assured a perfect symbol recovery for each user, regardless of the underlying frequency-selective channel. Moreover, some spreading code examples are provided, which satisfy the design rules. Finally, numerical simulations are carried out to show that the theoretical considerations developed herein provide useful guidelines for practical MC-CDMA system designs.

5.1 Introduction

In the previous chapter 4, we dealt with single-carrier DS-CDMA technology which nevertheless, at high data-rates, becomes impractical, due to both severe multipath-induced intersymbol interference (ISI) and synchronization difficulties. Thus, we chose a multicarrier scheme to overcame these drawbacks, as MC-CDMA technology [11, 83, 84]. Indeed MC-CDMA systems, employing *frequency-domain spreading* [83], which consists of copying each information symbol over the N subcarriers and multiplying it by a user-specific vector code, achieve ISI mitigation more efficiently than DS-CDMA systems because they transmit with a lower data-rate over multiple subcarriers.Moreover, due to

the lowered symbol rate, the synchronization task is easier in MC-CDMA networks, compared with a DS-CDMA system with similar processing gain. Furthermore, it has been shown in [85] that, at the expense of a reduced bandwidth efficiency, MUD techniques offer higher near-far resistance in MC-CDMA systems than in DS-CDMA ones.

Several MC-CDMA transmission schemes have been proposed in the literature, among which those based on *cyclic prefix* (CP) and *zero padding* (ZP) precoding techniques, items of interest in this chapter. In conventional CP-MC-CDMA systems, after multiplying each information symbol by a user-specific vector code, the resulting vector is subject to inverse fast Fourier transform (IFFT) and, finally, a CP of length L_{p} larger than the channel order L is inserted; at the receiver, the CP is discarded and the remaining part of the MC-CDMA symbol turns out to be free of interblock interference (IBI). Instead in ZP-MC-CDMA systems, for achieving deterministic IBI cancellation, the CP is replaced with ZP, by appending $L_{\mathrm{p}} > L$ zero symbols to each IFFT-precoded symbol block; in this case, IBI suppression is obtained without discarding any portion of the received signal. If the number of zero symbols is equal to the CP length, then CP- and ZP-based systems exhibit the same spectral efficiency. ZP precoding technique has been originally proposed [86] for OFDM systems since, unlike CP-based transmissions, it enables L-FIR perfect symbol recovery, even when the channel transfer function has nulls on (or close to) some subcarriers. Compared with CP precoding, the price to pay for such a capability is the slightly increased receiver complexity and the larger power amplifier backoff.

It is worth to observe that many L- and WL-MUD techniques, which are proposed in the DS-CDMA context 4, can be readily adapted to MC-CDMA systems. To suppress MAI with an affordable computational complexity and, simultaneously, achieve close-to-optimality performance (in the minimum bit-error-rate sense), one can resort to the FIR L-ZF (or linear decorrelating) and linear minimum mean-square error (L-MMSE) receivers. In addition we have underlined still in the previous chapters of this thesis (see section 2.4 in the chapter 2 and chapters 3,4 for more details) that by exploiting the possible improper or noncircular nature of the transmitted symbols, improved MAI suppression capabilities can be attained by adopting WL-FIR receiving structures, such as the WL-ZF and WL-MMSE receivers.

With reference to FIR L-MUD receiving structures, it is known [84, 87] that perfect symbol recovery is guaranteed in a ZP-based downlink, for any FIR channel of order $L < L_{\mathrm{p}}$, as long as the number of users is smaller than

the number of subcarriers (*underloaded systems*) and the code vectors are linearly independent. In general, a similar feature does not hold for CP-based transmissions. Thus, we show in section 5.3, following our paper [18], that universal L-ZF-MUD can be guaranteed even for the underloaded CP-MC-CDMA downlink, provided that the spreading codes are judiciously designed. On the other hand, a detailed study of the conditions assuring FIR WL-MUD perfect symbol recovery in both CP- and ZP-based systems is lacking. Consequently, we also show 5.4, following our papers [17, 18], that, if appropriate complex-valued spreading codes are employed, universal WL-ZF multiuser detectors can be designed even for overloaded CP-MC-CDMA and ZP-MC-CDMA systems[1].

5.2 Models for CP- and ZP-MC-CDMA systems

Let us consider the downlink of a MC-CDMA system employing N subcarriers and accommodating J users. The information symbol $b_j(n)$ emitted by the jth user in the nth ($n \in \mathbb{Z}$) symbol interval multiplies the *frequency-domain* spreading code $\mathbf{c}_j \stackrel{\triangle}{=} [c_j^{(0)}, c_j^{(1)}, \ldots, c_j^{(N-1)}]^T \in \mathbb{C}^N$, with $c_j^{(m)} \neq 0$, $\forall m \in \{0, 1, \ldots, N-1\}$ and $\forall j \in \{1, 2, \ldots, J\}$. The resulting N-length sequence is subject to the inverse discrete Fourier transform (IDFT), producing thus the block

$$\widetilde{\mathbf{u}}_j(n) = \mathbf{W}_{\text{IDFT}}\, \mathbf{c}_j\, b_j(n) \in \mathbb{C}^N, \tag{5.1}$$

where $\mathbf{W}_{\text{IDFT}} \in \mathbb{C}^{N\times N}$ denotes the IDFT matrix, with (m_1, m_2)th entry $\frac{1}{\sqrt{N}}\, e^{i\frac{2\pi}{N}\,(m_1-1)\,(m_2-1)}$, for $m_1, m_2 \in \{1, 2, \ldots, N\}$, and its inverse $\mathbf{W}_{\text{DFT}} \stackrel{\triangle}{=} \mathbf{W}_{\text{IDFT}}^{-1} = \mathbf{W}_{\text{IDFT}}^{H}$ defines the discrete Fourier transform (DFT) matrix. The equation (5.1) is different depending on if we adopt a CP- or ZP-linear precoding strategies:

Cyclic prefixing: after computing the IDFT, a CP of length $L_{\text{p}} \ll N$ is inserted at the beginning of $\widetilde{\mathbf{u}}_j(n)$, obtaining thus the vector

$$\mathbf{u}_{\text{cp},j}(n) = \mathbf{T}_{\text{cp}}\, \mathbf{W}_{\text{IDFT}}\, \mathbf{c}_j\, b_j(n)\,, \tag{5.2}$$

where $\mathbf{T}_{\text{cp}} \stackrel{\triangle}{=} [\mathbf{I}_{\text{cp}}^T, \mathbf{I}_N]^T \in \mathbb{R}^{P\times N}$ and $P \stackrel{\triangle}{=} L_{\text{p}} + N$, with $\mathbf{I}_{\text{cp}} \in \mathbb{R}^{L_{\text{p}}\times N}$ built by drawing out the last L_{p} rows of the identity matrix $\mathbf{I}_N$.

[1] It is worthwhile to observe that overloaded systems are of practical interest [88], for example, in bandwidth-efficient multiuser communication, where the bandwidth is at a premium.

Zero padding: after computing the IDFT, $L_{\rm p} \ll N$ trailing zeros are padded at the end of $\widetilde{\mathbf{u}}_j(n)$, obtaining thus the vector

$$\mathbf{u}_{{\rm zp},j}(n) = \mathbf{T}_{\rm zp}\, \mathbf{W}_{\rm IDFT}\, \mathbf{c}_j\, b_j(n)\,, \tag{5.3}$$

where $\mathbf{T}_{\rm zp} \stackrel{\triangle}{=} [\mathbf{I}_N, \mathbf{O}_{N\times L_{\rm p}}]^T \in \mathbb{R}^{P\times N}$.

In either cases, the blocks $\mathbf{u}_{{\rm cp},j}(n)$ and $\mathbf{u}_{{\rm zp},j}(n)$ are subject to parallel-to-serial conversion, and the resulting sequences feed a digital-to-analog converter, operating at rate $1/T_c = P/T_s$, where T_s and T_c denote the symbol and the sampling period, respectively. In the downlink, all the users are synchronous and propagate through a common frequency-selective channel that is modeled as a linear time-invariant system, whose channel impulse response $g_c(t)$ (including transmitting filter, physical channel and receiving filter) is *complex-valued*, that is, neither $\mathrm{Re}\{g_c(t)\}$ nor $\mathrm{Im}\{g_c(t)\}$ vanish identically, and spans $L+1$ sampling periods, i.e., $g_c(t) \equiv 0$, $\forall t \notin [0, L\,T_c]$, where $g_c(0), g_c(L\,T_c) \neq 0$, with $L < P$ within one symbol interval. In this case, the discrete-time channel $g(\ell) \stackrel{\triangle}{=} g_c(\ell\,T_c)$ turns out to be a FIR filter of order L, i.e., $g(\ell) \equiv 0$, $\forall \ell \notin \{0, 1, \ldots, L\}$, with $g(0), g(L) \neq 0$. Furthermore, we assume that the channel order L is not exactly known, but is upper bounded by $L_{\rm p}$, i.e., $L < L_{\rm p}$.

In a CP-based system, the IBI is deterministically removed by discarding the first $L_{\rm p}$ samples of each P-dimensional received block. Indeed, after CP removal, the kth ($k \in \mathbb{Z}$) received symbol block $\mathbf{r}_{\rm cp}(k) \in \mathbb{C}^N$ can be expressed (see, e.g., [84, 86]) as

$$\mathbf{r}_{\rm cp}(k) = \underbrace{\mathbf{\Theta}_{\rm cp}\, \mathbf{W}_{\rm IDFT}\, \mathbf{C}}_{\boldsymbol{\mathcal{G}}_{\rm cp}\in\mathbb{C}^{N\times J}}\, \mathbf{b}(k) + \mathbf{v}_{\rm cp}(k) = \boldsymbol{\mathcal{G}}_{\rm cp}\, \mathbf{b}(k) + \mathbf{v}_{\rm cp}(k)\,, \tag{5.4}$$

where $\mathbf{\Theta}_{\rm cp} \in \mathbb{C}^{N\times N}$ is the circulant [56] matrix having $\mathbf{\Omega}_{\rm cp}\,\mathbf{g}$ as its first column, with $\mathbf{\Omega}_{\rm cp} \stackrel{\triangle}{=} [\mathbf{I}_{L_{\rm p}}, \mathbf{O}_{L_{\rm p}\times(N-L_{\rm p})}]^T \in \mathbb{R}^{N\times L_{\rm p}}$ and $\mathbf{g} \stackrel{\triangle}{=} [g(0), g(1), \ldots, g(L), 0, \ldots, 0]^T \in \mathbb{C}^{L_{\rm p}}$, the vector $\mathbf{b}(k) \stackrel{\triangle}{=} [b_1(k), b_2(k), \ldots, b_J(k)]^T \in \mathbb{C}^J$ collects the symbols transmitted by the users as in the chapter 4, $\mathbf{C} \stackrel{\triangle}{=} [\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_J] \in \mathbb{C}^{N\times J}$ defines the frequency-domain *code matrix* and, finally, vector $\mathbf{v}_{\rm cp}(k) \in \mathbb{C}^N$ accounts for thermal noise.

In contrast, ZP-based precoding allows one to deterministically eliminate the IBI by retaining all the samples of each P-dimensional received block. Specifically, in a ZP-based system, the kth received symbol block $\mathbf{r}_{\rm zp}(k) \in \mathbb{C}^P$

is given by (see, e.g., [84, 86])

$$\mathbf{r}_{\text{zp}}(k) = \underbrace{\boldsymbol{\Theta}_{\text{zp}}\,\mathbf{W}_{\text{IDFT}}\,\mathbf{C}}_{\boldsymbol{\mathcal{G}}_{\text{zp}}\in\mathbb{C}^{P\times J}}\,\mathbf{b}(k) + \mathbf{v}_{\text{zp}}(k) = \boldsymbol{\mathcal{G}}_{\text{zp}}\,\mathbf{b}(k) + \mathbf{v}_{\text{zp}}(k)\,, \qquad (5.5)$$

where $\boldsymbol{\Theta}_{\text{zp}} \in \mathbb{C}^{P\times N}$ is the Toeplitz [56] matrix having $\boldsymbol{\Omega}_{\text{zp}}\,\mathbf{g}$ as first column, with $\boldsymbol{\Omega}_{\text{zp}} \triangleq [\mathbf{I}_{L_{\text{p}}}, \mathbf{O}_{L_{\text{p}}\times(P-L_{\text{p}})}]^T \in \mathbb{R}^{P\times L_{\text{p}}}$, and $[g(0), 0, \ldots, 0]$ as first row, whereas $\mathbf{v}_{\text{zp}}(k) \in \mathbb{C}^P$ accounts for thermal noise. For the sake of conciseness, we unify models (5.4) and (5.5) in the equivalent one

$$\mathbf{r}(k) \;=\; \boldsymbol{\mathcal{G}}\,\mathbf{b}(k) \;+\; \mathbf{v}(k), \;\; \text{with } \mathbf{r}(k), \mathbf{v}(k) \in \mathbb{C}^R \text{ and } \boldsymbol{\mathcal{G}} \in \mathbb{C}^{R\times J}, \quad (5.6)$$

where, for a CP-based system, $\mathbf{r}(k) = \mathbf{r}_{\text{cp}}(k)$, $\boldsymbol{\mathcal{G}} = \boldsymbol{\mathcal{G}}_{\text{cp}}$, $\mathbf{v}(k) = \mathbf{v}_{\text{cp}}(k)$, with $R = N$, whereas, for a ZP-based system, $\mathbf{r}(k) = \mathbf{r}_{\text{zp}}(k)$, $\boldsymbol{\mathcal{G}} = \boldsymbol{\mathcal{G}}_{\text{zp}}$, $\mathbf{v}(k) = \mathbf{v}_{\text{zp}}(k)$, with $R = P$. Hereinafter, we assume that:

a1) the transmitted symbols $b_j(n)$ are modeled as mutually independent zero-mean and independent identically-distributed (iid) random sequences, with second-order moments $\sigma_b^2 \triangleq \mathrm{E}[|b_j(n)|^2] > 0$ and $\varrho_b(n) \triangleq \mathrm{E}[b_j^2(n)]$;

a2) the noise vector $\mathbf{v}(k)$ is a zero-mean wide-sense stationary complex *proper* white random process, which is independent of $b_j(n)$, $\forall j \in \{1, 2, \ldots, J\}$, with autocorrelation matrix $\mathbf{R}_{\mathbf{vv}} \triangleq \mathrm{E}[\mathbf{v}(k)\,\mathbf{v}^H(k)] = \sigma_v^2\,\mathbf{I}_R$.

As regards assumption a1, we have still observed (see previous chapters) that, there exists a large family of modulation schemes of practical interest, such as, BPSK, DBPSK, M-ASK, OQPSK, OQAM, and binary CPM, MSK, GMSK, which are improper, i.e., $\varrho_b(n) \neq 0$, for any $n \in \mathbb{Z}$.

5.3 Perfect symbol recovery for L-MUD

In this section we consider, the problem of ZF detectability in FIR L-MUD, which can be used for both CP- and ZP-based systems, employing either proper or improper data symbols (although it is suboptimal in the latter case). These theoretical aspects strongly affect both the synthesis and the performance analysis of the L-ZF and L-MMSE multiuser detectors [45, 44].

To detect the transmitted symbol $b_j(k)$ of the jth user from the received vector (5.6), with $j \in \{1, 2, \ldots, J\}$, with a FIR L-MUD, we apply the input-output relationship that we have seen in chapter 4

$$y_j(k) = \mathbf{f}_j^H \mathbf{r}(k), \tag{5.7}$$

with $\mathbf{f}_j \in \mathbb{C}^R$ (which is followed by a decision device). In the absence of noise, the perfect or ZF symbol recovery condition $y_j(k) = b_j(k)$ leads to the system of linear equations $\boldsymbol{\mathcal{G}}^H \mathbf{f}_j = \mathbf{e}_j$, where $\mathbf{e}_j \stackrel{\triangle}{=} [\mathbf{0}_{j-1}^T, 1, \mathbf{0}_{J-j}^T]^T \in \mathbb{R}^J$, which is consistent (i.e., it admits at least one solution) for each user if and only if (iff) the *composite* channel matrix $\boldsymbol{\mathcal{G}}$ is full-column rank, i.e., $\text{rank}(\boldsymbol{\mathcal{G}}) = J$. Under this assumption, the *minimal norm* [24] solution of $\boldsymbol{\mathcal{G}}^H \mathbf{f}_j = \mathbf{e}_j$ is given by

$$\mathbf{f}_{\text{L-ZF},j} = \boldsymbol{\mathcal{G}} \, (\boldsymbol{\mathcal{G}}^H \boldsymbol{\mathcal{G}})^{-1} \mathbf{e}_j \, , \tag{5.8}$$

which defines the L-ZF or linear decorrelating multiuser detector. In the presence of noise, the L-ZF receiver perfectly suppresses the MAI at the price of noise enhancement. To better counteract the noise, one can resort to the L-MMSE multiuser detector [45, 44], which is defined as

$$\mathbf{f}_{\text{L-MMSE},j} = \arg \min_{\mathbf{f}_j \in \mathbb{C}^R} \text{E}[|b_j(k) - y_j(k)|^2] = \sigma_b^2 \, \mathbf{R}_{\mathbf{rr}}^{-1} \, \boldsymbol{\mathcal{G}} \, \mathbf{e}_j \, , \tag{5.9}$$

where $\mathbf{R}_{\mathbf{rr}} \stackrel{\triangle}{=} \text{E}[\mathbf{r}(k) \, \mathbf{r}^H(k)] \in \mathbb{C}^{R \times R}$ is the autocorrelation matrix of $\mathbf{r}(k)$ which, accounting for (5.6), and invoking assumptions a1 and a2, is given by

$$\mathbf{R}_{\mathbf{rr}} = \sigma_b^2 \, \boldsymbol{\mathcal{G}} \, \boldsymbol{\mathcal{G}}^H + \sigma_v^2 \, \mathbf{I}_R \, . \tag{5.10}$$

If $\boldsymbol{\mathcal{G}}$ is full-column rank, by resorting to the limit formula for the Moore-Penrose inverse [24], it can be seen that

$$\begin{aligned} \lim_{\sigma_v^2/\sigma_b^2 \to 0} \mathbf{f}_{\text{L-MMSE},j} &= \lim_{\sigma_v^2/\sigma_b^2 \to 0} \sigma_b^2 \, (\sigma_b^2 \, \boldsymbol{\mathcal{G}} \, \boldsymbol{\mathcal{G}}^H + \sigma_v^2 \, \mathbf{I}_R)^{-1} \, \boldsymbol{\mathcal{G}} \, \mathbf{e}_j \\ &= \boldsymbol{\mathcal{G}} \, (\boldsymbol{\mathcal{G}}^H \boldsymbol{\mathcal{G}})^{-1} \mathbf{e}_j = \mathbf{f}_{\text{L-ZF},j} \, , \end{aligned} \tag{5.11}$$

$\forall j \in \{1, 2, \ldots, J\}$, i.e., the L-MMSE receiver boils down to the L-ZF one[2]. In summary, the performance of the L-MMSE receiver in the high signal-to-noise

[2] More generally, when $\boldsymbol{\mathcal{G}}$ is possibly rank-deficient, it results that $\lim_{\sigma_v^2/\sigma_b^2 \to 0} \mathbf{f}_{\text{L-MMSE},j} = (\boldsymbol{\mathcal{G}}^H)^\dagger \, \mathbf{e}_j \stackrel{\triangle}{=} \mathbf{f}_{\text{L-LS},j}$, i.e., the L-MMSE detector ends up to the minimal-norm least-squares solution [24] of $\boldsymbol{\mathcal{G}}^H \mathbf{f}_j = \mathbf{e}_j$ [note that, when $\boldsymbol{\mathcal{G}}$ is full-column rank, one has $\mathbf{f}_{\text{L-LS},j} = \mathbf{f}_{\text{L-ZF},j}$ from (5.8)].

(SNR) region strongly depends on the existence of L-ZF solutions: indeed, if $\boldsymbol{\mathcal{G}}$ is not full-column rank, the performance curve of the L-MMSE multiuser detector exhibits a marked bit-error-rate (BER) floor (see Section 5.5), when $\sigma_v^2/\sigma_b^2 \to 0$. Motivated by this fact, the first step of our study consists of investigating whether the condition $\text{rank}(\boldsymbol{\mathcal{G}}) = J$ is satisfied, regardless of the frequency-selective channel.

As a matter of fact, for a ZP-based system [see (5.5)], the rank properties of $\boldsymbol{\mathcal{G}} = \boldsymbol{\mathcal{G}}_{\text{zp}} = \boldsymbol{\Theta}_{\text{zp}}\,\mathbf{W}_{\text{IDFT}}\,\mathbf{C}$ are easily characterized, since the Toeplitz matrix $\boldsymbol{\Theta}_{\text{zp}}$ is full-column rank for any FIR channel of order L [84, 87, 89]. Indeed, owing to nonsingularity of $\mathbf{W}_{\text{IDFT}}$, it results that

$$\text{rank}(\boldsymbol{\mathcal{G}}_{\text{zp}}) = \text{rank}(\mathbf{C}) \tag{5.12}$$

As stated in [87], the composite channel matrix $\boldsymbol{\mathcal{G}}_{\text{zp}}$ is always full-column rank and, thus, channel-irrespective L-FIR perfect symbol recovery is possible iff the vectors $\mathbf{c}_1$, $\mathbf{c}_2, \ldots, \mathbf{c}_J$ are linearly independent, that is, $\mathbf{C}$ is full-column rank. To this aim, one can for example use Walsh-Hadamard (WH) spreading codes, which are widely used in CDMA systems. This issue is analyzed in section 4.4.1 of the chapter 4. It is worth noting that condition $\text{rank}(\mathbf{C}) = J$ imposes that the number of users be smaller than or equal to the number of subcarriers, i.e., $J \leq N$: strictly speaking, L-ZF-MUD is exclusively targeted at underloaded systems. On the other hand, for a CP-based system [see (5.4)], the linear independence of the code vectors is not sufficient to assure that $\boldsymbol{\mathcal{G}} = \boldsymbol{\mathcal{G}}_{\text{cp}} = \boldsymbol{\Theta}_{\text{cp}}\,\mathbf{W}_{\text{IDFT}}\,\mathbf{C}$ be always full-column rank since, unlike $\boldsymbol{\Theta}_{\text{zp}}$, the circulant matrix $\boldsymbol{\Theta}_{\text{cp}}$ turns out to be singular for some FIR channels. However, after characterizing the rank properties of $\boldsymbol{\mathcal{G}}_{\text{cp}}$, we show [18] in Subsection 5.3.1 that, through appropriate design of user codes, the condition $\text{rank}(\boldsymbol{\mathcal{G}}_{\text{cp}}) = J$ can be guaranteed regardless of the underlying frequency-selective channel. Hence, channel-irrespective L-ZF-MUD is possible not only in a ZP-based system, but also in a CP-based one.

5.3.1 Rank characterization of $\boldsymbol{\mathcal{G}}_{\text{cp}}$ and universal code design for L-ZF-MUD

With reference to a CP-based system, let us study the rank properties of $\boldsymbol{\mathcal{G}}_{\text{cp}} = \boldsymbol{\Theta}_{\text{cp}}\,\mathbf{W}_{\text{IDFT}}\,\mathbf{C}$. Preliminarily, observe that $\boldsymbol{\mathcal{G}}_{\text{cp}}$ is full-column rank only if the number J of users is not larger than the number of subcarriers N, i.e., $J \leq N$. Thus, as for a ZP-based system, L-ZF-MUD is confined only to underloaded CP-based systems. Furthermore, by resorting to standard

eigenstructure concepts [84, 56], one has $\mathbf{\Theta}_{\text{cp}} = \mathbf{W}_{\text{IDFT}}\,\mathbf{\Gamma}_{\text{cp}}\,\mathbf{W}_{\text{DFT}}$, where the diagonal entries of $\mathbf{\Gamma}_{\text{cp}} \triangleq \text{diag}[\gamma_{\text{cp}}(0), \gamma_{\text{cp}}(1), \ldots, \gamma_{\text{cp}}(N-1)] \in \mathbb{C}^{N\times N}$ are the values of the channel transfer function $G(z) \triangleq \sum_{\ell=0}^{L} g(\ell)\, z^{-\ell}$ evaluated at the subcarriers $z_m \triangleq e^{i\frac{2\pi}{N}m}$, i.e., $\gamma_{\text{cp}}(m) = G(z_m)$, $\forall m \in \{0, 1, \ldots, N-1\}$. Henceforth, one obtains that $\boldsymbol{\mathcal{G}}_{\text{cp}} = \mathbf{W}_{\text{IDFT}}\,\mathbf{\Gamma}_{\text{cp}}\,\mathbf{C}$ and, since $\mathbf{W}_{\text{IDFT}}$ is nonsingular, it follows that

$$\text{rank}(\boldsymbol{\mathcal{G}}_{\text{cp}}) = \text{rank}(\mathbf{\Gamma}_{\text{cp}}\,\mathbf{C}). \tag{5.13}$$

The full-column rank property of matrix $\boldsymbol{\mathcal{G}}_{\text{cp}}$ is characterized by the following Theorem [18].

Theorem 5.1 (Rank characterization of $\boldsymbol{\mathcal{G}}_{\text{cp}}$) *If $\mathbf{C}$ is full-column rank and the channel transfer function $G(z)$ has $0 \leq M_z \leq L$ distinct zeros at $z_{m_1} = e^{i\frac{2\pi}{N}m_1}, z_{m_2} = e^{i\frac{2\pi}{N}m_2}, \ldots, z_{m_{M_z}} = e^{i\frac{2\pi}{N}m_{M_z}}$, with $m_1 \neq m_2 \neq \cdots \neq m_{M_z} \in \{0, 1, \ldots, N-1\}$, then the composite channel matrix $\boldsymbol{\mathcal{G}}_{cp}$ is full-column rank iff $[\mathbf{C}, \mathbf{S}_z] \in \mathbb{C}^{N\times(J+M_z)}$ is full-column rank, where $\mathbf{S}_z \triangleq [\mathbf{1}_{m_1}, \mathbf{1}_{m_2}, \ldots, \mathbf{1}_{m_{M_z}}] \in \mathbb{R}^{N\times M_z}$ is a full-column rank matrix, with $\mathbf{1}_m$ denoting the $(m+1)$th column of $\mathbf{I}_N$.*

Proof: See C.1

Some remarks are now in order concerning immediate implications of Theorem 5.1.

Remark 1: $\boldsymbol{\mathcal{G}}_{\text{cp}}$ may be rank deficient even if $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_J$ are linearly independent (see proof in C.1), i.e., $\mathbf{C}$ is full-column rank. However, if $G(z)$ has no zeros (i.e., $M_z = 0$) on the subcarriers $\{z_m\}_{m=0}^{N-1}$, that is, $\gamma_{\text{cp}}(m) \neq 0$, $\forall m \in \{0, 1, \ldots, N-1\}$, it results that $\mathbf{\Gamma}_{\text{cp}}$ is nonsingular and, consequently, $\text{rank}(\boldsymbol{\mathcal{G}}_{\text{cp}}) = \text{rank}(\mathbf{C})$. In other words, for a CP-based system, only if $G(z)$ has no zeros on the used subcarriers, the linear independence of the vectors $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_J$ becomes a necessary *and* sufficient condition for assuring the full-column rank property of $\boldsymbol{\mathcal{G}}_{\text{cp}}$. In this case, both CP-based and ZP-based systems are able to support up to N active users.

Remark 2: Unlike conventional CP-OFDM systems [84], the presence of channel zeros on some subcarriers does not prevent perfect symbol recovery. This result stems from the fact that, in MC-CDMA systems with frequency-domain spreading, each symbol is transmitted in *parallel* on all the subcarriers; therefore, if the $\overline{m}$-th subcarrier is hit by a channel zero, i.e., $\gamma_{\text{cp}}(\overline{m}) = 0$, the

transmitted symbol $b_j(k)$ can still be recovered from the other subcarriers. In contrast, in CP-OFDM systems, wherein each subcarrier conveys a different data symbol, if $G(z)$ exhibits a zero on a used subcarrier, the symbol transmitted on that subcarrier is permanently lost [84, 90].

Remark 3: condition $\text{rank}(\boldsymbol{\mathcal{G}}_{\text{cp}}) = J$ amounts to $\text{rank}([\mathbf{C}, \mathbf{S}_z]) = J + M_z$, which necessarily requires that $J \leq N - M_z$, with $0 < M_z \leq L < L_{\text{p}} \ll N$. Therefore, the number of active users that can be supported through L-ZF-MUD is decremented by one unit for any additional zero on the subcarriers[3]. In this case, the capacity (i.e. the maximum number of users that can be supported) of a CP-based downlink is smaller than that of a ZP-based system, which is equal to N regardless of the channel-zero configuration. In the worst case when $M_z = L$, i.e., all the channel zeros are located at the subcarriers, the maximum number of allowable users in a CP-based downlink is equal to $N - L$.

Theorem 5.1 evidences that, in contrast with ZP-based systems, the full-column rank property of $\boldsymbol{\mathcal{G}}_{\text{cp}}$ depends not only on the linear independence of $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_J$, but also on the presence of channel zeros located at the subcarriers $\{z_m\}_{m=0}^{N-1}$, whose number M_z and locations $m_1, m_2, \ldots, m_{M_z}$ are *unknown* at the receiver. In other words, by imposing the unique constraint that $\mathbf{C}$ be full-column rank, perfect symbol recovery in a CP-based system explicitly depends on the channel impulse response. However, the usefulness of Theorem 5.1 goes beyond this aspect and, most importantly, it allows us to provide universal code designs, assuring that $\boldsymbol{\mathcal{G}}_{\text{cp}}$ is full-column rank for *any* possible configuration of the channel zeros. To this aim, on the basis of Theorem 5.1, observing that $0 \leq M_z \leq L$ and any subset of linearly independent vectors is constituted by linearly independent vectors, we can state the following universal design constraint for the user codes in a CP-based system:

Condition D_{cp} (Universal code design for L-ZF-MUD in CP-MC-CDMA): Under the assumption that $\mathbf{C}$ is full-column rank, no linear combination of the columns of $\mathbf{C}$ can be expressed as linear combinations of the L distinct vectors $\mathbf{1}_{m_1}, \mathbf{1}_{m_2}, \ldots, \mathbf{1}_{m_L}$, for *any* $\{m_1, m_2, \ldots, m_L\} \subset \{0, 1, \ldots, N-1\}$.

[3]It is worth noting that, unlike L-ZF universal multiuser detectors, which do not exist for $J > N - M_z$, the L-MMSE multiuser detector can still be synthesized in the presence of noise even when $J > N - M_z$. However, as previously remarked, its performance is unsatisfactory in this case (see also Section 5.5). Thus, $N - M_z$ also represents the maximum number of users that a CP-based system can reliably manage when L-MMSE-MUD is employed at the receiver.

Equivalently, require that

$$\text{rank}([\mathbf{C}, \mathbf{S}_{\text{univ}}]) = J+L, \;\; \forall\{m_1, m_2, \ldots, m_L\} \subset \{0, \ldots, N-1\}, \quad (5.14)$$

where $\mathbf{S}_{\text{univ}} \stackrel{\triangle}{=} [\mathbf{1}_{m_1}, \mathbf{1}_{m_2}, \ldots, \mathbf{1}_{m_L}] \in \mathbb{R}^{N\times L}$ is a full-column rank matrix.

By virtue of Theorem 5.1, the composite channel matrix $\boldsymbol{\mathcal{G}}_{\text{cp}}$ turns out to be full-column rank for *any* FIR channel of order $L < L_{\text{p}}$ iff the code design constraint D_{cp} is fulfilled. Observe that D_{cp} is stronger than condition $\text{rank}(\mathbf{C}) = J$. In fact, D_{cp} implies that $\text{rank}(\mathbf{C}) = J$, whereas $\text{rank}(\mathbf{C}) = J$ does not imply D_{cp}. It is also apparent from (5.14) that, since $[\mathbf{C}, \mathbf{S}_{\text{univ}}] \in \mathbb{C}^{N\times(J+L)}$, fulfillment of D_{cp} imposes that

$$J \leq N - L, \quad (5.15)$$

i,e, no more than $N - L$ users can be handled by a CP-based system. Furthermore, it is worth noting that common WH spreading codes do not satisfy (5.14). To show this, as a simple counterexample, consider the case of two users (i.e., $J = 2$), which employs the following 8-length WH codes $\mathbf{c}_1 = [1, -1, 1, -1, 1, -1, 1, -1]^T$ and $\mathbf{c}_2 = [1, 1, -1, -1, 1, 1, -1, -1]^T$, obtained by picking the second and third columns of the Hadamard matrix of order $N = 8$. In this case, it is easily verified that $\mathbf{c}_1 + \mathbf{c}_2 = [2, 0, 0, -2, 2, 0, 0, -2]^T$. Hence, if the channel transfer function $G(z)$ has $M_z = 4$ zeros on the subcarriers z_0, z_3, z_4 and z_7, the corresponding matrix $\boldsymbol{\mathcal{G}}_{\text{cp}}$ is not full-column rank, since a particular linear combination of $\mathbf{c}_1$ and $\mathbf{c}_2$ can be expressed as the linear combination of the vectors $\mathbf{1}_0$, $\mathbf{1}_3$, $\mathbf{1}_4$ and $\mathbf{1}_7$:

$$\mathbf{c}_1 + \mathbf{c}_2 = 2\,\mathbf{1}_0 - 2\,\mathbf{1}_3 + 2\,\mathbf{1}_4 - 2\,\mathbf{1}_7. \quad (5.16)$$

Hence, WH spreading codes do not guarantee $\boldsymbol{\mathcal{G}}_{\text{cp}}$ to be full-column rank for any FIR channel of order $L < L_{\text{p}}$.

To design codes that instead fulfill D_{cp}, it is convenient to give an alternative interpretation of (5.14). Since it results [91] that

$$\text{rank}([\mathbf{C}, \mathbf{S}_{\text{univ}}]) = \text{rank}(\mathbf{S}_{\text{univ}}) + \text{rank}[(\mathbf{I}_N - \mathbf{S}_{\text{univ}}\,\mathbf{S}_{\text{univ}}^{-})\,\mathbf{C}], \quad (5.17)$$

with $\text{rank}(\mathbf{S}_{\text{univ}}) = L$ and $\mathbf{S}_{\text{univ}}^{-} = \mathbf{S}_{\text{univ}}^{T}$ [24], it follows that $\text{rank}([\mathbf{C}, \mathbf{S}_{\text{univ}}]) = J + L$ holds iff $\text{rank}[(\mathbf{I}_N - \mathbf{S}_{\text{univ}}\,\mathbf{S}_{\text{univ}}^{T})\,\mathbf{C}] = J$. It can be verified by direct inspection that all the L rows of the matrix $(\mathbf{I}_N - \mathbf{S}_{\text{univ}}\,\mathbf{S}_{\text{univ}}^{T})\,\mathbf{C}$ located in the positions $m_1 + 1, m_2 + 1, \ldots, m_L + 1$ are zero (all the entries are equal

to zero), whereas the $N-L$ remaining ones coincide with the corresponding rows of $\mathbf{C}$. Consequently, the condition $\text{rank}[(\mathbf{I}_N - \mathbf{S}_{\text{univ}}\,\mathbf{S}_{\text{univ}}^T)\,\mathbf{C}] = J$, for any $\{m_1, m_2, \ldots, m_L\} \subset \{0, 1, \ldots, N-1\}$, is equivalent to state that, among any $N-L$ rows of $\mathbf{C}$, a set of $J \leq N-L$ linearly independent rows can be selected. More formally, D_{cp} can be equivalently restated as follows:

Reformulation of Condition D_{cp}: Let vector $\boldsymbol{\omega}_\ell^T \triangleq [c_1^{(\ell)}, c_2^{(\ell)}, \ldots, c_J^{(\ell)}] \in \mathbb{C}^{1\times J}$ denote the $(\ell+1)$th row of $\mathbf{C}$, with $\ell \in \{0, 1, \ldots, N-1\}$; for any $\{m_1, m_2, \ldots, m_L\} \subset \{0, 1, \ldots, N-1\}$, there exists a subset of $J \leq N-L$ distinct indices $\{\ell_1, \ell_2, \ldots, \ell_J\} \subset \{0, 1, \ldots, N-1\} - \{m_1, m_2, \ldots, m_L\}$ such that the vectors $\boldsymbol{\omega}_{\ell_1}, \boldsymbol{\omega}_{\ell_2}, \ldots, \boldsymbol{\omega}_{\ell_J}$ are linearly independent.

It is worthwhile to observe that condition D_{cp} does not uniquely specify $\mathbf{C}$ and, thus, different universal codes can be built. For instance, condition D_{cp} can be accomplished by imposing that each row of $\mathbf{C}$ be a Vandermonde-like vector. Specifically, let us select $N \geq J+L$ nonzero numbers $\{\rho_\ell\}_{\ell=0}^{N-1}$ and build the code vectors $\mathbf{c}_j$ as

$$\mathbf{c}_j = \frac{1}{\sqrt{\chi_j}} \left[\rho_0^j, \rho_1^j, \ldots, \rho_{N-1}^j\right]^T, \quad \forall j \in \{1, 2, \ldots, J\}, \tag{5.18}$$

where the normalization by $1/\sqrt{\chi_j}$ has been introduced to ensure that $\|\mathbf{c}_j\|^2 = 1$ for each user. Relying on the properties of Vandermonde vectors [56], it can be easily verified that, provided that

$$\rho_0 \neq \rho_1 \neq \ldots \neq \rho_{N-1}, \tag{5.19}$$

any J rows of $\mathbf{C}$ are linearly independent, thus satisfying D_{cp}. An advantage of choosing the spreading vectors as in (5.18) is that, in this way, the code matrix $\mathbf{C}$ is uniquely characterized only by the N parameters $\{\rho_\ell\}_{\ell=0}^{N-1}$. For example, such numbers can be chosen equispaced on the unit circle, by setting $\rho_\ell = e^{-i\frac{2\pi}{N}\ell}$, $\forall \ell \in \{0, 1, \ldots, N-1\}$, thus obtaining

$$\mathbf{c}_j = \frac{1}{\sqrt{N}} \left[1, e^{-i\frac{2\pi}{N}j}, \ldots, e^{-i\frac{2\pi}{N}(N-1)j}\right]^T, \tag{5.20}$$

$\forall j \in \{1, 2, \ldots, J\}$. In this case, the spreading vector $\mathbf{c}_j$ turns out to be a Vandermonde (VM) vector (up to the power-controlling constant $1/\sqrt{N}$) and the columns of the resulting code matrix $\mathbf{C}$ coincide with some columns of the N-point DFT matrix $\mathbf{W}_{\text{DFT}}$. Obviously, since the VM code vectors (5.20) are linearly independent by construction, they also guarantee the existence of

L-ZF solutions for any FIR channel of order $L < L_{\rm p}$ in underloaded ZP-based systems.

Remark 4: Since the channel order L is seldom known in practice, one must resort to the upper bound $L < L_{\rm p}$ for synthesizing $\mathbf{C}$, i.e, one should use $L_{\rm p}$ instead of L in condition $\mathrm{D}_{\rm cp}$. So doing, the allowable number of users must obey $J \leq N - L_{\rm p}$, which is a more restrictive limit than $J \leq N - L$. In other words, requiring that the composite channel matrix $\boldsymbol{\mathcal{G}}_{\rm cp}$ be full-column rank for *any* FIR channel of order $L < L_{\rm p}$ poses a stronger limitation on system capacity.

5.4 Perfect symbol recovery for WL-MUD

In this section we extend the previous analysis to WL receivers, following our papers [17, 18]. With reference to the unified model (5.6), when the information symbols are improper, L-MUD does not fully exploit the second-order statistics (SOS) of the received vector $\mathbf{r}(k)$ because it does not take into account the conjugate autocorrelation matrix $\mathbf{R}_{\mathbf{r}\mathbf{r}^*}(k) \stackrel{\triangle}{=} \mathrm{E}[\mathbf{r}(k)\,\mathbf{r}^T(k)] \in \mathbb{C}^{R\times R}$ (see chapter 2 for more details). Invoking assumptions a1 and a2, we can write

$$\mathbf{R}_{\mathbf{r}\mathbf{r}^*}(k) = \varrho_b(k)\,\boldsymbol{\mathcal{G}}\,\boldsymbol{\mathcal{G}}^T. \tag{5.21}$$

As we have underlined in the previous chapters (**??**), the symbols $b_j(k)$ are improper in a large number of digital modulation schemes (BPSK, DBPSK, OQPSK, CPM, MSK, GMSK) and the improper nature of $b_j(k)$ can be seen as the consequence of a linear deterministic dependence existing between $b_j(k)$ and its conjugate version $b_j^*(k)$, i.e., $b_j^*(k) = e^{i\,2\pi\xi k}\,b_j(k)$, for any $k \in \mathbb{Z}$ and for *any* realization of $b_j(k)$. To conveniently exploit the improper nature of the transmitted symbols we must resort to WL-MUD structures (see chapter 2), which are characterized by the input-output relationship (4.19) that here is reported for simplicity

$$w_j(k) = \mathbf{f}_{j,1}^H\,\mathbf{r}(k) + \mathbf{f}_{j,2}^H\,\mathbf{r}^*(k) = \overline{\mathbf{f}}_j^H\,\widetilde{\mathbf{z}}(k), \tag{5.22}$$

where $\overline{\mathbf{f}}_j \stackrel{\triangle}{=} [\mathbf{f}_{j,1}^T, \mathbf{f}_{j,2}^T]^T \in \mathbb{C}^{2R}$ and the *augmented* vector $\widetilde{\mathbf{z}}(k) \stackrel{\triangle}{=} [\mathbf{r}^T(k), \mathbf{r}^H(k)]^T \in \mathbb{C}^{2R}$ is obtained by stacking $\mathbf{r}(k)$ and its complex conjugate version $\mathbf{r}^*(k)$. Moreover, note that, with reference to the above-mentioned improper modulations techniques, the following linear deterministic relationship holds: $\mathbf{b}^*(k) = e^{i\,2\pi\xi k}\,\mathbf{b}(k)$, for any $k \in \mathbb{Z}$ which, substituted in (5.6),

yields $\mathbf{r}^*(k) = e^{i\,2\pi\xi k}\,\boldsymbol{\mathcal{G}}^*\,\mathbf{b}(k) + \mathbf{v}^*(k)$. The latter relation shows that the (possible) conjugate cyclostationarity of $b_j(k)$ can be deterministically compensated for by performing a *derotation* [92] of $\mathbf{r}^*(k)$ before evaluating $\widetilde{\mathbf{z}}(k)$, that is, by considering the modified input-output relationship

$$w_j(k) = \mathbf{f}_{j,1}^H\,\mathbf{r}(k) + \mathbf{f}_{j,2}^H\,\mathbf{r}^*(k)\,e^{-i\,2\pi\xi k} = \overline{\mathbf{f}}_j^H\,\mathbf{z}(k)\,, \tag{5.23}$$

where the augmented and derotated vector $\mathbf{z}(k) \in \mathbb{C}^{2R}$ is given by

$$\begin{aligned}\mathbf{z}(k) \stackrel{\triangle}{=} \begin{bmatrix} \mathbf{r}(k) \\ \mathbf{r}^*(k)\,e^{-i\,2\pi\xi k} \end{bmatrix} &= \underbrace{\begin{bmatrix} \boldsymbol{\mathcal{G}} \\ \boldsymbol{\mathcal{G}}^* \end{bmatrix}}_{\boldsymbol{\mathcal{H}}\in\mathbb{C}^{2R\times J}}\mathbf{b}(k) + \underbrace{\begin{bmatrix} \mathbf{v}(k) \\ \mathbf{v}^*(k)\,e^{-i\,2\pi\xi k} \end{bmatrix}}_{\mathbf{w}(k)\in\mathbb{C}^{2R}} \\ &= \boldsymbol{\mathcal{H}}\,\mathbf{b}(k) + \mathbf{w}(k)\,. \end{aligned} \tag{5.24}$$

Following the same lines that we have indicated in the previous chapters, in the absence of noise, the ZF condition $w_j(k) = b_j(k)$ leads to the system of linear equations $\boldsymbol{\mathcal{H}}^H\,\overline{\mathbf{f}}_j = \mathbf{e}_j$, which is consistent for each user iff the augmented channel matrix $\boldsymbol{\mathcal{H}}$ is full-column rank, i.e., $\mathrm{rank}(\boldsymbol{\mathcal{H}}) = J$; under this assumption, the *minimal norm* [24] solution of $\boldsymbol{\mathcal{H}}^H\overline{\mathbf{f}}_j = \mathbf{e}_j$ is

$$\overline{\mathbf{f}}_{\text{WL-ZF},j} = \boldsymbol{\mathcal{H}}\,(\boldsymbol{\mathcal{H}}^H\boldsymbol{\mathcal{H}})^{-1}\mathbf{e}_j\,, \tag{5.25}$$

which defines the WL-ZF or WL decorrelating multiuser detector. In the presence of noise, one can more suitably resort to the WL-MMSE multiuser detector [3, 50, 54], which is defined as

$$\overline{\mathbf{f}}_{\text{WL-MMSE},j} = \arg\min_{\overline{\mathbf{f}}_j\in\mathbb{C}^R} \mathrm{E}[|b_j(k) - w_j(k)|^2] = \sigma_b^2\,\mathbf{R}_{\mathbf{zz}}^{-1}\,\boldsymbol{\mathcal{H}}\,\mathbf{e}_j\,, \tag{5.26}$$

where $\mathbf{R}_{\mathbf{zz}} \stackrel{\triangle}{=} \mathrm{E}[\mathbf{z}(k)\,\mathbf{z}^H(k)] \in \mathbb{C}^{2R\times 2R}$ is the autocorrelation matrix of $\mathbf{z}(k)$ which, accounting for (5.24), and invoking assumptions a1 and a2, is given by

$$\mathbf{R}_{\mathbf{zz}} = \sigma_b^2\,\boldsymbol{\mathcal{H}}\,\boldsymbol{\mathcal{H}}^H + \sigma_v^2\,\mathbf{I}_{2R}\,. \tag{5.27}$$

Reasoning as in precedence for the L-MMSE multiuser detector, it is readily seen that, if $\boldsymbol{\mathcal{H}}$ is full-column rank[4], the WL-MMSE multiuser detector ends up to the WL-ZF one in the limit $\sigma_v^2/\sigma_b^2 \to 0$.

[4] More generally, when $\boldsymbol{\mathcal{H}}$ is possibly rank-deficient, then $\lim_{\sigma_v^2/\sigma_b^2\to 0}\overline{\mathbf{f}}_{\text{WL-MMSE},j} = (\boldsymbol{\mathcal{H}}^H)^\dagger\mathbf{e}_j \stackrel{\triangle}{=} \overline{\mathbf{f}}_{\text{WL-LS},j}$, i.e., the WL-MMSE multiuser detector ends up to the minimal-norm least-squares solution [24] of $\boldsymbol{\mathcal{H}}^H\overline{\mathbf{f}}_j = \mathbf{e}_j$ [note that, when $\boldsymbol{\mathcal{H}}$ is full-column rank, one has $\overline{\mathbf{f}}_{\text{WL-LS},j} = \overline{\mathbf{f}}_{\text{WL-ZF},j}$ from (5.25)].

Henceforth, similarly to the condition rank$(\boldsymbol{\mathcal{G}}) = J$ for L-MUD, the full-column rank property of $\boldsymbol{\mathcal{H}}$ not only assures the existence of WL-ZF solutions, but also allows the WL-MMSE multiuser detector to satisfactorily work in the high SNR region. Such a condition, i.e., rank$(\boldsymbol{\mathcal{H}}) = J$, is esplored in the chapter 4 with reference to a DS-CDMA system. The full column rank property of $\boldsymbol{\mathcal{H}}$ is thoroughly studied in Subsection 5.4.1, with reference to both CP- and ZP-based systems. In particular, by taking advantage of the results derived in Section 5.3, we will show that, if the user codes are judiciously designed, the condition rank$(\boldsymbol{\mathcal{H}}) = J$ can also be guaranteed when the number of users exceeds the number of subcarriers, regardless of the underlying frequency-selective channel.

5.4.1 Rank characterization of $\boldsymbol{\mathcal{H}}$ and universal code design for WL-ZF-MUD

From a unified perspective, observe that rank$(\boldsymbol{\mathcal{H}}) = J$ iff the null spaces of the matrices $\boldsymbol{\mathcal{G}}$ and $\boldsymbol{\mathcal{G}}^*$ intersect only trivially, that is, $\mathcal{N}(\boldsymbol{\mathcal{G}}) \cap \mathcal{N}(\boldsymbol{\mathcal{G}}^*) = \{\mathbf{0}_J\}$. It can be easily verified that, if $\boldsymbol{\mathcal{G}}$ is full-column rank, which necessarily requires that $J \leq N$ (underloaded systems), then this condition is trivially satisfied and, hence, the augmented matrix $\boldsymbol{\mathcal{H}}$ is full-column rank as well. Remarkably, the converse statement is not true, that is, $\boldsymbol{\mathcal{H}}$ may be full-column rank even in overloaded MC-CDMA systems, i.e., when the number J of users is larger then the number N of subcarriers and, thus, $\boldsymbol{\mathcal{G}}$ is inherently rank-deficient. In the latter case, the code vectors $\{\mathbf{c}_j\}_{j=1}^{J}$ *cannot* be linearly independent, thus giving rank$(\boldsymbol{\mathcal{G}}) \leq N$, which in its turn implies that the dimension of the subspace $\mathcal{N}(\boldsymbol{\mathcal{G}})$ is nonnull and is equal to $J - \text{rank}(\boldsymbol{\mathcal{G}})$. Now, we analyze this aspect separating ZP- and CP- cases.

ZP-based downlink

Let us consider a ZP-based system [see (5.5)], wherein $\boldsymbol{\mathcal{G}} = \boldsymbol{\mathcal{G}}_{\text{zp}} = \boldsymbol{\Theta}_{\text{zp}}\, \mathbf{W}_{\text{IDFT}}\, \mathbf{C}$. In this case, the augmented channel matrix $\boldsymbol{\mathcal{H}}$ assumes the form

$$\boldsymbol{\mathcal{H}} = \boldsymbol{\mathcal{H}}_{\text{zp}} = \begin{bmatrix} \boldsymbol{\mathcal{G}}_{\text{zp}} \\ \boldsymbol{\mathcal{G}}_{\text{zp}}^* \end{bmatrix} = \underbrace{\begin{bmatrix} \boldsymbol{\Theta}_{\text{zp}}\, \mathbf{W}_{\text{IDFT}} & \mathbf{O}_{P\times N} \\ \mathbf{O}_{P\times N} & \boldsymbol{\Theta}_{\text{zp}}^*\, \mathbf{W}_{\text{IDFT}}^* \end{bmatrix}}_{\overline{\boldsymbol{\Xi}}_{\text{zp}} \in \mathbb{C}^{2P\times 2N}} \underbrace{\begin{bmatrix} \mathbf{C} \\ \mathbf{C}^* \end{bmatrix}}_{\overline{\mathbf{C}} \in \mathbb{C}^{2N\times J}} = \overline{\boldsymbol{\Xi}}_{\text{zp}}\, \overline{\mathbf{C}}. \tag{5.28}$$

It can be shown [56] that

$$\text{rank}(\overline{\boldsymbol{\Xi}}_{\text{zp}}) = \text{rank}(\boldsymbol{\Theta}_{\text{zp}}\,\mathbf{W}_{\text{IDFT}}) + \text{rank}(\boldsymbol{\Theta}^{*}_{\text{zp}}\,\mathbf{W}^{*}_{\text{IDFT}}) = 2\,N \tag{5.29}$$

and, therefore

$$\text{rank}(\boldsymbol{\mathcal{H}}_{\text{zp}}) = \text{rank}(\overline{\mathbf{C}}). \tag{5.30}$$

In other words, let $\overline{\mathbf{c}}_j \stackrel{\triangle}{=} [\mathbf{c}_j^T, \mathbf{c}_j^H]^T \in \mathbb{C}^{2N}$ define the *augmented* code vector of the jth user, for $j \in \{1, 2, \ldots, J\}$, the matrix $\boldsymbol{\mathcal{H}}_{\text{zp}}$ is full-column rank iff the code vectors $\overline{\mathbf{c}}_1, \overline{\mathbf{c}}_2, \ldots, \overline{\mathbf{c}}_J$ are linearly independent. In other words, a necessary and sufficient condition guaranteeing the existence of WL-ZF solutions for ZP-based system is that the augmented code matrix $\overline{\mathbf{C}}$ is full-column rank. It is worthwhile to observe that the augmented code vectors $\{\overline{\mathbf{c}}_j\}_{j=1}^{J}$ can be linearly independent even if the code vectors $\mathbf{c}_j$ are linearly dependent, which surely happens when $J > N$. In this regard, we provide the following Lemma.

Lemma 5.1 (Rank characterization of $\overline{\mathbf{C}}$) *If $J \leq 2\,N$, then the augmented frequency-domain code matrix $\overline{\mathbf{C}}$ is full-column rank iff there are no conjugate pairs of nonzero vectors belonging to $\mathcal{N}(\mathbf{C})$.*

Proof: See C.2.

In underloaded scenarios, wherein the code vectors $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_J$ can be linearly independent, it follows that $\mathcal{N}(\mathbf{C}) = \{\mathbf{0}_J\}$ and, thus, the augmented matrix $\overline{\mathbf{C}}$ turns out to be full-column rank, too. Therefore, from now on, we focus attention on the more interesting overloaded environments, wherein $N < J \leq 2\,N$. In this case, $\mathbf{C}$ is a wide matrix and, assuming without loss of generality that its first N column $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_N$ are linearly independent, its remaining $J - N$ columns $\mathbf{c}_{N+1}, \mathbf{c}_{N+2}, \ldots, \mathbf{c}_J$ can be expressed as a linear combination of the first N ones, thus obtaining the following decomposition

$$\mathbf{C} = [\mathbf{C}_{\text{left}} \;\; \mathbf{C}_{\text{left}}\,\boldsymbol{\Pi}] = \mathbf{C}_{\text{left}}\,[\mathbf{I}_N \;\; \boldsymbol{\Pi}]\,, \tag{5.31}$$

where $\mathbf{C}_{\text{left}} \stackrel{\triangle}{=} [\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_N] \in \mathbb{C}^{N\times N}$ is nonsingular and $\boldsymbol{\Pi} \in \mathbb{C}^{N\times(J-N)}$ is a tall matrix. Due to nonsingularity of $\mathbf{C}_{\text{left}}$, it follows that $\mathcal{N}(\mathbf{C}) = \mathcal{N}([\mathbf{I}_N \;\; \boldsymbol{\Pi}])$. Furthermore, it can be verified that the general form of two vectors $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2 \in \mathbb{C}^J$ belonging to $\mathcal{N}([\mathbf{I}_N \;\; \boldsymbol{\Pi}])$ and, thus, to $\mathcal{N}(\mathbf{C})$, is given by

$$\boldsymbol{\alpha}_1 = \begin{bmatrix} -\boldsymbol{\Pi} \\ \mathbf{I}_{J-N} \end{bmatrix} \boldsymbol{\vartheta}_1 \quad \text{and} \quad \boldsymbol{\alpha}_2 = \begin{bmatrix} -\boldsymbol{\Pi} \\ \mathbf{I}_{J-N} \end{bmatrix} \boldsymbol{\vartheta}_2\,, \tag{5.32}$$

with arbitrary $\boldsymbol{\vartheta}_1, \boldsymbol{\vartheta}_2 \in \mathbb{C}^{J-N}$. By virtue of Lemma 5.1, the augmented code matrix $\overline{\mathbf{C}}$ is not full-column rank iff there exist at least two nonzero vectors $\boldsymbol{\vartheta}_1$ and $\boldsymbol{\vartheta}_2$ such that $\boldsymbol{\alpha}_1 = \boldsymbol{\alpha}_2^*$, which amounts to $\boldsymbol{\vartheta}_1 = \boldsymbol{\vartheta}_2^*$ and $(\boldsymbol{\Pi} - \boldsymbol{\Pi}^*)\,\boldsymbol{\vartheta}_1 = \mathbf{0}_N$. In its turn this second equation can be equivalently written as $\mathrm{Im}\{\boldsymbol{\Pi}\}\,\boldsymbol{\vartheta}_1 = \mathbf{0}_N$. Therefore, if the imaginary part of $\boldsymbol{\Pi}$ is full-column rank, then $\boldsymbol{\alpha}_1 = \boldsymbol{\alpha}_2^*$ is satisfied iff $\boldsymbol{\alpha}_1 = \boldsymbol{\alpha}_2 = \mathbf{0}_J$ which, accounting for Lemma 5.1, assures that $\mathrm{rank}(\overline{\mathbf{C}}) = J$. Summarizing this result, we can state the following universal code design strategy for a ZP-based overloaded system:

Condition $\overline{\mathrm{D}}_{zp}$ (Universal code design for WL-ZF in ZP-MC-CDMA): Let $N < J \leq 2\,N$ and $\mathbf{C}_{\text{left}} \stackrel{\triangle}{=} [\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_N] \in \mathbb{C}^{N\times N}$ be nonsingular, the code matrix has the form $\mathbf{C} = \mathbf{C}_{\text{left}}\,[\mathbf{I}_N \;\; \boldsymbol{\Pi}]$, where $\boldsymbol{\Pi} \in \mathbb{C}^{N\times(J-N)}$ is a tall matrix, whose imaginary part $\mathrm{Im}\{\boldsymbol{\Pi}\}$ is full-column rank.

Some interesting remarks regarding fulfillment of condition $\overline{\mathrm{D}}_{\text{zp}}$ can be drawn at this point.

Remark 5: To begin with, observe that the code design $\overline{\mathrm{D}}_{\text{zp}}$, which represents a necessary and sufficient condition in order to guarantee $\mathrm{rank}(\overline{\mathbf{C}}) = J$, is universal, in the sense that it allows $\boldsymbol{\mathcal{H}}_{\text{zp}}$ to be full-column rank for *any* FIR channel of order $L < L_{\text{p}}$. If this universal code constraint is fulfilled, then channel-irrespective WL-ZF-MUD is guaranteed up to $2\,N$ users, which is exactly the double of the number of users that can be managed in a ZP-based system employing L-ZF-MUD.

Remark 6: If the spreading codes are real-valued, i.e., $\mathbf{C}^* = \mathbf{C}$, the matrix $\boldsymbol{\Pi}$ is real-valued as well, i.e., $\mathrm{Im}\{\boldsymbol{\Pi}\} = \mathbf{O}_{N\times(J-N)}$ and, consequently, condition $\overline{\mathrm{D}}_{\text{zp}}$ is not satisfied. Thus, employing real-valued code vectors (e.g., WH spreading) implies necessarily that, similarly to L-ZF-MUD, the existence of WL-ZF solutions can be guaranteed only in underloaded MC-CDMA systems. On the other hand, if complex-valued code vectors are employed, then $\overline{\mathbf{C}}$ can be full-column rank even in overloaded systems, where $\mathbf{C}$ is not full-column rank.

Remark 7: Although they are complex-valued, the VM code vectors given by (5.20) do not satisfy $\overline{\mathrm{D}}_{\text{zp}}$ when $N < J \leq 2\,N$: indeed, it is easily shown that, in this case, the following decomposition holds $\mathbf{C} = \mathbf{W}_{\text{DFT}}\,[\mathbf{I}_N \;\; \mathbf{J}]$, where $\mathbf{J} \stackrel{\triangle}{=} [\mathbf{1}_1, \mathbf{1}_2, \ldots, \mathbf{1}_{J-N}] \in \mathbb{R}^{N\times(J-N)}$ is real-valued and, in this case, $\mathrm{Im}\{\boldsymbol{\Pi}\} = \mathrm{Im}\{\mathbf{J}\} = \mathbf{O}_{N\times(J-N)}$ is rank-deficient. Hence, the VM codes (5.20) do not ensure channel-independent WL-ZF-MUD in an overloaded ZP-

based downlink.

Besides allowing one to readily check whether a given set of spreading sequences assures the existence of WL-ZF solutions for any FIR channel of order $L < L_{\rm p}$, condition $\overline{\rm D}_{\rm zp}$ provides a direct procedure to build universal codes for ZP-based overloaded systems. Among several options that can be pursued, we devise here a simple universal code design relying on WH spreading. Specifically, let $\boldsymbol{\mathcal{W}}_N \in \mathbb{R}^{N\times N}$ denote the common Hadamard matrix of order N: in underloaded scenarios, i.e., when $J \leq N$, one can choose the spreading vectors $\{\mathbf{c}_j\}_{j=1}^{J}$ as the columns of $\frac{1}{\sqrt{N}}\boldsymbol{\mathcal{W}}_N$ (the normalization by $1/\sqrt{N}$ assures that $\|\mathbf{c}_j\|^2 = 1$ for each user); on the other hand, in an overloaded downlink, wherein $N < J \leq 2\,N$, the code matrix $\mathbf{C}$ can be chosen as follows

$$\mathbf{C} = \frac{1}{\sqrt{N}}\Big(\boldsymbol{\mathcal{W}}_N\,[\mathbf{I}_N \;\; i\,\mathbf{J}]\Big)\,, \tag{5.33}$$

which, as it is immediately seen, satisfies condition $\overline{\rm D}_{\rm zp}$. In this way, the spreading vectors of the first N users have elements confined to the two values $\{\pm 1/\sqrt{N}\}$, whereas the entries of the code vectors of the remaining users take on the two values $\{\pm i/\sqrt{N}\}$. In conclusion, we can state that the adoption of the code matrix (5.33), which comes from a simple modification of the conventional WH spreading technique, guarantees WL-ZF-MUD in both underloaded and overloaded ZP-based downlink, for any FIR channel of order $L < L_{\rm p}$.

CP-based downlink

Let us consider a CP-based system [see (5.4)], wherein $\boldsymbol{\mathcal{G}} = \boldsymbol{\mathcal{G}}_{\rm cp}$ can be equivalently expressed as $\boldsymbol{\mathcal{G}}_{\rm cp} = \mathbf{W}_{\rm IDFT}\,\boldsymbol{\Gamma}_{\rm cp}\,\mathbf{C}$. In this case, one has

$$\begin{aligned}\boldsymbol{\mathcal{H}} = \boldsymbol{\mathcal{H}}_{\rm cp} = \begin{bmatrix}\boldsymbol{\mathcal{G}}_{\rm cp}\\ \boldsymbol{\mathcal{G}}_{\rm cp}^{*}\end{bmatrix} &= \underbrace{\begin{bmatrix}\mathbf{W}_{\rm IDFT} & \mathbf{O}_{N\times N}\\ \mathbf{O}_{N\times N} & \mathbf{W}_{\rm IDFT}^{*}\end{bmatrix}}_{\overline{\mathbf{W}}_{\rm IDFT}\in\mathbb{C}^{2N\times 2N}} \cdot \underbrace{\begin{bmatrix}\boldsymbol{\Gamma}_{\rm cp} & \mathbf{O}_{N\times N}\\ \mathbf{O}_{N\times N} & \boldsymbol{\Gamma}_{\rm cp}^{*}\end{bmatrix}}_{\overline{\boldsymbol{\Gamma}}_{\rm cp}\in\mathbb{C}^{2N\times 2N}} \underbrace{\begin{bmatrix}\mathbf{C}\\ \mathbf{C}^{*}\end{bmatrix}}_{\overline{\mathbf{C}}\in\mathbb{C}^{2N\times J}}\\ &= \overline{\mathbf{W}}_{\rm IDFT}\,\overline{\boldsymbol{\Gamma}}_{\rm cp}\,\overline{\mathbf{C}}\,.\end{aligned} \tag{5.34}$$

Since $\mathrm{rank}(\overline{\mathbf{W}}_{\rm IDFT}) = \mathrm{rank}(\mathbf{W}_{\rm IDFT}) + \mathrm{rank}(\mathbf{W}_{\rm IDFT}^{*}) = 2\,N$, it results that $\mathrm{rank}(\boldsymbol{\mathcal{H}}_{\rm cp}) = \mathrm{rank}(\overline{\boldsymbol{\Gamma}}_{\rm cp}\,\overline{\mathbf{C}})$ and, hence, we can directly investigate the rank properties of $\overline{\boldsymbol{\Gamma}}_{\rm cp}\,\overline{\mathbf{C}}$. As a first remark, observe that, in order for $\overline{\boldsymbol{\Gamma}}_{\rm cp}\,\overline{\mathbf{C}}$ to be full-column rank, the matrix $\overline{\mathbf{C}}$ must necessarily be full-column rank, i.e., $J \leq 2\,N$ and $\mathrm{rank}(\overline{\mathbf{C}}) = J$. Therefore, differently from the ZP case, linear independence of the augmented code vector $\overline{\mathbf{c}}_1, \overline{\mathbf{c}}_2, \ldots, \overline{\mathbf{c}}_J$ is a necessary

but not sufficient condition in order to have rank($\boldsymbol{\mathcal{H}}_{\text{cp}}$) $= J$. Consequently, to allow $\boldsymbol{\mathcal{H}}_{\text{cp}}$ to be full-column rank even in overloaded scenarios, as a first constraint on the user codes, we have to impose that the matrix $\mathbf{C}$ be synthesized according to $\overline{\mathrm{D}}_{\text{zp}}$, which represents a necessary and sufficient condition in order to have rank($\overline{\mathbf{C}}$) $= J$, when $N < J \leq 2\,N$. This implies that any spreading technique, which enables channel-irrespective WL perfect symbol recovery for a CP-based downlink, can also be employed for the same purpose in a ZP-based system. The full-column rank property of the matrix $\boldsymbol{\mathcal{H}}_{\text{cp}}$ is characterized by the following Theorem.

Theorem 5.2 (Rank characterization of $\boldsymbol{\mathcal{H}}_{\text{cp}}$) *If $\overline{\mathbf{C}}$ is full-column rank and the channel transfer function $G(z)$ has $0 \leq M_z \leq L$ distinct zeros on the subcarriers $z_{m_1} = e^{i\frac{2\pi}{N}m_1}, z_{m_2} = e^{i\frac{2\pi}{N}m_2}, \ldots, z_{m_{M_z}} = e^{i\frac{2\pi}{N}m_{M_z}}$, with $m_1 \neq m_2 \neq \cdots \neq m_{M_z} \in \{0, 1, \ldots, N-1\}$, then the augmented channel matrix $\boldsymbol{\mathcal{H}}_{cp}$ is full-column rank iff $[\overline{\mathbf{C}}, \overline{\mathbf{S}}_z] \in \mathbb{C}^{2N\times(J+2M_z)}$ is full-column rank, where $\overline{\mathbf{S}}_z \stackrel{\triangle}{=} diag[\mathbf{S}_z, \mathbf{S}_z] \in \mathbb{R}^{2N\times 2M_z}$ is full-column rank and $\mathbf{S}_z \in \mathbb{R}^{N\times M_z}$ has been previously defined in Theorem 5.1.*

Proof: The proof is similar in spirit with that of Theorem 5.1 and, thus, is omitted.

Theorem 5.2 suggests the following two additional remarks:

Remark 8: As a first consequence, if the channel transfer function $G(z)$ has no zeros on the subcarriers $\{z_m\}_{m=0}^{N-1}$, i.e., $M_z = 0$, then, similarly to a ZP-based system, the linear independence of the augmented code vectors $\overline{\mathbf{c}}_1, \overline{\mathbf{c}}_2, \ldots, \overline{\mathbf{c}}_J$ becomes a necessary *and* sufficient condition for the existence of WL-ZF solutions in a CP-based downlink. In this case, a CP-based downlink can support up to $2N$ active users, which is equal to the system capacity of a ZP-based downlink employing WL-ZF-MUD. Instead, in the presence of channel zeros on some subcarriers, $\boldsymbol{\mathcal{H}}_{\text{cp}}$ can still be full-column rank. However, in this case, provided that $\overline{\mathbf{C}}$ is full-column rank, the existence of WL-ZF solutions explicitly depends on the channel-zero configuration.

Remark 9: Most importantly, unlike the condition rank($[\mathbf{C}, \mathbf{S}_z]$) $= J + M_z$ of Theorem 5.1 (see also Remark 3), the condition rank($[\overline{\mathbf{C}}, \overline{\mathbf{S}}_z]$) $= J + 2\,M_z$ can be satisfied even when the number of users is larger than the number of subcarriers. Specifically, rank($[\overline{\mathbf{C}}, \overline{\mathbf{S}}_z]$) $= J + 2\,M_z$ necessarily requires that $2N \geq J + 2\,M_z$, that is, the number J of active users must not be larger than $2\,(N - M_z)$, with $0 < M_z \leq L < L_{\text{p}} \ll N$. Hence, similarly to a ZP-based system, WL-ZF-MUD allows a CP-based downlink to support a

number of users that is exactly the double of the number of users that can be accommodated when L-ZF-MUD is employed. However, in the latter case, the allowable number of users is decremented by *two* units for any additional zero on the subcarriers and is smaller than $2N$, which represents the system capacity of a ZP-based system employing WL-ZF-MUD. In the worst case, when all the channel zeros are located at the subcarriers, the maximum number of allowable users in a CP-based downlink is $2\,(N-L)$.

Similarly to Theorem 5.1, the most important implication of Theorem 5.2 regards the fact that it enlightens how to single out universal code designs, which assure that $\boldsymbol{\mathcal{H}}_{\text{cp}}$ be full-column rank for any possible configuration of the channel zeros. With this goal in mind, paralleling the arguments that led to condition D_{cp} in Subsection 5.3.1, the following code design represents a necessary and sufficient condition ensuring that $\boldsymbol{\mathcal{H}}_{\text{cp}}$ is full-column rank for any possible configuration of the channel zeros:

Condition $\overline{\text{D}}_{\text{cp}}$ (Universal code design for WL-ZF in CP-MC-CDMA): Define the full-column rank matrix $\overline{\mathbf{S}}_{\text{univ}} \stackrel{\triangle}{=} \text{diag}[\mathbf{S}_{\text{univ}}, \mathbf{S}_{\text{univ}}] \in \mathbb{R}^{2N\times 2L}$, where $\mathbf{S}_{\text{univ}} \in \mathbb{R}^{N\times L}$ has been previously defined in condition D_{cp}, then, $\forall\{m_1, m_2, \ldots, m_L\} \subset \{0, 1, \ldots, N-1\}$,

$$\begin{aligned} &\text{rank}([\overline{\mathbf{C}}, \overline{\mathbf{S}}_{\text{univ}}]) = J + 2\,L \quad \text{or, equivalently,} \\ &\text{rank}[(\mathbf{I}_{2N} - \overline{\mathbf{S}}_{\text{univ}}\,\overline{\mathbf{S}}_{\text{univ}}^{T})\,\overline{\mathbf{C}}] = J\,. \end{aligned} \tag{5.35}$$

The price to pay for imposing that the matrix $[\overline{\mathbf{C}}, \overline{\mathbf{S}}_{\text{univ}}] \in \mathbb{C}^{2N\times(J+2L)}$ be full-column rank is a reduction of the system capacity (see Remark 9) because the universal code design $\overline{\text{D}}_{\text{cp}}$ can be devised for a maximum number of $2\,(N-L)$ users. It should be observed that $\overline{\text{D}}_{\text{cp}}$ is stronger than condition $\overline{\text{D}}_{\text{zp}}$: indeed, $\overline{\text{D}}_{\text{cp}}$ necessarily requires that $\text{rank}(\overline{\mathbf{C}}) = J$; on the other hand, $\text{rank}(\overline{\mathbf{C}}) = J$ is not sufficient to assure fulfillment of $\overline{\text{D}}_{\text{cp}}$. On the other hand, it is noteworthy that, if condition D_{cp} is satisfied, which is possible as long as $J \leq N-L$, then $\overline{\text{D}}_{\text{cp}}$ is surely fulfilled, too. Therefore, by imposing the unique constraint that the N parameters $\{\rho_\ell\}_{\ell=0}^{N-1}$ be distinct, the code vectors (5.18) guarantee, up to $N-L$ users, the existence of universal WL-ZF solutions. However, in its present form, condition $\overline{\text{D}}_{\text{cp}}$ does not help us give a direct procedure for synthesizing universal spreading codes when $N-L < J \leq 2\,(N-L)$. Nevertheless, taking into account that the matrix $(\mathbf{I}_{2N} - \overline{\mathbf{S}}_{\text{univ}}\,\overline{\mathbf{S}}_{\text{univ}}^{T})\,\overline{\mathbf{C}} \in \mathbb{C}^{2N\times J}$ is obtained from $\overline{\mathbf{C}}$ by setting to zero all the entries of its $2\,L$ rows located in the positions

$m_1+1, m_2+1, \ldots, m_L+1, m_1+N+1, m_2+N+1, \ldots, m_L+N+1$, with reference to the specific case wherein $N-L < J \leq 2\,(N-L)$, we can reformulate condition $\overline{\mathrm{D}}_{\mathrm{cp}}$ in this way:

Reformulated Condition $\overline{D}_{cp}$ when $N-L < J \leq 2\,(N-L)$: Let $\boldsymbol{\omega}_\ell^T \triangleq [c_1^{(\ell)}, c_2^{(\ell)}, \ldots, c_J^{(\ell)}] \in \mathbb{C}^{1\times J}$ denote the $(\ell+1)$th row of $\mathbf{C}$, with $\ell \in \{0,1,\ldots,N-1\}$; when $N-L < J \leq 2\,(N-L)$, for any subset of distinct indices $\{\ell_1, \ell_2, \ldots, \ell_{N-L}\} \subset \{0,1,\ldots,N-1\}$, there exists J linearly independent vectors from the total set $\boldsymbol{\omega}_{\ell_1}, \boldsymbol{\omega}_{\ell_2}, \ldots, \boldsymbol{\omega}_{\ell_{N-L}}, \boldsymbol{\omega}_{\ell_1}^*, \boldsymbol{\omega}_{\ell_2}^*, \ldots, \boldsymbol{\omega}_{\ell_{N-L}}^*$.

Reformulation of condition $\overline{\mathrm{D}}_{\mathrm{cp}}$ allows one to readily check out that the code vectors (5.18) can still fulfill $\overline{\mathrm{D}}_{\mathrm{cp}}$ when $N-L < J \leq 2\,(N-L)$, provided that, in addition to $\rho_0 \neq \rho_1 \neq \cdots \neq \rho_{N-1}$, further constraints on the parameters $\{\rho_\ell\}_{\ell=0}^{N-1}$ are imposed. More precisely, relying on the properties of Vandermonde vectors [56], it is not difficult to prove that condition $\overline{\mathrm{D}}_{\mathrm{cp}}$ is surely satisfied if, besides requiring that the parameters $\{\rho_\ell\}_{\ell=0}^{N-1}$ be distinct, one additionally imposes that

$$\rho_{\ell_1} \neq \rho_{\ell_2}^*, \quad \forall \ell_1, \ell_2 \in \{0,1,\ldots,N-1\}, \tag{5.36}$$

which means that the number $\{\rho_\ell\}_{\ell=0}^{N-1}$ must be complex-valued and cannot be pairwise conjugate. Additionally, it can be immediately inferred that the VM codes (5.20) cannot satisfy the code design $\overline{\mathrm{D}}_{\mathrm{cp}}$ since, in this case, it turns out that $\rho_\ell = \rho_{N-\ell}^*$, $\forall \ell \in \{0,1,\ldots,N-1\}$. Furthermore, it can be verified by direct inspection that the code matrix given by (5.33) does not satisfy condition $\overline{\mathrm{D}}_{\mathrm{cp}}$ (see also Section 5.5) and, thus, contrary to the ZP case, such a spreading technique does not guarantee the existence of universal WL-ZF solutions in CP-based systems.

To develop a family of codes fulfilling condition $\overline{\mathrm{D}}_{\mathrm{cp}}$, we restrict our attention to the spreading vectors (5.18) and, in particular, we start from the N-point DFT codes (5.20), whereby $\rho_\ell = e^{-i\frac{2\pi}{N}\ell}$, $\forall \ell \in \{0,1,\ldots,N-1\}$. To obtain a set of N complex-valued parameters $\{\rho_\ell\}_{\ell=0}^{N-1}$ equispaced on the unit circle, which are not pairwise conjugate, it is sufficient to introduce a suitable rotation by setting $\rho_\ell = e^{-i\left(\frac{2\pi}{N}\ell-\theta\right)}$, $\forall \ell \in \{0,1,\ldots,N-1\}$ and $\theta \in (0, 2\,\pi)$, thus getting the code vectors

$$\mathbf{c}_j = \frac{1}{\sqrt{N}}\left[e^{-i(-\theta)\,j}, e^{-i\left(\frac{2\pi}{N}-\theta\right)j}, \ldots, e^{-i\left(\frac{2\pi}{N}(N-1)-\theta\right)j}\right]^T, \tag{5.37}$$

$\forall j \in \{1, 2, \ldots, J\}$, where, in order to fulfill the constraint $\rho_{\ell_1} \neq \rho^*_{\ell_2}$, $\forall \ell_1, \ell_2 \in \{0, 1, \ldots, N-1\}$, the angle rotation θ must obey the following condition:

$$\theta \neq \frac{\pi}{N}(\ell_1 + \ell_2) + h\,\pi, \quad \forall \ell_1, \ell_2 \in \{0, 1, \ldots, N-1\} \text{ and } \forall h \in \mathbb{Z}. \quad (5.38)$$

Note that the spreading vectors (5.37) differ from those in (5.20) only for the multiplicative scalar $e^{-i(-\theta)j}$. The code vectors (5.37) satisfy the condition $\overline{\mathrm{D}}_{\mathrm{cp}}$ and, hence, they ensure universal WL perfect symbol recovery not only when $J \leq N - L$, but also when $N - L < J \leq 2(N - L)$, in both CP- and ZP-based systems. Finally, observe that, when L is replaced with L_{p}, universal WL-ZF-MUD is still possible in a CP-based system, with the difference that perfect symbol recovery can be guaranteed to at most $2(N - L_{\mathrm{p}})$ users, whose number, although does not depend on the channel order, is however smaller than $2(N - L)$.

5.5 Numerical performance analysis

To corroborate our theoretical analysis, we resort to Monte Carlo computer simulations in this section. Specifically, we consider that, without loss of generality, the desired user is the first one ($j = 1$) and, moreover, we assume that $\mathbf{g}$ is exactly known at the receiver.

In all the experiments, the following simulation setting is adopted. The CP- and ZP-based MC-CDMA systems employ $N = 16$ subcarriers, with $L_{\mathrm{p}} = 4$ and OQPSK improper symbol modulation. Both systems use four different frequency-domain spreading sequences: the common WH spreading codes; the VM spreading vectors given by (5.20); the complex-valued WH (CWH) code vectors given by (5.33); the rotated VM (RVM) code vectors given by (5.37), with $\theta = \pi/32$. The baseband discrete-time multipath channel $\{g(\ell)\}_{\ell=0}^{L}$ is a FIR filter of order $L = 3$, whose transfer function is given by

$$G(z) = (1 - \zeta_1\, z^{-1})\,(1 - \zeta_2\, z^{-1})\,(1 - \zeta_3\, z^{-1})\,, \quad (5.39)$$

where the group $(\zeta_1, \zeta_2, \zeta_3)$ of its three zeros assumes a different configuration in each Monte Carlo run. During the first 16 runs, we set $\zeta_1 = e^{i\frac{2\pi}{N}m_1}$ (one zero on the subcarriers), where, in each run, m_1 takes on a different value in $\{0, 1, \ldots, N-1\}$, whereas the magnitudes and phases of ζ_2 and ζ_3, which are modeled as mutually independent random variables uniformly distributed over the intervals $(0, 2)$ and $(0, 2\,\pi)$, respectively, are randomly and independently generated from run to run. During the subsequent $\binom{16}{2} = 120$ runs, we set

$\zeta_1 = e^{i\frac{2\pi}{N}m_1}$ and $\zeta_2 = e^{i\frac{2\pi}{N}m_2}$ (two zeros on the subcarriers), where, in each run, m_1 and m_2 take on a different value in $\{0, 1, \ldots, N-1\}$, with $m_1 \neq m_2$, whereas the magnitude and phase of ζ_3, which are modeled as mutually independent random variables uniformly distributed over the intervals $(0, 2)$ and $(0, 2\,\pi)$, respectively, are randomly and independently generated from run to run. During the last $\binom{16}{3} = 560$ runs, we set $\zeta_1 = e^{i\frac{2\pi}{N}m_1}$, $\zeta_2 = e^{i\frac{2\pi}{N}m_2}$ and $\zeta_3 = e^{i\frac{2\pi}{N}m_3}$ (three zeros on the subcarriers), where, in each run, m_1, m_2 and m_3 take on a different value in $\{0, 1, \ldots, N-1\}$, with $m_1 \neq m_2 \neq m_3$. In this way, one obtains $16 + 120 + 560 = 696$ independent channel realizations. According to assumption A2, the entries of the noise vector $\mathbf{v}(k)$ [see eq. (5.6)] are modeled as zero-mean independent identically-distributed (iid) complex circular Gaussian random variables, with variance σ_v^2, and the SNR of the desired user is defined as $\text{SNR} \triangleq (\sigma_b^2 \, \|\mathbf{c}_1\|^2)/\sigma_v^2$ (since $\|\mathbf{c}_j\|^2 = 1$, $\forall j \in \{1, 2, \ldots, J\}$, all the users undergo the same SNR).

For both CP- and ZP-based systems, employing the aforementioned four different spreading sequences, we carried out a comparative performance study of the L-ZF, L-MMSE, WL-ZF and WL-MMSE detectors[5]. At first sight, it seems that the synthesis of the L-ZF detector given by (5.8), which does not depend on the statistics of the received data, requires knowledge of the spreading codes of all the active users, which is an unreasonable requirement in the downlink. However, following the same lines of [44], this problem can be circumvented by implementing the L-ZF detector by means of the following SOS-based *subspace* representation

$$\mathbf{f}_{\text{L-ZF},1} = \mathbf{V}_s \, (\mathbf{\Lambda}_s - \sigma_v^2 \, \mathbf{I}_J)^{-1} \, \mathbf{V}_s^H \, \mathbf{\Xi}_1 \, \mathbf{g} \,, \tag{5.40}$$

where $\mathbf{V}_s \in \mathbb{C}^{R \times J}$ collects the eigenvectors associated with the J largest eigenvalues of $\mathbf{R}_{\mathbf{rr}}$ (arranged in descending order), which represents the diagonal entries of $\mathbf{\Lambda}_s \triangleq \text{diag}[\lambda_1, \lambda_2, \ldots, \lambda_J] \in \mathbb{R}^{J \times J}$, whereas: for a CP-based system ($R = N$), $\mathbf{\Xi}_1 = \mathbf{\Xi}_{\text{cp},1} \triangleq \widetilde{\mathbf{\Phi}}_{\text{cp},1} \, \mathbf{\Omega}_{\text{cp}} \in \mathbb{C}^{N \times L_{\text{p}}}$ is a known full-column rank matrix, with $\widetilde{\mathbf{\Phi}}_{\text{cp},1} \in \mathbb{C}^{N \times N}$ being a nonsingular circulant matrix, whose first column is given by $\widetilde{\mathbf{c}}_1 \triangleq \mathbf{W}_{\text{IDFT}} \, \mathbf{c}_1 \in \mathbb{C}^N$; for a ZP-based system ($R = P$), $\mathbf{\Xi}_1 = \mathbf{\Xi}_{\text{zp},1} \triangleq \widetilde{\mathbf{\Phi}}_{\text{zp},1} \, \mathbf{\Omega}_{\text{zp}} \in \mathbb{C}^{P \times L_{\text{p}}}$, where

[5] In the sequel, for notational convenience, a particular detector, which operates in a system employing a given set of spreading sequences, will be synthetically referred to through the acronym of the detector followed by the acronym of the code enclosed in round brackets; for example, the notation "L-ZF (WH)" means that the L-ZF detector is used at the receiver and, at the same time, WH spreading codes are employed at the transmitter.

$\widetilde{\mathbf{\Phi}}_{\text{zp},1} \in \mathbb{C}^{P\times P}$ is a known lower triangular Toeplitz [56] matrix having as first column $[\widetilde{\mathbf{c}}_1^T, 0, \ldots, 0]^T$ and as first row $[\widetilde{c}_1^{(0)}, 0, \ldots, 0]$. In the subspace-based form (5.40), apart from $\mathbf{g}$ and the eigenstructure of $\mathbf{R}_{\mathbf{rr}}$ (which can be consistently estimated from the received data), the synthesis of the L-ZF detector requires only knowledge of the desired code vector $\mathbf{c}_1$. For a fair comparison, we implemented the subspace-based version of the L-MMSE detector defined in (5.9), which can be expressed as [44, 93] (see also chapter 4 for more details)

$$\mathbf{f}_{\text{L-MMSE},1} = \mathbf{V}_s\, \mathbf{\Lambda}_s^{-1}\, \mathbf{V}_s^H\, \mathbf{\Xi}_1\, \mathbf{g}\,. \tag{5.41}$$

The derivations reported in [44, 93], which exclusively consider linear receiving structure, can be suitably extended to obtain the subspace versions of the WL-ZF and WL-MMSE detectors given by (5.25) and (5.26), respectively, thus obtaining (for the sake of brevity, we omit the mathematical details)

$$\mathbf{f}_{\text{WL-ZF},1} = \mathbf{U}_s\, (\mathbf{\Sigma}_s - \sigma_v^2\, \mathbf{I}_J)^{-1}\, \mathbf{U}_s^H \begin{bmatrix} \mathbf{\Xi}_1\, \mathbf{g} \\ \mathbf{\Xi}_1^*\, \mathbf{g}^* \end{bmatrix}\,, \tag{5.42}$$

$$\mathbf{f}_{\text{WL-MMSE},1} = \mathbf{U}_s\, \mathbf{\Sigma}_s^{-1}\, \mathbf{U}_s^H \begin{bmatrix} \mathbf{\Xi}_1\, \mathbf{g} \\ \mathbf{\Xi}_1^*\, \mathbf{g}^* \end{bmatrix}\,, \tag{5.43}$$

where $\mathbf{U}_s \in \mathbb{C}^{2R\times J}$ collects the eigenvectors associated with the J largest eigenvalues $\mu_1, \mu_2, \ldots, \mu_J$ of $\mathbf{R}_{\mathbf{zz}}$ (arranged in descending order) and $\mathbf{\Sigma}_s \stackrel{\triangle}{=} \text{diag}[\mu_1, \mu_2, \ldots, \mu_J] \in \mathbb{R}^{J\times J}$. In all the experiments, sample estimates of the eigenvectors and eigenvalues (including the noise variance σ_v^2 needed for the synthesis of the ZF detectors) of $\mathbf{R}_{\mathbf{rr}}$ and $\mathbf{R}_{\mathbf{zz}}$ were obtained in batch-mode from the sample autocorrelation matrices $\widehat{\mathbf{R}}_{\mathbf{rr}}$ and $\widehat{\mathbf{R}}_{\mathbf{zz}}$, respectively, by using a data record of $K = 500$ symbols. Finally, as performance measure, we resorted to the average BER (ABER) at the output of the considered receivers: after estimating the detector weight vectors on the basis of the given data record, for each of the 696 Monte Carlo run (wherein, besides the channel impulse response, independent sets of noise and data sequences were randomly generated), an independent record of $K_{\text{aber}} = 10^5$ symbols was considered to evaluate the ABER.

Experiment 5.1 (ABER versus SNR) *:* in the first group of experiments, we evaluated the performances of the considered receivers as a function of the SNR ranging from 0 to 20 dB.

In the first two experiments, we preliminarily studied the performances of the L-ZF and L-MMSE detectors: since linear receivers can work only when

$J \leq N$, we considered in these experiments underloaded CP- and ZP-based systems, with $J = 10$ active users. In Fig. 5.1, we considered a CP-based sys-

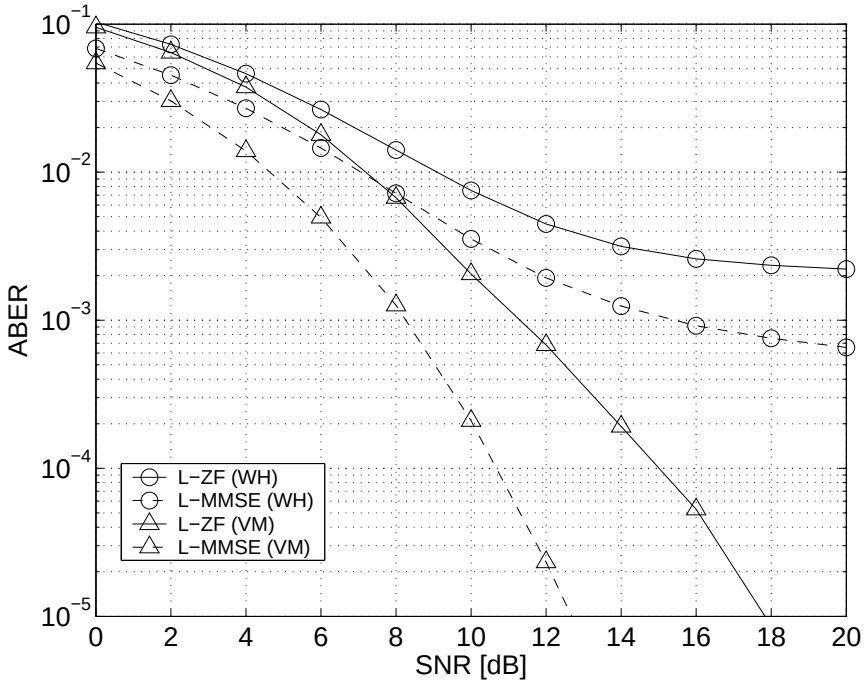

Figure 5.1: ABER versus SNR (CP-based downlink, underloaded system with $J = 10$ users, linear receiving structures).

tem employing either WH or VM spreading codes[6]. In this case, it is apparent from Fig. 5.1 that the performances of both the "L-ZF (WH)" and "L-MMSE (WH)" detectors exhibit a marked floor in the high SNR region, which is the natural consequence of the fact that, for a CP-based downlink, WH spreading sequences do not ensure the existence of L-ZF solutions when the channel transfer function exhibits zeros located on the subcarriers. On the other hand, when VM codes are used, perfect symbol recovery in the absence of noise is guaranteed regardless of the channel zero locations. In fact, as it is shown in

[6]The results regarding CWH code vectors are not reported since, for underloaded systems, they end up to the WH spreading sequences; additionally, in the same scenario, we do not report the results concerning the RVM spreading vectors since they are very similar to those presented for the VM code vectors.

Fig. 5.1, the curves of both the "L-ZF (VM)" and "L-MMSE (VM)" detectors go down very quickly as the SNR increases, thus assuring a huge performance gain with respect to the "L-ZF (WH)" and "L-MMSE (WH)" receivers. The

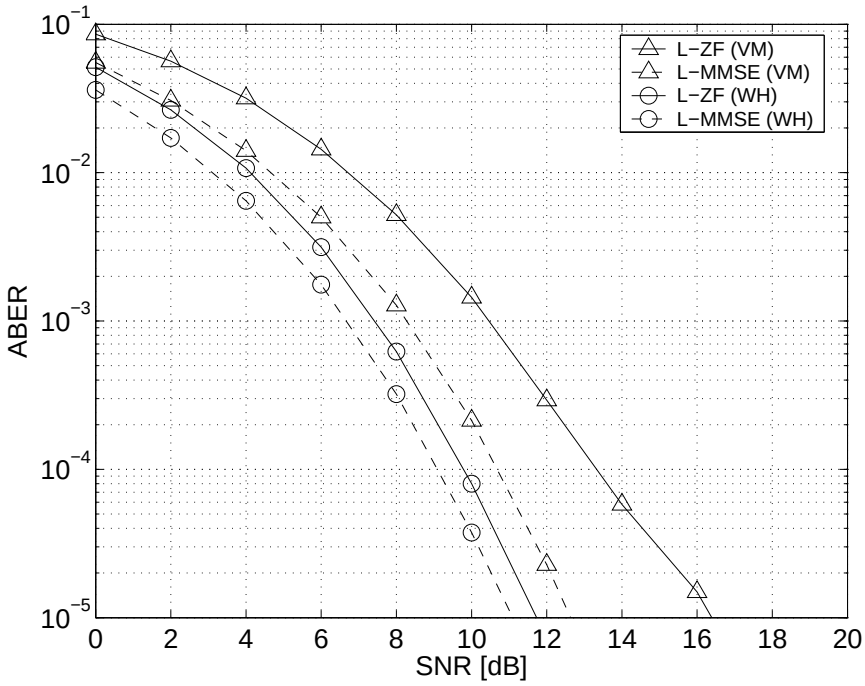

Figure 5.2: ABER versus SNR (ZP-based downlink, underloaded system with $J = 10$ users, linear receiving structures).

results of Fig. 5.2 were instead obtained by considering a ZP-based downlink. In this scenario, both WH and VM codes assure the existence of L-ZF solutions for any FIR channel of order $L < L_{\mathrm{p}}$. Indeed, as it is apparent from Fig. 5.2, the performances of all the receivers under comparison rapidly improve for increasing values of the SNR. As regards the L-ZF receivers, it is noteworthy that the "L-ZF (WH)" detector performs better than the "L-ZF (VM)" one: specifically, with respect to the "L-ZF (VM)" receiver, the "L-ZF (WH)" detector saves about 4 dB in transmitter power, for a target ABER of 10^{-4}. This means that, in comparison with VM spreading, WH codes lead to a reduced noise enhancement at the receiver output. Anyway, this performance gap is substantially halved if one brings the same comparison between the performances of

the "L-MMSE (WH)" and "L-MMSE (VM)" detectors.

In the following two experiments, we investigated the performances of the WL-ZF and WL-MMSE detectors: since WL receivers can work even when $J > N$, we simulated in these experiments overloaded CP- and ZP-based systems with $J = 20$ active users. With reference to a CP-based system, re-

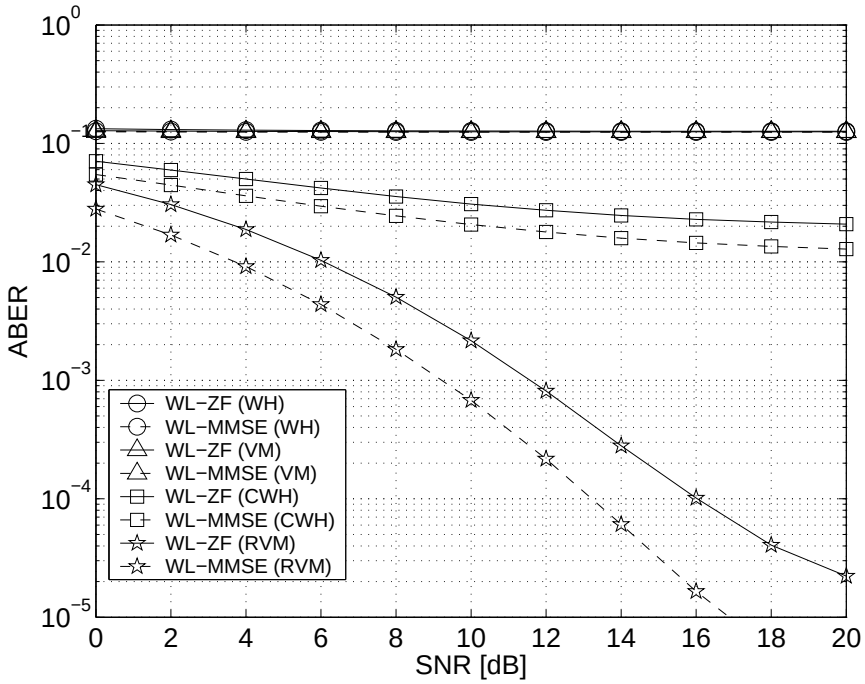

Figure 5.3: ABER versus SNR (CP-based downlink, overloaded system with $J = 20$ users, WL receiving structures).

sults of Fig. 5.3 show that the "WL-ZF (WH)", "WL-MMSE (WH)", "WL-ZF (VM)" and "WL-MMSE (VM)" receivers do not work at all. As previously pointed out in Remarks 6 and 7, these catastrophic performances arise since, not only the WH spreading codes, but also the VM code vectors do not assure the full-column rank property of the augmented code matrix $\overline{\mathbf{C}}$ in overloaded environments, which is a necessary condition for the existence of WL-ZF solutions in CP-based systems. In addition, since the code matrix given by (5.33) does not satisfy condition $\overline{\mathrm{D}}_{\mathrm{cp}}$, the curves of both the "WL-ZF (CWH)" and "WL-MMSE (CWH)" detectors exhibit an unacceptable floor for moderate-

to-high values of the SNR. In contrast, it can be seen from the same figure that the proposed RVM spreading vectors (5.37), which ensure the existence of universal WL-ZF solutions in both CP- and ZP-based overloaded systems, allow the "WL-ZF (RVM)" and "WL-MMSE (RVM)" receivers to work very well. Fig. 5.4 reports the ABER curves of the receivers under comparison for

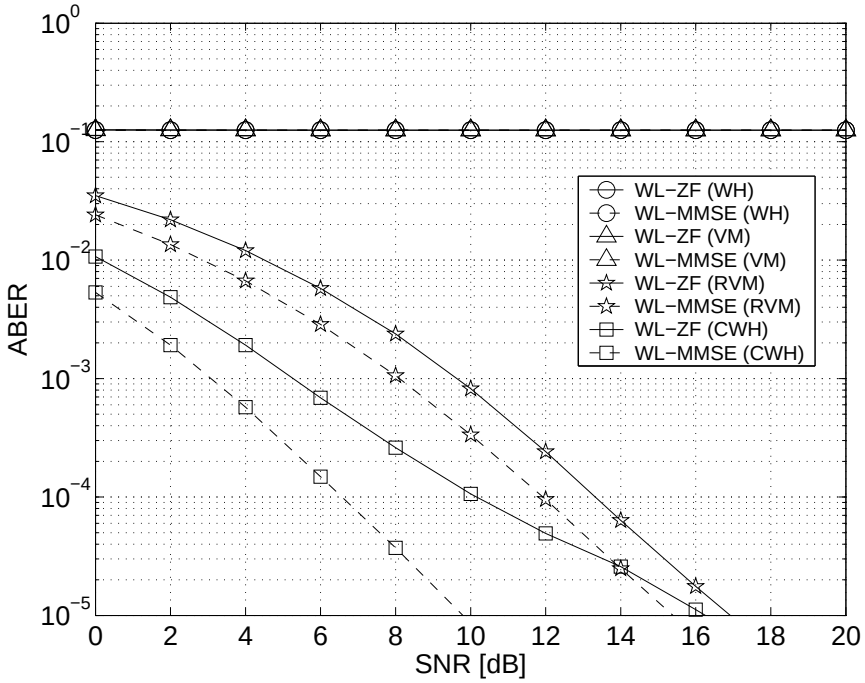

Figure 5.4: ABER versus SNR (ZP-based downlink, overloaded system with $J = 20$ users, WL receiving structures).

a ZP-based system. We recall that, in this case, the full-column rank property of $\overline{\mathbf{C}}$ is a necessary and sufficient condition for the existence of WL-ZF solutions. Indeed, besides corroborating the uselessness of the "WL-ZF (WH)", "WL-MMSE (WH)", "WL-ZF (VM)" and "WL-MMSE (VM)" receivers in the considered overloaded setting, results of Fig. 5.4 confirm that both the proposed CWH and RVM code vectors ensure the existence of universal WL-ZF solutions, by showing that the curves of the "WL-ZF (CWH)", "WL-MMSE (CWH)", "WL-ZF (RVM)" and "WL-MMSE (RVM)" rapidly fall away as the SNR goes up. In particular, as already evidenced in the linear case, due to

noise amplification effects, the "WL-ZF (CWH)" and "WL-MMSE (CWH)" detectors perform better than the corresponding counterparts "WL-ZF (RVM)" and "WL-MMSE (RVM)", expecially for low SNR values, by guaranteeing a significant saving in transmitter power, for a given value of the ABER.

Experiment 5.2 (ABER versus number of users) *:* in the second group of experiments, the performances of the considered receivers were studied as a function of the number J of active users, by setting SNR $= 10$ dB. As previously done, we investigated the performances of linear and WL receivers separately.

Fig. 5.5 and Fig. 5.6 report the performances of the L-ZF and L-MMSE detectors, when they are employed in both CP- and ZP-based underloaded systems, which use either WH or VM spreading sequences (the observation made in footnote 6 still applies to this case). With reference to a CP-based

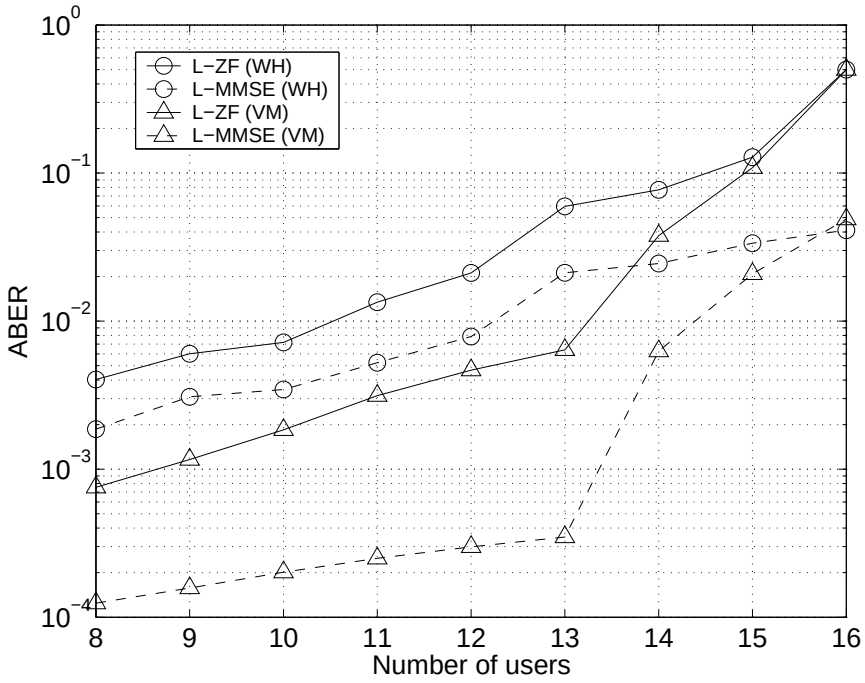

Figure 5.5: ABER versus number J of users (CP-based downlink, SNR $= 10$ dB, linear receiving structures).

system, results of Fig. 5.5 shows that, as long as the number of users is less than

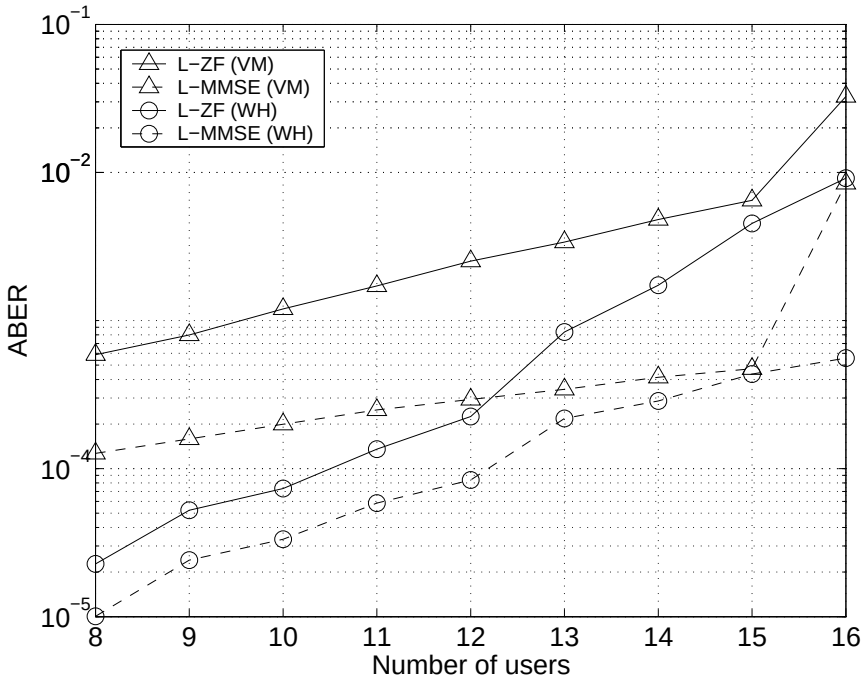

Figure 5.6: ABER versus number J of users (ZP-based downlink, SNR $= 10$ dB, linear receiving structures).

the threshold $N - L = 13$ (see Remarks 3 and 4), the "L-ZF (VM)" and "L-MMSE (VM)" detectors significantly outperform their "L-ZF (WH)" and "L-MMSE (WH)" corresponding counterparts. However, as soon as the number of active users gets over $J = 13$, in which case universal perfect symbol recovery in the absence of noise cannot be guaranteed, the performances of both the "L-ZF (VM)" and "L-MMSE (VM)" detectors rapidly deteriorate as the system load grows, by approaching the curves of the "L-ZF (WH)" and "L-MMSE (WH)" receivers. On the other hand, it can be seen from Fig. 5.6 that, for a ZP-based downlink, wherein the linear independence of the code vectors is a sufficient and necessary condition for assuring up to N users the existence of universal ZF solutions, all the receivers under comparison enable to achieve a greater system capacity than a CP-based system. In particular, according with the results of Fig. 5.2, the WH spreading sequences allow both the "L-ZF (WH)" and "L-MMSE (WH)" detectors to outperform their "L-ZF (VM)" and

"L-MMSE (VM)" counterparts, respectively, for all the considered values of J. It is worthwhile to note that, with $J = N = 16$ users, the "L-MMSE (WH)" detector is able to assure an ABER of about $5 \cdot 10^{-4}$ at its output, whereas the "L-MMSE (VM)" one exhibits competitive performances, i.e., less than $5 \cdot 10^{-4}$, only up to 15 users.

In the last two experiments, we investigated the performances of the WL-ZF and WL-MMSE detectors as a function of the number J of users, ranging from an underloaded ($J \leq N$) to an overloaded ($J > N$) system.

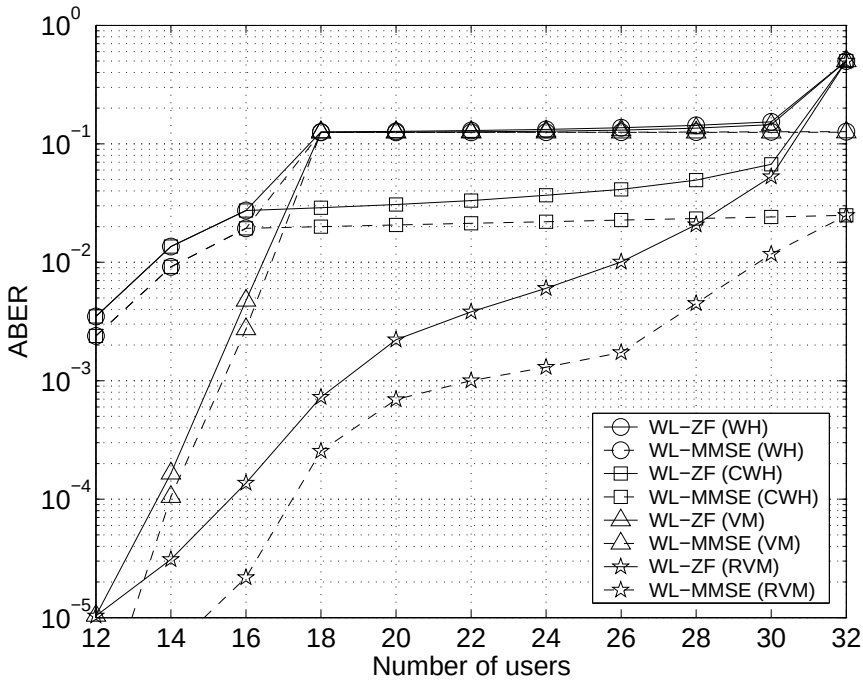

Figure 5.7: ABER versus number J of users (CP-based downlink, SNR $= 10$ dB, WL receiving structures).

For a CP-based downlink, it can be seen from Fig. 5.7 that, paying no attention to the uninteresting cases of WH and CWH spreading sequences, which do not guarantee channel-irrespective perfect symbol recovery in both underloaded and overloaded CP-based systems, the "WL-ZF (VM)" and "WL-MMSE (VM)" detectors perform comparably to the "WL-ZF (RVM)" and "WL-MMSE (RVM)" ones only for $J = 12$ active users. Beyond this value,

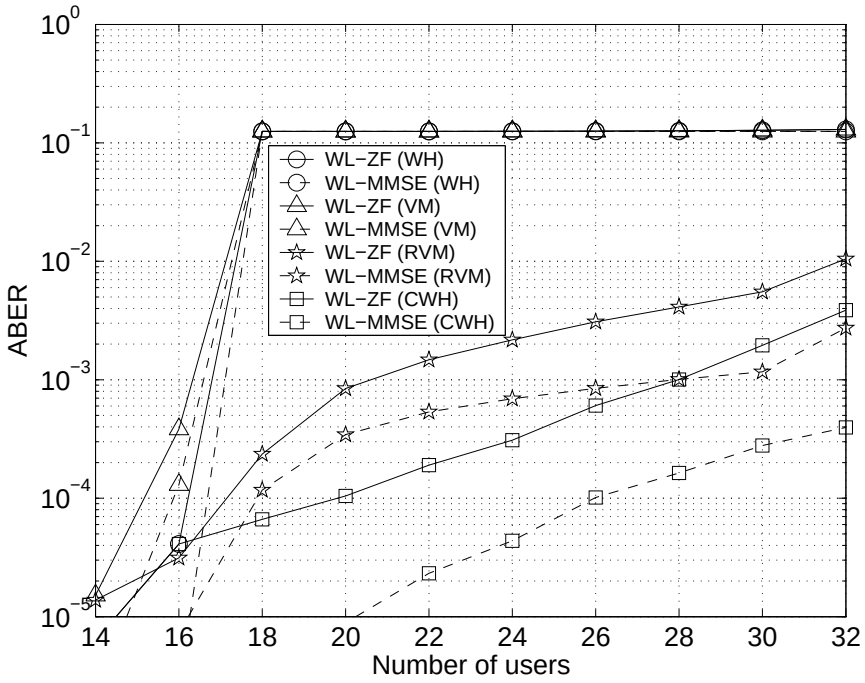

Figure 5.8: ABER versus number J of users (ZP-based downlink, SNR $= 10$ dB, WL linear receiving structures).

while the performances of the "WL-ZF (VM)" and "WL-MMSE (VM)" receivers get worse very quickly, both the "WL-ZF (RVM)" and "WL-MMSE (RVM)" detectors still work satisfactorily up to $2\,(N - L) = 26$ users (see Remarks 9 and 11), by exhibiting ABER values less than or equal to 10^{-2} and $2 \cdot 10^{-3}$, respectively. Beyond the threshold $J = 26$, whereupon the existence of universal WL-ZF solutions cannot be ensured, the performances of the "WL-ZF (RVM)" and "WL-MMSE (RVM)" detectors rapidly worsen as J increases, and became comparable to those of the "WL-ZF (CWH)" and "WL-MMSE (CWH)" receivers. Finally, with reference to a ZP-based system, the curves depicted in Fig. 5.8 evidence that the performances of the "WL-ZF (WH)", "WL-MMSE (WH)", "WL-ZF (VM)" and "WL-MMSE (VM)" receivers are very poor when the system becomes overloaded. Furthermore, it is apparent that the proposed RVM and CWH code vectors allow both the WL-ZF and WL-MMSE receiver to manage a number of users which is significantly

larger than the number of subcarriers. Remarkably, with $J = 2\,N = 32$ users, the "WL-MMSE (CWH)" detector is able to assure an ABER of $4 \cdot 10^{-4}$ at its output, whereas the ABER performance of the "WL-MMSE (VM)" is below 10^{-3} up to 30 users. On the basis of these experiments, we maintain that, among the different spreading techniques considered herein, the RVM code vectors turn out to be the best choice for both underloaded and overloaded CP-based systems, equipped with both linear and WL receiving structure, whereas the CWH spreading vectors allow both linear and WL detectors to exhibit the best performances in both underloaded and overloaded ZP-based systems.

Conclusions

In this thesis, the role of the Widely-Linear processing in the narrowband and wideband systems has been proposed.

With reference to narrowband systems, we analyzed the constant modulus cost function under the general assumptions that improper modulation schemes of practical interest are employed and the baseband equivalent of the channel impulse response is complex-valued. This study allows one to determine a broad family of undesired minima of the CM cost function, which do not lead to perfect symbol recovery in the absence of noise. Successively, in this contest, we applied widely-linear approach providing the mathematical conditions assuring perfect symbol recovery in the absence of noise. Furthermore, we enlightened that, similarly to the L-FS-CM equalizer, the performances of WL-CM equalizers suffer from the presence of undesired global minima. To overcome this drawback we proposed to resort to a constraint WL-CM equalizer.

In the context of wideband systems, with reference to DS-CDMA technique, we developed performance comparisons between ideal and data-estimated WL-MOE and L-MOE receivers. With reference to the ideal implementation, we investigated the relative performances of the WL-MOE and L-MOE receivers in the high-SNR regime. In this case, we provided a necessary and sufficient condition on the spreading codes, which allows the WL-MOE receiver to achieve perfect MAI suppression even in overloaded downlink configurations. As regards the data-estimated versions of the WL-MOE and L-MOE receivers, we derived easily interpretable formulas, which allow one to obtain clear insights about the effects of different parameters on performances. In a nutshell, compared with the L-MOE one, the performance of the WL-MOE receiver turns out to be more sensitive to finite-sample-size effects, and the performance gains predicted by the theory can be achieve in practice only by resorting to the more sophisticated subspace-based implementation. Then, to assess of the effects of channel-estimation errors we have extended

the previous analysis. Specifically, we presented a comprehensive performance comparison between different versions of the L- and WL-MOE receivers with blind CE, when both the ACM and the channel impulse response of the desired user are estimated from a finite sample-size. This analysis allows to conclude that with reference to their subspace-based implementations, for moderate-to-high values of the SNR, errors in estimating the L-SUB-CE and WL-SUB-CE receivers are essentially due to ACM estimation. The same is not true for the L-SMI-CE and WL-SMI-CE receivers, implemented by using the sample ACM directly, for which CE errors undesirably combine with ACM errors; however, compared with the known-channel case, CE errors adversely affect the SINR performances of L-SMI-CE and WL-SMI-CE receivers in a similar way. Therefore, when considering finite sample-size implementation, the more sophisticated subspace-based implementation is an effective method to assure that the WL-MOE receiver (with or without CE) significantly outperform (for low-to-moderate values of the SNR) its linear counterpart. In this case, for a given channel length, the WL-MOE receiver allows one to work with an increased number of users, which makes it a viable choice in heavily-congested DS-CDMA networks.

Finally, in the last part of this thesis, we tackled the problem of deriving mathematical conditions guaranteeing perfect symbol recovery in the absence of noise for either CP-based or ZP-based MC-CDMA downlink transmissions, which employ frequency-domain symbol-spreading. The conditions derived are channel-independent and are expressed in terms of relatively simple system design constraints, regarding the maximum number of allowable users and their spreading sequences. Specifically, it was first shown that, similarly to a ZP-based MC-CDMA downlink and differently from CP-OFDM systems, L-ZF-MUD, which is confined only to underloaded systems and can be used when transmitted symbols are either proper or improper, can be guaranteed for a CP-based MC-CDMA downlink, even when the channel transfer function exhibits nulls on some used subcarriers. On the other hand, when the information-bearing symbols are improper, it was further shown that, for both CP- and ZP-based systems, WL-ZF-MUD allows one to successfully operate even in overloaded scenarios, by doubling the system capacity, regardless of the channel zero locations. However, such an increased throughput can be achieved as long as appropriate complex-valued spreading codes are used.

Basing on the above results, the suggestions for future work are twofold. As regard to narro- and wide-band systems, future work could include an extension of the performance analyses to time-variant channels, since in such an

environment new issues arise in channel equalization. Moreover, more attention should be devoted to the role of widely linear processing in ultrawideband systems, which represent a suitable technology to achieve high data rates in wireless communications.

Appendix A

Constant Modulus Equalizers

A.1 Proof of Theorem 3.1

We distinguish the following groups of stationary points.

G0. The only vector belonging to this group is $\overline{\mathbf{q}}_0 = \mathbf{O}_K$, which trivially fulfills $\widetilde{\mathbf{g}}(\mathbf{q}) = \mathbf{O}_K$.

G1. In this group, there are all the vectors satisfying $\widetilde{\mathbf{g}}(\mathbf{q}) = \mathbf{O}_K$ ensuring an ISI-free equalizer output, that is, they exhibit only one nonzero entry $\overline{q}_{i_1}$, with $i_1 \in \{0, 1, \ldots, K-1\}$, i.e., $\overline{\mathbf{q}} = \overline{q}_{i_1}\, \mathbf{e}_{i_1}$.

Vector $\overline{\mathbf{q}} = \overline{q}_{i_1}\, \mathbf{e}_{i_1}$ satisfies $\widetilde{\mathbf{g}}(\mathbf{q}) = \mathbf{O}_K$ if and only if (iff) $\left[(\kappa_s + 3\,\sigma_s^4)\,|\overline{q}_{i_1}|^2 - \gamma_s\,\sigma_s^2\right] \overline{q}_{i_1}\, e^{j\,2\pi\beta i_1} = 0$; since $\overline{q}_{i_1}\, e^{j\,2\pi\beta i_1} \neq 0$ and $\kappa_s = \gamma_s\,\sigma_s^2 - 3\,\sigma_s^4$, this equation is equivalent to $\gamma_s\,\sigma_s^2\,|\overline{q}_{i_1}|^2 - \gamma_s\,\sigma_s^2 = 0 \Leftrightarrow |\overline{q}_{i_1}| = 1$. Thus, the general expression of the vectors belonging to this group is

$$\overline{\mathbf{q}}_1 = e^{j\,\theta}\, \mathbf{e}_{i_1}, \quad \text{with } \theta \in [0, 2\pi) \text{ and } i_1 \in \{0, 1, \ldots, K-1\}. \tag{A.1}$$

G2. In this group, there are all the vectors satisfying $\widetilde{\mathbf{g}}(\mathbf{q}) = \mathbf{O}_K$ leading to an ISI-contaminated equalizer output, i.e., the number of their nonzero entries is greater than one.

To prove the existence of undesired local minima, it is sufficient to focus attention on those solutions of $\widetilde{\mathbf{g}}(\mathbf{q}) = \mathbf{O}_K$ possessing only two nonzero entries $\overline{q}_{i_1}$ and $\overline{q}_{i_2}$, with $i_1 \neq i_2 \in \{0, 1, \ldots, K-1\}$, i.e., $\overline{\mathbf{q}} = \overline{q}_{i_1}\, \mathbf{e}_{i_1} + \overline{q}_{i_2}\, \mathbf{e}_{i_2}$. After some algebraic manipulations, it can been seen that $\overline{\mathbf{q}}$ fulfills $\widetilde{\mathbf{g}}(\mathbf{q}) = \mathbf{O}_K$ iff

the following system is satisfied:

$$\begin{cases} [(\kappa_s + \sigma_s^4)\,|\overline{q}_{i_1}|^2 + 2\,\sigma_s^4\,(|\overline{q}_{i_1}|^2 + |\overline{q}_{i_2}|^2) - \gamma_s\,\sigma_s^2]|\overline{q}_{i_1}|^2 \\ \qquad +\sigma_s^4\,(\overline{q}_{i_1}^*\,\overline{q}_{i_2})^2\,e^{j\,2\pi\beta(i_1-i_2)} = 0 \\ \left[(\kappa_s + \sigma_s^4)\,|\overline{q}_{i_2}|^2 + 2\,\sigma_s^4\,(|\overline{q}_{i_1}|^2 + |\overline{q}_{i_2}|^2) - \gamma_s\,\sigma_s^2\right]|\overline{q}_{i_2}|^2 \\ \qquad +\sigma_s^4\,(\overline{q}_{i_1}\,\overline{q}_{i_2}^*)^2\,e^{-j\,2\pi\beta(i_1-i_2)} = 0 \end{cases} \tag{A.2}$$

Since the quantities enclosed in square brackets are real-valued, fulfillment of (A.2) requires that $(\overline{q}_{i_1}\,\overline{q}_{i_2}^*)^2\,e^{-j\,2\pi\beta(i_1-i_2)}$ be a real number, which happens when

$$\angle\overline{q}_{i_1} - \angle\overline{q}_{i_2} = \pi\,\beta\,(i_1 - i_2) + \pi\,\ell_{i_1,i_2} \tag{A.3}$$

or when

$$\angle\overline{q}_{i_1} - \angle\overline{q}_{i_2} = \pi\,\beta\,(i_1 - i_2) + \frac{\pi}{2} + \pi\,\ell_{i_1,i_2}, \tag{A.4}$$

with $\ell_{i_1,i_2} \in \mathbb{Z}$.

In these cases, system (A.2) can be split up into the two different systems

$$\begin{cases} \gamma_s\,|\overline{q}_{i_1}|^2 + \sigma_s^2\,\delta_{i_1,i_2}\,|\overline{q}_{i_2}|^2 = \gamma_s \\ \sigma_s^2\,\delta_{i_1,i_2}\,|\overline{q}_{i_1}|^2 + \gamma_s\,|\overline{q}_{i_2}|^2 = \gamma_s \end{cases} \quad , \quad \text{with } \delta_{i_1,i_2} = 1, 3\,, \tag{A.5}$$

which involve only the magnitudes of $\overline{q}_{i_1}$ and $\overline{q}_{i_2}$. Specifically, it results that $\delta_{i_1,i_2} = 3$ when $(\overline{q}_{i_1}\,\overline{q}_{i_2}^*)^2\,e^{-j\,2\pi\beta(i_1-i_2)}$ is positive, whereas one has $\delta_{i_1,i_2} = 1$ when $(\overline{q}_{i_1}\,\overline{q}_{i_2}^*)^2\,e^{-j\,2\pi\beta(i_1-i_2)}$ is negative.

By resorting to the Cramer's rule, it is easily seen that, if $\gamma_s \neq \sigma_s^2\,\delta_{i_1,i_2}$, the solution of system (A.5) is unique and is given by $|\overline{q}_{i_1}|^2 = |\overline{q}_{i_2}|^2 = \gamma_s/(\gamma_s + \sigma_s^2\,\delta_{i_1,i_2})$. On the other hand, when $\gamma_s = \sigma_s^2\,\delta_{i_1,i_2}$, system (A.5) admits an infinite number of solutions characterized by the relation $|\overline{q}_{i_1}|^2 + |\overline{q}_{i_2}|^2 = 1$.

In summary, the general expressions of the vectors belonging to this group are given by

$$\overline{\mathbf{q}}_{2,1} = e^{j\,\theta}\sqrt{\gamma_s/(\gamma_s + 3\,\sigma_s^2)}\cdot[\mathbf{e}_{i_1} + (-1)^{\ell_{i_1,i_2}}\,e^{j\,\pi\beta(i_2-i_1)}\,\mathbf{e}_{i_2}] \tag{A.6}$$

and

$$\overline{\mathbf{q}}_{2,2} = e^{j\,\theta}\sqrt{\gamma_s/(\gamma_s + \sigma_s^2)}\cdot[\mathbf{e}_{i_1} - j\,(-1)^{\ell_{i_1,i_2}}e^{j\,\pi\beta(i_2-i_1)}\,\mathbf{e}_{i_2}], \tag{A.7}$$

for $\sigma_s^2 < \gamma_s < 3\,\sigma_s^2$, whereas, for $\gamma_s = 3\,\sigma_s^2$, one obtains

$$\overline{\mathbf{q}}_{2,3} = e^{j\,\theta} \cdot [\rho\,\mathbf{e}_{i_1} + (-1)^{\ell_{i_1,i_2}}\,e^{j\,\pi\beta(i_2-i_1)}\,\sqrt{1-\rho^2}\;\mathbf{e}_{i_2}],\;\; \text{for } \gamma_s = 3\,\sigma_s^2 \tag{A.8}$$

and, finally, for $\gamma_s = \sigma_s^2$, one has

$$\overline{\mathbf{q}}_{2,4} = e^{j\,\theta}\cdot[\rho\,\mathbf{e}_{i_1} - j\,(-1)^{\ell_{i_1,i_2}}\,e^{j\,\pi\beta(i_2-i_1)}\,\sqrt{1-\rho^2}\;\mathbf{e}_{i_2}],\;\; \text{for } \gamma_s = \sigma_s^2 \tag{A.9}$$

with $\theta \in [0, 2\pi)$, $\ell_{i_1,i_2} \in \mathbb{Z}$, $i_1 \neq i_2 \in \{0, 1, \ldots, K-1\}$ and $0 < \rho < 1$.

At this point, to find the local minima of $\widetilde{J}_{\text{cm}}(\mathbf{q})$, we have to study the positive definiteness of $\widetilde{\boldsymbol{\mathcal{H}}}(\overline{\mathbf{q}})$ given by (3.16), evaluated at each of the stationary points previously derived.

G0. Since $\widetilde{\boldsymbol{\mathcal{H}}}(\overline{\mathbf{q}}_0) = -2\,\gamma_s\,\sigma_s^2\,\mathbf{I}_K$, the cost function $\widetilde{J}_{\text{cm}}(\mathbf{q})$ has a local maximum at $\overline{\mathbf{q}}_0 = \mathbf{O}_K$.

G1. The matrix $\widetilde{\boldsymbol{\mathcal{H}}}(\overline{\mathbf{q}}_1)$ turns out to be diagonal, with diagonal entries $\{\widetilde{\boldsymbol{\mathcal{H}}}(\overline{\mathbf{q}}_1)\}_{i_1+1,i_1+1} = 2\,\sigma_s^2\,\gamma_s$ and $\{\widetilde{\boldsymbol{\mathcal{H}}}(\overline{\mathbf{q}}_1)\}_{i+1,i+1} = -2\,\sigma_s^2\,(\gamma_s - 2\,\sigma_s^2)$, for $i \in \{0, 1, \ldots, K-1\} - \{i_1\}$. Hence, if $\gamma_s \geq 2\,\sigma_s^2$, the diagonal entries of $\widetilde{\boldsymbol{\mathcal{H}}}(\overline{\mathbf{q}}_1)$ take on both positive and negative values and, consequently, the vector $\overline{\mathbf{q}}_1$ is a saddle point. On the other hand, in accordance with assumption A1, if $\gamma_s < 2\,\sigma_s^2$, the diagonal entries of the diagonal matrix $\widetilde{\boldsymbol{\mathcal{H}}}(\overline{\mathbf{q}}_1)$ are all positive and, thus, $\widetilde{J}_{\text{cm}}(\mathbf{q})$ has a local minimum at $\overline{\mathbf{q}}_1$.

G2.

First, the matrix $\widetilde{\boldsymbol{\mathcal{H}}}(\overline{\mathbf{q}}_{2,1})$ is nonsingular, with diagonal entries

$$\{\widetilde{\boldsymbol{\mathcal{H}}}(\overline{\mathbf{q}}_{2,1})\}_{i+1,i+1} = 2\,\gamma_s\,(\kappa_s + 2\,\sigma_s^4)/(\gamma_s + 3\,\sigma_s^2),\;\; \text{for}\, i \in \{i_1, i_2\}, \tag{A.10}$$

and

$$\{\widetilde{\boldsymbol{\mathcal{H}}}(\overline{\mathbf{q}}_{2,1})\}_{i+1,i+1} = -2\,\gamma_s\,(\kappa_s + 2\,\sigma_s^4)/(\gamma_s + 3\,\sigma_s^2),\\ \text{for}\, i \in \{0, 1, \ldots, K-1\} - \{i_1, i_2\}. \tag{A.11}$$

It is apparent that, regardless of κ_s, the matrix $\widetilde{\boldsymbol{\mathcal{H}}}(\overline{\mathbf{q}}_{2,1})$ cannot be positive definite since its diagonal entries take on both positive and negative values and, thus, $\widetilde{J}_{\text{cm}}(\mathbf{q})$ has a saddle point at $\overline{\mathbf{q}}_{2,1}$.

Second, it can be seen that $\widetilde{\boldsymbol{\mathcal{H}}}(\overline{\mathbf{q}}_{2,2})$ turns out to be diagonal, with diagonal entries

$$\{\widetilde{\boldsymbol{\mathcal{H}}}(\overline{\mathbf{q}}_{2,2})\}_{i+1,i+1} = 2\,\gamma_s\,\sigma_s^2,\;\; \text{for}\, i \in \{i_1, i_2\}, \tag{A.12}$$

and

$$\{\widetilde{\boldsymbol{\mathcal{H}}}(\overline{\mathbf{q}}_{2,2})\}_{i+1,i+1} = -2\,\gamma_s\,\kappa_s/(\gamma_s+\sigma_s^2),\ \text{ for}\, i \in \{0,1,\ldots,K-1\}-\{i_1,i_2\}. \tag{A.13}$$

If assumption A2 is fulfilled, i.e., $\kappa_s < 0$, the diagonal matrix $\widetilde{\boldsymbol{\mathcal{H}}}(\overline{\mathbf{q}}_{2,2})$ is positive definite since its diagonal entries are all positive and, hence, $\widetilde{J}_{\text{cm}}(\mathbf{q})$ has a local minimum at $\overline{\mathbf{q}}_{2,2}$.

Third, it can be verified that $\widetilde{\boldsymbol{\mathcal{H}}}(\overline{\mathbf{q}}_{2,3})$ is diagonal, with diagonal entries

$$\{\widetilde{\boldsymbol{\mathcal{H}}}(\overline{\mathbf{q}}_{2,3})\}_{i+1,i+1} = 6\,\sigma_s^4,\ \text{ for } i \in \{i_1,i_2\}, \tag{A.14}$$

$$\{\widetilde{\boldsymbol{\mathcal{H}}}(\overline{\mathbf{q}}_{2,3})\}_{i+1,i+1} = -2\,\sigma_s^4,\ \text{ for}\, i \in \{0,1,\ldots,K-1\}-\{i_1,i_2\}. \tag{A.15}$$

Since the diagonal entries of $\widetilde{\boldsymbol{\mathcal{H}}}(\overline{\mathbf{q}}_{2,3})$ assume both positive and negative values, $\widetilde{J}_{\text{cm}}(\mathbf{q})$ has a saddle point at $\overline{\mathbf{q}}_{2,3}$.

Finally, it results that $\widetilde{\boldsymbol{\mathcal{H}}}(\overline{\mathbf{q}}_{2,4})$ is a diagonal matrix, with positive diagonal entries

$$\{\widetilde{\boldsymbol{\mathcal{H}}}(\overline{\mathbf{q}}_{2,4})\}_{i+1,i+1} = 2\,\sigma_s^4,\ \text{ for}\, i \in \{0,1,\ldots,K-1\} \tag{A.16}$$

and, thus, $\widetilde{J}_{\text{cm}}(\mathbf{q})$ has a local minimum at $\overline{\mathbf{q}}_{2,4}$.

Appendix B

Equalization Techniques for DS-CDMA Systems

B.1 Proof of Lemma 4.1

Any vector $\mathbf{f}_j \in \mathbb{C}^{2N}$ can be uniquely decomposed as $\mathbf{f}_j = \mathbf{f}_{j,\mathrm{s}} + \mathbf{f}_{j,\mathrm{a}}$, where we defined the *symmetric part* $\mathbf{f}_{j,\mathrm{s}} \in \mathcal{S} \stackrel{\triangle}{=} \{\mathbf{f} = [\mathbf{f}_1^T, \mathbf{f}_2^T]^T \in \mathbb{C}^{2N} \,|\, \mathbf{f}_1 = \mathbf{f}_2^* \in \mathbb{C}^N\}$ and the *antisymmetric part* $\mathbf{f}_{j,\mathrm{a}} \in \mathcal{A} \stackrel{\triangle}{=} \{\mathbf{f} = [\mathbf{f}_1^T, \mathbf{f}_2^T]^T \in \mathbb{C}^{2N} \,|\, \mathbf{f}_1 = -\mathbf{f}_2^* \in \mathbb{C}^N\}$. Since both $\mathbf{h}_j$ and $\mathbf{q}_j(k)$ in (4.22) are symmetric, i.e., they belong to $\mathcal{S}$, one has $\mathrm{Re}[\mathbf{f}_j^H \mathbf{h}_j] = \mathbf{f}_{j,\mathrm{s}}^H \mathbf{h}_j$ and $\mathrm{Re}[\mathbf{f}_j^H \mathbf{q}_j(k)] = \mathbf{f}_{j,\mathrm{s}}^H \mathbf{q}_j(k)$ in (4.24), that is, the SINR (4.24) is not affected by the antisymmetric part $\mathbf{f}_{j,\mathrm{a}}$. Hence, the weight vector $\mathbf{f}_{j,\text{max-SINR}}$ maximizing $\mathrm{SINR}(\mathbf{f}_j)$ given by (4.24) can equivalently be obtained by maximizing the following constrained cost function:

$$\mathrm{SINR}'(\mathbf{f}_j) \stackrel{\triangle}{=} \frac{|\mathbf{f}_j^H \mathbf{h}_j|^2}{\mathrm{E}[|\mathbf{f}_j^H \mathbf{q}_j(k)|^2]} = \frac{|\mathbf{f}_j^H \mathbf{h}_j|^2}{\mathbf{f}_j^H \mathbf{R}_{\mathbf{q}_j \mathbf{q}_j} \mathbf{f}_j}\,, \quad \text{subject to } \mathbf{f}_j \in \mathcal{S}. \tag{B.1}$$

Note that in general $\mathrm{SINR}(\mathbf{f}_j) \neq \mathrm{SINR}'(\mathbf{f}_j)$, but they coincide for $\mathbf{f}_j \in \mathcal{S}$. The unconstrained maximization of $\mathrm{SINR}'(\mathbf{f}_j)$ leads [56] to the solution $\mathbf{f}'_{j,\text{max-SINR}} = \gamma_j \mathbf{R}_{\mathbf{q}_j \mathbf{q}_j}^{-1} \mathbf{h}_j$, with $\gamma_j \in \mathbb{C} - \{0\}$. At this point, we have to impose that $\mathbf{f}'_{j,\text{max-SINR}}$ satisfies the constraint $\mathbf{f}'_{j,\text{max-SINR}} \in \mathcal{S}$. To this respect, it can be verified that $\mathbf{R}_{\mathbf{q}_j \mathbf{q}_j}^{-1} \mathbf{h}_j \in \mathcal{S}$, hence, fulfillment of the constraint is ensured by imposing that γ_j be real, i.e., $\gamma_j = \gamma_j^*$. In conclusion, we can state that the general expression of the weight vector $\mathbf{f}_{j,\text{max-SINR}}$ maximizing $\mathrm{SINR}(\mathbf{f}_j)$

is given by $\mathbf{f}_{j,\text{max-SINR}} = \xi_j \, \mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}^{-1} \mathbf{h}_j$, with $\xi_j \stackrel{\triangle}{=} \text{Re}[\gamma_j] \in \mathbb{R} - \{0\}$. The corresponding maximum value of $\text{SINR}(\mathbf{f}_j)$ turns out to be $\text{SINR}(\mathbf{f}_{j,\text{max-SINR}}) = \mathbf{h}_j^H \mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}^{-1} \mathbf{h}_j$.

B.2 Relationships between $\overline{\text{SINR}}_{j,\text{max}}$, $\text{SINR}_{j,\text{L-MOE}}$ and $\text{SINR}_{j,\text{WL-MOE}}$ in the high-SNR regime

First of all, let us derive the expression of $\overline{\text{SINR}}_{j,\text{max}}$ [see (4.18)] in terms of σ_v^2. Under assumptions **(a1)–(a2)**, one has $\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j} = \overline{\mathbf{\Phi}}_j \, \overline{\mathbf{\Phi}}_j^H + \sigma_v^2 \, \mathbf{I}_N$. Hence, by resorting to the EVD of $\overline{\mathbf{\Phi}}_j \, \overline{\mathbf{\Phi}}_j^H$, one obtains $\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j} = \mathbf{V}_{j,s} \, \mathbf{\Sigma}_{j,s} \, \mathbf{V}_{j,s}^H + \sigma_v^2 \, \mathbf{I}_N$, where $\mathbf{V}_{j,s} \in \mathbb{C}^{N \times r_j}$ collects the eigenvectors associated with the r_j nonnull eigenvalues $\mu_{j,1}, \mu_{j,2}, \ldots, \mu_{j,r_j}$ of $\overline{\mathbf{\Phi}}_j \, \overline{\mathbf{\Phi}}_j^H$ (arranged in decreasing order), with $r_j \stackrel{\triangle}{=} \text{rank}(\overline{\mathbf{\Phi}}_j) \leq \min\{N, J-1\}$ and $\mathbf{\Sigma}_{j,s} \stackrel{\triangle}{=} \text{diag}(\mu_{j,1}, \mu_{j,2}, \ldots, \mu_{j,r_j}) \in \mathbb{R}^{r_j \times r_j}$. Relying on this decomposition and reasoning as in [94], the following series expansion of $\overline{\text{SINR}}_{j,\text{max}}$ holds:

$$\overline{\text{SINR}}_{j,\text{max}} = \boldsymbol{\phi}_j^H \, \mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}^{-1} \boldsymbol{\phi}_j = \frac{\boldsymbol{\phi}_j^H \mathbf{V}_{j,n} \, \mathbf{V}_{j,n}^H \, \boldsymbol{\phi}_j}{\sigma_v^2} + \boldsymbol{\phi}_j^H \mathbf{V}_{j,s} \, \mathbf{\Sigma}_{j,s}^{-1} \, \mathbf{V}_{j,s}^H \, \boldsymbol{\phi}_j + o(\sigma_v^2), \tag{B.2}$$

where $\mathbf{V}_{j,n} \in \mathbb{C}^{N \times (N-r_j)}$ collects the eigenvectors of $\overline{\mathbf{\Phi}}_j \, \overline{\mathbf{\Phi}}_j^H$ associated with its $N - r_j$ null eigenvalues. Eq. (B.2) shows that, as $\sigma_v^2 \to 0$, $\overline{\text{SINR}}_{j,\text{max}} \to +\infty$ if and only if (iff) $\boldsymbol{\phi}_j^H \mathbf{V}_{j,n} \, \mathbf{V}_{j,n}^H \, \boldsymbol{\phi}_j \neq 0$, which implies that $\boldsymbol{\phi}_j \notin \mathcal{N}(\mathbf{V}_{j,n}^H) \equiv \mathcal{R}(\overline{\mathbf{\Phi}}_j)$. It is noteworthy that this condition holds, $\forall j \in \{1, 2, \ldots, J\}$, iff the matrix $\mathbf{\Phi} \in \mathbb{C}^{N \times J}$ is full-column rank, i.e., $\text{rank}(\mathbf{\Phi}) = J$, which imposes that the number of users J must be smaller than or equal to the processing gain N (underloaded system). On the other hand, when $\boldsymbol{\phi}_j$ belongs to $\mathcal{R}(\overline{\mathbf{\Phi}}_j)$, it results that $\lim_{\sigma_v^2 \to 0} \overline{\text{SINR}}_{j,\text{max}} = \boldsymbol{\phi}_j^H \mathbf{V}_{j,s} \, \mathbf{\Sigma}_{j,s}^{-1} \, \mathbf{V}_{j,s}^H \, \boldsymbol{\phi}_j$, which evidences that, as $\sigma_v^2 \to 0$, $\overline{\text{SINR}}_{j,\text{max}}$ takes on a finite value.

At this point, we are able to establish the relationship existing between $\overline{\text{SINR}}_{j,\text{max}}$ and $\text{SINR}_{j,\text{L-MOE}}$ [see (4.18) and (4.50)], in the limiting case of vanishingly small noise. Preliminarily, we observe that, under assumptions **(a1)–(a2)**, one has $\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j^*} = \overline{\mathbf{\Phi}}_j \, \overline{\mathbf{\Phi}}_j^T$. By substituting (4.16) in (4.50) and

accounting for (4.18), after some algebraic manipulations, one obtains

$$\lim_{\sigma_v^2\to 0}\frac{\text{SINR}_{j,\text{L-MOE}}}{\overline{\text{SINR}}_{j,\max}}=\frac{2}{1+\lim\limits_{\sigma_v^2\to 0}\frac{\text{Re}\left[\boldsymbol{\phi}_j^H\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}^{-1}\overline{\boldsymbol{\Phi}}_j\,(\overline{\boldsymbol{\Phi}}_j^H\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}^{-1})^*\,\boldsymbol{\phi}_j^*\right]}{\overline{\text{SINR}}_{j,\max}}}\,. \tag{B.3}$$

By resorting to the limit formula for the Moore-Penrose inverse [24], it can be seen that $\lim_{\sigma_v^2\to 0}\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}^{-1}\overline{\boldsymbol{\Phi}}_j=(\overline{\boldsymbol{\Phi}}_j^H)^\dagger$ and $\lim_{\sigma_v^2\to 0}\overline{\boldsymbol{\Phi}}_j^H\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}^{-1}=(\overline{\boldsymbol{\Phi}}_j)^\dagger$. Consequently, we get $\lim_{\sigma_v^2\to 0}\text{Re}\big[\boldsymbol{\phi}_j^H\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}^{-1}\overline{\boldsymbol{\Phi}}_j\,(\overline{\boldsymbol{\Phi}}_j^H\mathbf{R}_{\mathbf{p}_j\mathbf{p}_j}^{-1})^*\,\boldsymbol{\phi}_j^*\big]=\text{Re}\big[\boldsymbol{\phi}_j^H(\overline{\boldsymbol{\Phi}}_j^H)^\dagger\,(\overline{\boldsymbol{\Phi}}_j^*)^\dagger\,\boldsymbol{\phi}_j^*\big]$, which can only assume finite values. Therefore, based on the previous discussion regarding the asymptotic expression of $\overline{\text{SINR}}_{j,\max}$, by virtue of (B.2) and (B.3), we can conclude that, if $\boldsymbol{\Phi}$ is full-column rank, then

$$\lim_{\sigma_v^2\to 0}\frac{\text{SINR}_{j,\text{L-MOE}}}{\overline{\text{SINR}}_{j,\max}}=2\,,\quad\forall j\in\{1,2,\ldots,J\}\,, \tag{B.4}$$

which additionally implies that, as $\sigma_v^2\to 0$, $\text{SINR}_{j,\text{L-MOE}}\to+\infty$, $\forall j\in\{1,2,\ldots,J\}$.

Let us now derive the expression of $\text{SINR}_{j,\text{WL-MOE}}$ [see (4.49)] in terms of σ_v^2. Under assumptions **(a1)**–**(a2)**, one has $\mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}=\overline{\mathbf{H}}_j\,\overline{\mathbf{H}}_j^H+\sigma_v^2\,\mathbf{I}_{2N}$. Reasoning as previously done for $\overline{\text{SINR}}_{j,\max}$, we express $\text{SINR}_{j,\text{WL-MOE}}$ explicitly in terms of σ_v^2 as follows:

$$\begin{aligned}\text{SINR}_{j,\text{WL-MOE}}&=\mathbf{h}_j^H\,\mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}^{-1}\,\mathbf{h}_j\\&=\frac{\mathbf{h}_j^H\mathbf{U}_{j,n}\,\mathbf{U}_{j,n}^H\,\mathbf{h}_j}{\sigma_v^2}+\mathbf{h}_j^H\mathbf{U}_{j,s}\,\boldsymbol{\Lambda}_{j,s}^{-1}\,\mathbf{U}_{j,s}^H\,\mathbf{h}_j+o(\sigma_v^2),\end{aligned} \tag{B.5}$$

where $\mathbf{U}_{j,s}\in\mathbb{C}^{2N\times\nu_j}$ collects the eigenvectors associated with the ν_j nonnull eigenvalues $\lambda_{j,1},\lambda_{j,2},\ldots,\lambda_{j,\nu_j}$ of $\overline{\mathbf{H}}_j\,\overline{\mathbf{H}}_j^H$ (arranged in decreasing order), with $\nu_j\stackrel{\triangle}{=}\text{rank}(\overline{\mathbf{H}}_j)\leq\min\{2\,N,J-1\}$ and $\boldsymbol{\Lambda}_{j,s}\stackrel{\triangle}{=}\text{diag}(\lambda_{j,1},\lambda_{j,2},\ldots,\lambda_{j,\nu_j})\in\mathbb{R}^{\nu_j\times\nu_j}$, whereas $\mathbf{U}_{j,n}\in\mathbb{C}^{2N\times(2N-\nu_j)}$ collects the eigenvectors of $\overline{\mathbf{H}}_j\,\overline{\mathbf{H}}_j^H$ associated with its $2\,N-\nu_j$ null eigenvalues. It can be argued from (B.5) that, as $\sigma_v^2\to 0$, $\text{SINR}_{j,\text{WL-MOE}}\to+\infty$ iff $\mathbf{h}_j^H\mathbf{U}_{j,n}\,\mathbf{U}_{j,n}^H\,\mathbf{h}_j\neq 0$, which implies that $\mathbf{h}_j\notin\mathcal{N}(\mathbf{U}_{j,n}^H)\equiv\mathcal{R}(\overline{\mathbf{H}}_j)$. On the other hand, when $\mathbf{h}_j$ belongs to $\mathcal{R}(\overline{\mathbf{H}}_j)$, it results that, as $\sigma_v^2\to 0$, $\text{SINR}_{j,\text{WL-MOE}}$ takes on the finite value $\mathbf{h}_j^H\mathbf{U}_{j,s}\,\boldsymbol{\Lambda}_{j,s}^{-1}\,\mathbf{U}_{j,s}^H\,\mathbf{h}_j$. Therefore, since condition $\mathbf{h}_j\notin\mathcal{R}(\overline{\mathbf{H}}_j)$ holds, $\forall j\in$

$\{1, 2, \ldots, J\}$, iff the augmented matrix $\mathbf{H} = [\mathbf{\Phi}^T, \mathbf{\Phi}^H]^T \in \mathbb{C}^{2N \times J}$ is full-column rank, we maintain that, in the absence of noise, the WL-MOE receiver is able to achieve perfect MAI suppression for *each* active user iff rank$(\mathbf{H}) = J$. The matrix $\mathbf{H}$ turns out to be full-column rank iff the null spaces of the matrices $\mathbf{\Phi}$ and $\mathbf{\Phi}^*$ intersect only trivially (see, e.g., [95]), that is, $\mathcal{N}(\mathbf{\Phi}) \cap \mathcal{N}(\mathbf{\Phi}^*) = \{\mathbf{0}_J\}$. If $\mathbf{\Phi}$ is full-column rank, which necessarily requires that $J \leq N$ (underloaded system), this condition is trivially satisfied and, hence, the augmented matrix $\mathbf{H}$ is full-column rank as well. However, the converse statement is not true, that is, $\mathbf{H}$ may be full-column rank even when $N < J \leq 2\,N$ (overloaded system). To point out a first consequence of this result, let us focus attention on the case when $N < J \leq 2\,N$. In this overloaded scenario, the matrix $\mathbf{\Phi}$ cannot be full-column rank and, thus, it results that, as $\sigma_v^2 \to 0$, SINR$_{j,\text{L-MOE}}$ takes on a finite value. In contrast, since $\mathbf{H}$ can still be full-column rank in an overloaded system, relying on the results provided before, we can infer that, if $\mathbf{H}$ is full-column rank, then

$$\lim_{\sigma_v^2 \to 0} \frac{\text{SINR}_{j,\text{WL-MOE}}}{\text{SINR}_{j,\text{L-MOE}}} = +\infty\,, \tag{B.6}$$

$\forall j \in \{1, 2, \ldots, J\}$, with $N < J \leq 2\,N$. Let us now consider an underloaded scenario ($J \leq N$) and assume that $\mathbf{\Phi}$ is full-column rank. Since in this case the matrix $\mathbf{H}$ is full-column rank, too, it follows that both SINR$_{j,\text{L-MOE}}$ and SINR$_{j,\text{WL-MOE}}$ diverge, in the limiting case of vanishingly small noise, and thus $\lim_{\sigma_v^2 \to 0} \frac{\text{SINR}_{j,\text{WL-MOE}}}{\text{SINR}_{j,\text{L-MOE}}}$ assumes an indeterminate form. To overcome this mathematical difficulty, we preliminary develop the relationship existing between SINR$_{j,\text{WL-MOE}}$ and $\overline{\text{SINR}}_{j,\text{max}}$ in the high-SNR regime, by resorting to the series expansions (B.2) and (B.5). So doing, we get:

$$\lim_{\sigma_v^2 \to 0} \frac{\text{SINR}_{j,\text{WL-MOE}}}{\overline{\text{SINR}}_{j,\text{max}}} = \frac{\mathbf{h}_j^H\,\mathbf{U}_{j,n}\,\mathbf{U}_{j,n}^H\,\mathbf{h}_j}{\boldsymbol{\phi}_j^H\,\mathbf{V}_{j,n}\,\mathbf{V}_{j,n}^H\,\boldsymbol{\phi}_j} = \frac{\|\mathbf{U}_{j,n}^H\,\mathbf{h}_j\|^2}{\|\mathbf{V}_{j,n}^H\,\boldsymbol{\phi}_j\|^2}\,, \tag{B.7}$$

where, since both $\mathbf{\Phi}$ and $\mathbf{H}$ are full-column rank, it follows that $\|\mathbf{V}_{j,n}^H\,\boldsymbol{\phi}_j\| \neq 0$ and $\|\mathbf{U}_{j,n}^H\,\mathbf{h}_j\| \neq 0$, $\forall j \in \{1, 2, \ldots, J\}$. It is worth observing that $\mathbf{V}_{j,n}\,\mathbf{V}_{j,n}^H$ and $\mathbf{U}_{j,n}\,\mathbf{U}_{j,n}^H$ represent the orthogonal projections [24] onto the subspaces $\mathcal{R}^{\perp}(\overline{\mathbf{\Phi}}_j)$ and $\mathcal{R}^{\perp}(\overline{\mathbf{H}}_j)$, respectively, which can be equivalently expressed [24] as $\mathbf{V}_{j,n}\,\mathbf{V}_{j,n}^H = \mathbf{I}_N - \overline{\mathbf{\Phi}}_j\,(\overline{\mathbf{\Phi}}_j^H\,\overline{\mathbf{\Phi}}_j)^{-1}\,\overline{\mathbf{\Phi}}_j^H$ and $\mathbf{U}_{j,n}\,\mathbf{U}_{j,n}^H = \mathbf{I}_{2N} - \overline{\mathbf{H}}_j\,(\overline{\mathbf{H}}_j^H\,\overline{\mathbf{H}}_j)^{-1}\,\overline{\mathbf{H}}_j^H$. By substituting this two relations in (B.7), and remembering that $\overline{\mathbf{H}}_j = [\overline{\mathbf{\Phi}}_j^T, \overline{\mathbf{\Phi}}_j^H]^T$ and $\mathbf{h}_j = [\boldsymbol{\phi}_j^T, \boldsymbol{\phi}_j^H]^T$, after some algebraic

manipulations, one has:

$$\lim_{\sigma_v^2 \to 0} \frac{\mathrm{SINR}_{j,\text{WL-MOE}}}{\overline{\mathrm{SINR}}_{j,\max}} = 2 \cdot \frac{\|\boldsymbol{\phi}_j\|^2 - \mathrm{Re}[\boldsymbol{\phi}_j^H \overline{\boldsymbol{\Phi}}_j] \, \{\mathrm{Re}[\overline{\boldsymbol{\Phi}}_j^H \overline{\boldsymbol{\Phi}}_j]\}^{-1} \, \mathrm{Re}[\overline{\boldsymbol{\Phi}}_j^H \boldsymbol{\phi}_j]}{\|\boldsymbol{\phi}_j\|^2 - \boldsymbol{\phi}_j^H \overline{\boldsymbol{\Phi}}_j \, (\overline{\boldsymbol{\Phi}}_j^H \overline{\boldsymbol{\Phi}}_j)^{-1} \, \overline{\boldsymbol{\Phi}}_j^H \boldsymbol{\phi}_j} \, . \tag{B.8}$$

Therefore, if $\boldsymbol{\Phi}$ is full-column rank, accounting for (B.4) and (B.8), we can state that:

$$\begin{aligned} \lim_{\sigma_v^2 \to 0} \frac{\mathrm{SINR}_{j,\text{WL-MOE}}}{\mathrm{SINR}_{j,\text{L-MOE}}} &= \lim_{\sigma_v^2 \to 0} \frac{\mathrm{SINR}_{j,\text{WL-MOE}}}{\overline{\mathrm{SINR}}_{j,\max}} \cdot \lim_{\sigma_v^2 \to 0} \frac{\overline{\mathrm{SINR}}_{j,\max}}{\mathrm{SINR}_{j,\text{L-MOE}}} \\ &= \frac{\|\boldsymbol{\phi}_j\|^2 - \mathrm{Re}[\boldsymbol{\phi}_j^H \overline{\boldsymbol{\Phi}}_j] \, \{\mathrm{Re}[\overline{\boldsymbol{\Phi}}_j^H \overline{\boldsymbol{\Phi}}_j]\}^{-1} \, \mathrm{Re}[\overline{\boldsymbol{\Phi}}_j^H \boldsymbol{\phi}_j]}{\|\boldsymbol{\phi}_j\|^2 - \boldsymbol{\phi}_j^H \overline{\boldsymbol{\Phi}}_j \, (\overline{\boldsymbol{\Phi}}_j^H \overline{\boldsymbol{\Phi}}_j)^{-1} \, \overline{\boldsymbol{\Phi}}_j^H \boldsymbol{\phi}_j} \, . \end{aligned} \tag{B.9}$$

B.3 Proof of Theorem 4.1

Accounting for (4.52) and (4.54), and exploiting the fact that $\boldsymbol{\Theta}^* \, \boldsymbol{\Theta} = \mathbf{I}_N$, one has:

$$\mathbf{H} = \begin{bmatrix} \mathbf{G} & \mathbf{O}_{N \times N} \\ \mathbf{O}_{N \times N} & \mathbf{G}^* \end{bmatrix} \begin{bmatrix} \mathbf{C} \\ \mathbf{C}^* \, (\boldsymbol{\Theta}^2)^* \end{bmatrix} \boldsymbol{\Theta} \, \mathbf{A} \, , \tag{B.10}$$

which, as a consequence of the nonsingularity of matrices $\mathbf{G}$, $\mathbf{A}$ and $\boldsymbol{\Theta}$, implies that $\mathrm{rank}(\mathbf{H}) = \mathrm{rank}([\mathbf{C}^T, (\mathbf{C} \, \boldsymbol{\Theta}^2)^H]^T)$. In its turn, the matrix $[\mathbf{C}^T, (\mathbf{C} \, \boldsymbol{\Theta}^2)^H]^T \in \mathbb{C}^{2N \times J}$ is full-column rank iff $\mathcal{N}(\mathbf{C}) \cap \mathcal{N}[\mathbf{C}^* \, (\boldsymbol{\Theta}^2)^*] = \{\mathbf{0}_J\}$. At this point, let us characterize the null spaces of $\mathbf{C}$ and $\mathbf{C}^* \, (\boldsymbol{\Theta}^2)^*$, when $N < J \leq 2 \, N$. In this overloaded case, by assuming without loss of generality that the first N column $\mathbf{c}_1$, $\mathbf{c}_2, \ldots, \mathbf{c}_N$ of $\mathbf{C}$ are linearly independent, its remaining $J - N$ columns $\mathbf{c}_{N+1}$, $\mathbf{c}_{N+2}, \ldots, \mathbf{c}_J$ can be expressed as a linear combination of the first N ones, thus obtaining the following decomposition $\mathbf{C} = \mathbf{C}_{\text{left}} \, [\mathbf{I}_N, \boldsymbol{\Pi}]$, where $\mathbf{C}_{\text{left}} \stackrel{\triangle}{=} [\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_N] \in \mathbb{C}^{N \times N}$ is nonsingular and $\boldsymbol{\Pi} \in \mathbb{C}^{N \times (J-N)}$ is a tall matrix. Due to nonsingularity of $\mathbf{C}_{\text{left}}$, it follows that $\mathcal{N}(\mathbf{C}) = \mathcal{N}([\mathbf{I}_N, \boldsymbol{\Pi}])$. Hence, it can be verified that the general forms of a vector $\boldsymbol{\alpha}_1 \in \mathbb{C}^J$ belonging to $\mathcal{N}(\mathbf{C})$ and a vector $\boldsymbol{\alpha}_2 \in \mathbb{C}^J$ belonging to $\mathcal{N}[\mathbf{C}^* \, (\boldsymbol{\Theta}^2)^*]$ are given by

$$\boldsymbol{\alpha}_1 = \begin{bmatrix} -\boldsymbol{\Pi} \\ \mathbf{I}_{J-N} \end{bmatrix} \boldsymbol{\vartheta}_1 \quad \text{and} \quad \boldsymbol{\alpha}_2 = \boldsymbol{\Theta}^2 \begin{bmatrix} -\boldsymbol{\Pi}^* \\ \mathbf{I}_{J-N} \end{bmatrix} \boldsymbol{\vartheta}_2 \, , \tag{B.11}$$

with arbitrary $\boldsymbol{\vartheta}_1, \boldsymbol{\vartheta}_2 \in \mathbb{C}^{J-N}$. By virtue of (B.11), the matrix $\mathbf{H}$ is not full-column rank iff there exist at least two nonzero vectors $\boldsymbol{\vartheta}_1$ and

$\boldsymbol{\vartheta}_2$ such that $\boldsymbol{\alpha}_1 = \boldsymbol{\alpha}_2$, which amounts to $\boldsymbol{\Pi}\,\boldsymbol{\vartheta}_1 = \boldsymbol{\Theta}_1^2\,\boldsymbol{\Pi}^*\boldsymbol{\vartheta}_2$ and $\boldsymbol{\vartheta}_1 = \boldsymbol{\Theta}_2^2\,\boldsymbol{\vartheta}_2$, with $\boldsymbol{\Theta}_1 \stackrel{\triangle}{=} \mathrm{diag}(e^{i\,\theta_1}, e^{i\,\theta_2}, \ldots, e^{i\,\theta_N}) \in \mathbb{C}^{N\times N}$ and $\boldsymbol{\Theta}_2 \stackrel{\triangle}{=} \mathrm{diag}(e^{i\,\theta_{N+1}}, e^{i\,\theta_{N+2}}, \ldots, e^{i\,\theta_J}) \in \mathbb{C}^{(J-N)\times(J-N)}$. By substituting the second relation in the first one and observing that $\boldsymbol{\Theta}_1^2$ is nonsingular, one obtains $[\boldsymbol{\Pi}^* - (\boldsymbol{\Theta}_1^2)^*\,\boldsymbol{\Pi}\,\boldsymbol{\Theta}_2^2]\,\boldsymbol{\vartheta}_2 = \mathbf{0}_N$, which shows that, if the matrix $\boldsymbol{\Pi}^* - (\boldsymbol{\Theta}_1^2)^*\,\boldsymbol{\Pi}\,\boldsymbol{\Theta}_2^2 \in \mathbb{C}^{N\times(J-N)}$ is full-column rank, then $\boldsymbol{\alpha}_1 = \boldsymbol{\alpha}_2$ is satisfied iff $\boldsymbol{\vartheta}_1 = \boldsymbol{\vartheta}_2 = \mathbf{0}_{J-N}$. This assures that $\mathrm{rank}(\mathbf{H}) = J$, since it means that $\mathcal{N}(\mathbf{C}) \cap \mathcal{N}[\mathbf{C}^*\,(\boldsymbol{\Theta}^2)^*] = \{\mathbf{0}_J\}$.

B.4 Proof of Lemma 4.2

First, let us consider the SMI implementation of the WL-MOE receiver. By substituting (4.20) in (4.63), the sample autocorrelation matrix $\widehat{\mathbf{R}}_{\mathbf{zz}}$ of the augmented vector $\mathbf{z}(k)$ can be expressed as

$$\widehat{\mathbf{R}}_{\mathbf{zz}} = \mathbf{h}_j\,\mathbf{h}_j^H + \mathbf{h}_j\,\widehat{\mathbf{r}}_{\mathbf{q}_j b_j}^H + \widehat{\mathbf{r}}_{\mathbf{q}_j b_j}\mathbf{h}_j^H + \widehat{\mathbf{R}}_{\mathbf{q}_j\mathbf{q}_j}\,, \tag{B.12}$$

where $\widehat{\mathbf{r}}_{\mathbf{q}_j b_j} \stackrel{\triangle}{=} \frac{1}{K}\sum_{k=0}^{K-1}\mathbf{q}_j(k)\,b_j(k)$ and $\widehat{\mathbf{R}}_{\mathbf{q}_j\mathbf{q}_j} \stackrel{\triangle}{=} \frac{1}{K}\sum_{k=0}^{K-1}\mathbf{q}_j(k)\,\mathbf{q}_j^H(k)$ represent sample estimates of the cross-correlation between the disturbance vector $\mathbf{q}_j(k)$ and the desired symbol $b_j(k)$, and the autocorrelation matrix of $\mathbf{q}_j(k)$, respectively. It is shown in [75] that, for moderate-to-high values of the sample size, i.e., $K \geq 6\,N$, the predominant cause of SINR degradation is represented by $\widehat{\mathbf{r}}_{\mathbf{q}_j b_j}$ and, thus, replacing $\widehat{\mathbf{R}}_{\mathbf{q}_j\mathbf{q}_j}$ with $\mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}$ in (B.12) has a very marginal effect on the SINR. Therefore, remembering that $\mathbf{R}_{\mathbf{zz}} = \mathbf{h}_j\mathbf{h}_j^H + \mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}$, eq. (B.12) can be rewritten as $\widehat{\mathbf{R}}_{\mathbf{zz}} = \mathbf{R}_{\mathbf{zz}} + \mathbf{h}_j\,\widehat{\mathbf{r}}_{\mathbf{q}_j b_j}^H + \widehat{\mathbf{r}}_{\mathbf{q}_j b_j}\mathbf{h}_j^H$. Its inverse admits [24] the following first-order approximation $\widehat{\mathbf{R}}_{\mathbf{zz}}^{-1} \approx \mathbf{R}_{\mathbf{zz}}^{-1} - \mathbf{R}_{\mathbf{zz}}^{-1}(\mathbf{h}_j\,\widehat{\mathbf{r}}_{\mathbf{q}_j b_j}^H + \widehat{\mathbf{r}}_{\mathbf{q}_j b_j}\mathbf{h}_j^H)\,\mathbf{R}_{\mathbf{zz}}^{-1}$, which can be substituted in (4.65), thus obtaining

$$\mathbf{f}_{j,\text{WL-SMI}} \approx \mathbf{f}_{j,\text{WL-MOE}} \underbrace{-\mathbf{P}_{j,\text{WL}}\,\mathbf{R}_{\mathbf{zz}}^{-1}\,\widehat{\mathbf{r}}_{\mathbf{q}_j b_j}}_{\boldsymbol{\delta}\mathbf{f}_{j,\text{WL-SMI}}} = \mathbf{f}_{j,\text{WL-MOE}} + \boldsymbol{\delta}\mathbf{f}_{j,\text{WL-SMI}}\,, \tag{B.13}$$

with $\mathbf{P}_{j,\text{WL}} \stackrel{\triangle}{=} \mathbf{I}_{2N} - (\mathbf{h}_j^H\,\mathbf{R}_{\mathbf{zz}}^{-1}\,\mathbf{h}_j)^{-1}\,\mathbf{R}_{\mathbf{zz}}^{-1}\,\mathbf{h}_j\,\mathbf{h}_j^H = \mathbf{I}_{2N} - (\mathbf{h}_j^H\,\mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}^{-1}\,\mathbf{h}_j)^{-1}\,\mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}^{-1}\,\mathbf{h}_j\,\mathbf{h}_j^H \in \mathbb{C}^{N\times N}$, where here and in the sequel the symbol $\approx$ denotes *first-order equality*, i.e., we neglect all the summands that tend to zero, as the sample size K approaches infinity, faster than the norm of the corresponding perturbation term. It is easily verified that $\mathbf{P}_{j,\text{WL}}\,\mathbf{R}_{\mathbf{zz}}^{-1} = \mathbf{P}_{j,\text{WL}}\,\mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}^{-1}$.

At this point, we focus attention on the subspace implementation of the WL-MOE receiver. Preliminary, we recall that the EVD of $\widehat{\mathbf{R}}_{\mathbf{zz}}$ is given by (4.64) that here is reported for simplicity

$$\widehat{\mathbf{R}}_{\mathbf{zz}} = \widehat{\mathbf{U}}_s \, \widehat{\mathbf{\Lambda}}_s \, \widehat{\mathbf{U}}_s^H + \widehat{\mathbf{U}}_n \, \widehat{\mathbf{\Lambda}}_n \, \widehat{\mathbf{U}}_n^H \, , \tag{B.14}$$

where $\widehat{\mathbf{U}}_s$, $\widehat{\mathbf{\Lambda}}_s$, $\widehat{\mathbf{U}}_n$ and $\widehat{\mathbf{\Lambda}}_n$ are sample estimates of $\mathbf{U}_s$, $\mathbf{\Lambda}_s$, $\mathbf{U}_n$ and $\sigma_v^2 \, \mathbf{I}_N$, respectively. When $\mathbf{R}_{\mathbf{zz}}$ is estimated from the received data as in (4.63), for a sufficiently large sample size K, the estimate can be decomposed as $\widehat{\mathbf{R}}_{\mathbf{zz}} = \mathbf{R}_{\mathbf{zz}} + \boldsymbol{\delta}\mathbf{R}_{\mathbf{zz}}$, where $\boldsymbol{\delta}\mathbf{R}_{\mathbf{zz}}$ is a *small* additive perturbation (in the Frobenius norm sense). Consequently, the matrices $\widehat{\mathbf{U}}_s$ and $\widehat{\mathbf{\Lambda}}_s$ can be written [77, 76] as $\widehat{\mathbf{U}}_s = \mathbf{U}_s + \boldsymbol{\delta}\mathbf{U}_s$ and $\widehat{\mathbf{\Lambda}}_s = \mathbf{\Lambda}_s + \boldsymbol{\delta}\mathbf{\Lambda}_s$, where $\boldsymbol{\delta}\mathbf{U}_s$ and $\boldsymbol{\delta}\mathbf{\Lambda}_s$ represent the resulting perturbation in the estimated signal subspace, whose norm is of the order of $\|\boldsymbol{\delta}\mathbf{R}_{\mathbf{zz}}\|$. It results [77, 76] that $\boldsymbol{\delta}\mathbf{U}_s \approx \mathbf{U}_n \mathbf{U}_n^H \boldsymbol{\delta}\mathbf{R}_{\mathbf{zz}} \, \mathbf{\Omega}_{\text{WL}}^{-1}$, with $\mathbf{\Omega}_{\text{WL}} \stackrel{\triangle}{=} \mathbf{\Lambda}_{\mathbf{s}} - \sigma_v^2 \, \mathbf{I}_J$, and $\boldsymbol{\delta}\mathbf{\Lambda}_s \approx \mathbf{U}_s^H \boldsymbol{\delta}\mathbf{R}_{\mathbf{zz}} \, \mathbf{U}_s$. By substituting the above expressions of $\widehat{\mathbf{U}}_s$ and $\widehat{\mathbf{\Lambda}}_s$ in (4.66), and remembering that $\mathbf{P}_{j,\text{WL}} \, \mathbf{R}_{\mathbf{zz}}^{-1} = \mathbf{P}_{j,\text{WL}} \, \mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}^{-1}$, we get:

$$\begin{aligned} \mathbf{f}_{j,\text{WL-SUB}} &\approx \mathbf{f}_{j,\text{WL-MOE}} \underbrace{- \left(\mathbf{P}_{j,\text{WL}} \, \mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}^{-1} - \gamma_{j,\text{WL}} \, \mathbf{U}_n \mathbf{U}_n^H \right) \widehat{\mathbf{r}}_{\mathbf{q}_j b_j}}_{\delta \mathbf{f}_{j,\text{WL-SUB}}} \\ &= \mathbf{f}_{j,\text{WL-MOE}} + \boldsymbol{\delta}\mathbf{f}_{j,\text{WL-SUB}} \, , \end{aligned} \tag{B.15}$$

where $\gamma_{j,\text{WL}} \stackrel{\triangle}{=} \sigma_v^{-2} + (\mathbf{h}_j^H \, \mathbf{R}_{\mathbf{zz}}^{-1} \, \mathbf{h}_j)^{-1} \mathbf{h}_j^H \mathbf{U}_s \, \mathbf{\Omega}_{\text{WL}}^{-1} \mathbf{U}_s^H \mathbf{R}_{\mathbf{zz}}^{-1} \mathbf{h}_j$.

B.5 Proof of Lemma 4.4

It is shown in the proof of the Lemma 4.2 in B.4 that, for moderate-to-high values of the sample size, i.e., $K \geq 6\,N$, the sample ACM given by (4.63) can be decomposed as $\widehat{\mathbf{R}}_{\mathbf{zz}} = \mathbf{R}_{\mathbf{zz}} + \boldsymbol{\delta}\mathbf{R}_{\mathbf{zz}}$, where $\boldsymbol{\delta}\mathbf{R}_{\mathbf{zz}} \stackrel{\triangle}{=} \mathbf{h}_j \, \widehat{\mathbf{r}}_{\mathbf{q}_j b_j}^H + \widehat{\mathbf{r}}_{\mathbf{q}_j b_j} \mathbf{h}_j^H \in \mathbb{C}^{2N \times 2N}$, with $\widehat{\mathbf{r}}_{\mathbf{q}_j b_j} \stackrel{\triangle}{=} \frac{1}{K} \sum_{k=0}^{K-1} \mathbf{q}_j(k) \, b_j(k)$. Consequently, the inverse of the sample ACM admits the first-order approximation $\widehat{\mathbf{R}}_{\mathbf{zz}}^{-1} \approx \mathbf{R}_{\mathbf{zz}}^{-1} - \mathbf{R}_{\mathbf{zz}}^{-1} \, \boldsymbol{\delta}\mathbf{R}_{\mathbf{zz}} \, \mathbf{R}_{\mathbf{zz}}^{-1}$, where in the sequel the symbol $\approx$ denotes first-order equality.

First, let us consider the SMI-CE implementation (4.109) of the WL-MOE receiver. Substituting the previous approximation of $\widehat{\mathbf{R}}_{\mathbf{zz}}^{-1}$ and $\widehat{\mathbf{h}}_j = \mathbf{h}_j + \boldsymbol{\delta}\mathbf{h}_j$ in (4.109), after some algebraic manipulations, one obtains the first-order

approximation of the weight vector

$$\begin{aligned}\mathbf{f}_{j,\text{WL-SMI-CE}} \approx \mathbf{f}_{j,\text{WL-MOE}} &\underbrace{-\mathbf{P}_{j,\text{WL}}\,\mathbf{R}_{\mathbf{zz}}^{-1}\,\boldsymbol{\delta}\mathbf{R}_{\mathbf{zz}}\,\mathbf{f}_{j,\text{WL-MOE}}}_{\boldsymbol{\delta}\mathbf{f}_{j,\text{WL-SMI-CE}}^{(1)}} \\ &+ \underbrace{(\mathbf{h}_j^H\,\mathbf{R}_{\mathbf{zz}}^{-1}\,\mathbf{h}_j)^{-1}\,\mathbf{R}_{\mathbf{zz}}^{-1}\,\boldsymbol{\delta}\mathbf{h}_j - 2\,\text{Re}(\mathbf{f}_{j,\text{WL-MOE}}^H\,\boldsymbol{\delta}\mathbf{h}_j)\,\mathbf{f}_{j,\text{WL-MOE}}}_{\boldsymbol{\delta}\mathbf{f}_{j,\text{WL-SMI-CE}}^{(2)}} \\ &= \mathbf{f}_{j,\text{WL-MOE}} + \boldsymbol{\delta}\mathbf{f}_{j,\text{WL-SMI-CE}}^{(1)} + \boldsymbol{\delta}\mathbf{f}_{j,\text{WL-SMI-CE}}^{(2)}\,, \end{aligned} \tag{B.16}$$

with $\mathbf{P}_{j,\text{WL}} \overset{\triangle}{=} \mathbf{I}_{2N} - (\mathbf{h}_j^H\,\mathbf{R}_{\mathbf{zz}}^{-1}\,\mathbf{h}_j)^{-1}\,\mathbf{R}_{\mathbf{zz}}^{-1}\,\mathbf{h}_j\,\mathbf{h}_j^H = \mathbf{I}_{2N} - (\mathbf{h}_j^H\,\mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}^{-1}\,\mathbf{h}_j)^{-1}\,\mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}^{-1}\,\mathbf{h}_j\,\mathbf{h}_j^H \in \mathbb{C}^{2N\times 2N}$. Observe that, taking into account (4.27), the matrix $\mathbf{P}_{j,\text{WL}}$ can be equivalently expressed as $\mathbf{P}_{j,\text{WL}} = \mathbf{I}_{2N} - \mathbf{f}_j\,\mathbf{h}_j^H$. Substituting the expression of $\boldsymbol{\delta}\mathbf{R}_{\mathbf{zz}}$ in $\boldsymbol{\delta}\mathbf{f}_{j,\text{WL-SMI-CE}}^{(1)}$, and observing that $\mathbf{P}_{j,\text{WL}}\,\mathbf{R}_{\mathbf{zz}}^{-1}\,\mathbf{h}_j = \mathbf{0}_{2N}$, $\mathbf{h}_j^H\mathbf{f}_{j,\text{WL-MOE}} = 1$ and $\mathbf{P}_{j,\text{WL}}\,\mathbf{R}_{\mathbf{zz}}^{-1} = \mathbf{P}_{j,\text{WL}}\,\mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}^{-1}$, one has

$$\delta\mathbf{f}_{j,\text{WL-SMI-CE}}^{(1)} = -\underbrace{\mathbf{P}_{j,\text{WL}}\,\mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}^{-1}}_{\boldsymbol{\Gamma}_{j,\text{WL}}\in\mathbb{C}^{2N\times 2N}}\,\widehat{\mathbf{r}}_{\mathbf{q}_j b_j} = -\boldsymbol{\Gamma}_{j,\text{WL}}\,\widehat{\mathbf{r}}_{\mathbf{q}_j b_j}\,. \tag{B.17}$$

Since both $\mathbf{f}_{j,\text{WL-MOE}}^H$ and $\boldsymbol{\delta}\mathbf{h}_j$ exhibit the CS property, the scalar $\mathbf{f}_{j,\text{WL-MOE}}^H\,\boldsymbol{\delta}\mathbf{h}_j$ is real and, thus, $\text{Re}(\mathbf{f}_{j,\text{WL-MOE}}^H\,\boldsymbol{\delta}\mathbf{h}_j)\,\mathbf{f}_{j,\text{WL-MOE}} = (\mathbf{f}_{j,\text{WL-MOE}}^H\,\boldsymbol{\delta}\mathbf{h}_j)\,\mathbf{f}_{j,\text{WL-MOE}} = (\mathbf{f}_{j,\text{WL-MOE}}\,\mathbf{f}_{j,\text{WL-MOE}}^H)\,\boldsymbol{\delta}\mathbf{h}_j$. Consequently,

$$\begin{aligned}\delta\mathbf{f}_{j,\text{WL-SMI-CE}}^{(2)} &= \underbrace{\left[(\mathbf{h}_j^H\,\mathbf{R}_{\mathbf{zz}}^{-1}\,\mathbf{h}_j)^{-1}\,\mathbf{R}_{\mathbf{zz}}^{-1} - 2\,\mathbf{f}_{j,\text{WL-MOE}}\,\mathbf{f}_{j,\text{WL-MOE}}^H\right]}_{\boldsymbol{\Delta}_{j,\text{WL}}\in\mathbb{C}^{2N\times 2N}}\,\boldsymbol{\delta}\mathbf{h}_j \\ &= \boldsymbol{\Delta}_{j,\text{WL}}\,\boldsymbol{\delta}\mathbf{h}_j\,. \end{aligned} \tag{B.18}$$

At this point, we focus attention on the SUB-CE implementation (4.110) of the WL-MOE receiver. When the EVD is applied to the sample ACM $\widehat{\mathbf{R}}_{\mathbf{zz}}$ given by (4.63), for a sufficiently large sample size K, the matrices $\widehat{\mathbf{U}}_s$ and $\widehat{\boldsymbol{\Lambda}}_s$ can be decomposed [76, 77] as $\widehat{\mathbf{U}}_s = \mathbf{U}_s + \boldsymbol{\delta}\mathbf{U}_s$ and $\widehat{\boldsymbol{\Lambda}}_s = \boldsymbol{\Lambda}_s + \boldsymbol{\delta}\boldsymbol{\Lambda}_s$, where $\boldsymbol{\delta}\mathbf{U}_s$ and $\boldsymbol{\delta}\boldsymbol{\Lambda}_s$ represent the resulting perturbation in the estimated signal subspace, whose norm is of the order of $\|\boldsymbol{\delta}\mathbf{R}_{\mathbf{zz}}\|$. Moreover, it results [76, 77] that $\boldsymbol{\delta}\mathbf{U}_s \approx \mathbf{U}_n\mathbf{U}_n^H\,\boldsymbol{\delta}\mathbf{R}_{\mathbf{zz}}\,\mathbf{U}_s\,\boldsymbol{\Omega}_{\text{WL}}^{-1}$, with $\boldsymbol{\Omega}_{\text{WL}} \overset{\triangle}{=} \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_J) \in$

$\mathbb{R}^{J\times J}$, and $\boldsymbol{\delta\Lambda}_s \approx \mathbf{U}_s^H \boldsymbol{\delta}\mathbf{R}_{\mathbf{zz}}\,\mathbf{U}_s$. Consequently, we can write

$$\widehat{\mathbf{U}}_s\,\widehat{\boldsymbol{\Lambda}}_s^{-1}\,\widehat{\mathbf{U}}_s^H \approx \mathbf{U}_s\,\boldsymbol{\Lambda}_s^{-1}\,\mathbf{U}_s^H + \mathbf{U}_s\,\boldsymbol{\Lambda}_s^{-1}\,\boldsymbol{\delta}\mathbf{U}_s^H + \\ - \mathbf{U}_s\,\boldsymbol{\Lambda}_s^{-1}\,\boldsymbol{\delta\Lambda}_s\,\boldsymbol{\Lambda}_s^{-1}\,\mathbf{U}_s^H + \boldsymbol{\delta}\mathbf{U}_s\,\boldsymbol{\Lambda}_s^{-1}\,\mathbf{U}_s^H\,. \quad \text{(B.19)}$$

Observe that, since $\mathbf{U}_n^H\,\mathbf{h}_j = \mathbf{0}_{2N-J}$, one has $\boldsymbol{\delta}\mathbf{U}_s^H\,\mathbf{h}_j = \mathbf{0}_J$. Hence, using (B.19), accounting for the first-order perturbations of $\mathbf{U}_s$ and $\boldsymbol{\Lambda}_s$, and remembering that $\widehat{\mathbf{h}}_j = \mathbf{h}_j + \boldsymbol{\delta}\mathbf{h}_j$, one obtains

$$\widehat{\mathbf{h}}_j^H\,\widehat{\mathbf{U}}_s\,\widehat{\boldsymbol{\Lambda}}_s^{-1}\,\widehat{\mathbf{U}}_s^H\widehat{\mathbf{h}}_j \approx \mathbf{h}_j^H\,\mathbf{U}_s\,\boldsymbol{\Lambda}_s^{-1}\,\mathbf{U}_s^H\,\mathbf{h}_j - \mathbf{h}_j^H\,\mathbf{U}_s\,\boldsymbol{\Lambda}_s^{-1}\,\boldsymbol{\delta\Lambda}_s\,\boldsymbol{\Lambda}_s^{-1}\,\mathbf{U}_s^H\,\mathbf{h}_j \\ + 2\,\mathrm{Re}[\mathbf{h}_j^H\,\mathbf{U}_s\,\boldsymbol{\Lambda}_s^{-1}\,\mathbf{U}_s^H\,\boldsymbol{\delta}\mathbf{h}_j]\,. \quad \text{(B.20)}$$

Substituting (B.19) and (B.20) in (4.110), after some tedious but straightforward algebra, the first-order approximation of the weight vector can be concisely written as

$$\mathbf{f}_{j,\text{WL-SUB-CE}} \approx \mathbf{f}_{j,\text{WL-MOE}} + \boldsymbol{\delta}\mathbf{f}_{j,\text{WL-SUB-CE}}^{(1)} + \boldsymbol{\delta}\mathbf{f}_{j,\text{WL-SUB-CE}}^{(2)}\,, \quad \text{(B.21)}$$

where

$$\boldsymbol{\delta}\mathbf{f}_{j,\text{WL-SUB-CE}}^{(1)} \triangleq -\{\mathbf{P}_{j,\text{WL}}\mathbf{R}_{\mathbf{zz}}^{-1}\boldsymbol{\delta}\mathbf{R}_{\mathbf{zz}} + \\ - \mathbf{U}_n\,\mathbf{U}_n^H\boldsymbol{\delta}\mathbf{R}_{\mathbf{zz}}\left[\sigma_v^{-2}\,\mathbf{I}_{2N} + \mathbf{U}_s\,\boldsymbol{\Omega}_{\text{WL}}^{-1}\,\mathbf{U}_s^H\right]\}\mathbf{f}_{j,\text{WL-MOE}}, \quad \text{(B.22)}$$

$$\boldsymbol{\delta}\mathbf{f}_{j,\text{WL-SUB-CE}}^{(2)} \triangleq (\mathbf{h}_j^H\mathbf{R}_{\mathbf{zz}}^{-1}\,\mathbf{h}_j)^{-1}\left[\mathbf{R}_{\mathbf{zz}}^{-1} - \sigma_v^{-2}\,\mathbf{U}_n\,\mathbf{U}_n^H\right] + \boldsymbol{\delta}\mathbf{h}_j \\ - 2\,\mathrm{Re}(\mathbf{f}_{j,\text{WL-MOE}}^H\,\boldsymbol{\delta}\mathbf{h}_j)\,\mathbf{f}_{j,\text{WL-MOE}}\,. \quad \text{(B.23)}$$

Then, substituting the expression of the perturbation $\boldsymbol{\delta}\mathbf{R}_{\mathbf{zz}}$ in (B.22), remembering again that $\mathbf{U}_n^H\,\mathbf{h}_j = \mathbf{0}_{2N-J}$, $\mathbf{h}_j^H\mathbf{f}_{j,\text{WL-MOE}} = 1$, $\mathbf{P}_{j,\text{WL}}\,\mathbf{R}_{\mathbf{zz}}^{-1}\,\mathbf{h}_j = \mathbf{0}_{2N}$ and $\mathbf{P}_{j,\text{WL}}\,\mathbf{R}_{\mathbf{zz}}^{-1} = \mathbf{P}_{j,\text{WL}}\,\mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}^{-1}$, one gets

$$\boldsymbol{\delta}\mathbf{f}_{j,\text{WL-SUB-CE}}^{(1)} = -\underbrace{\left(\mathbf{P}_{j,\text{WL}}\,\mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}^{-1} - \gamma_{j,\text{WL}}\,\mathbf{U}_n\mathbf{U}_n^H\right)}_{\boldsymbol{\Gamma}_{j,\text{WL}}\in\mathbb{C}^{2N\times 2N}}\widehat{\mathbf{r}}_{\mathbf{q}_j b_j} = -\boldsymbol{\Gamma}_{j,\text{WL}}\,\widehat{\mathbf{r}}_{\mathbf{q}_j b_j}\,, \quad \text{(B.24)}$$

where $\gamma_{j,\text{WL}} \triangleq \sigma_v^{-2} + (\mathbf{h}_j^H\,\mathbf{R}_{\mathbf{zz}}^{-1}\,\mathbf{h}_j)^{-1}\mathbf{h}_j^H\mathbf{U}_s\,\boldsymbol{\Omega}_{\text{WL}}^{-1}\,\mathbf{U}_s^H\,\mathbf{R}_{\mathbf{zz}}^{-1}\,\mathbf{h}_j$. Moreover, using again the fact that $\mathrm{Re}(\mathbf{f}_{j,\text{WL-MOE}}^H\,\boldsymbol{\delta}\mathbf{h}_j)\,\mathbf{f}_{j,\text{WL-MOE}} =$

$(\mathbf{f}_{j,\text{WL-MOE}}\,\mathbf{f}_{j,\text{WL-MOE}}^H)\,\boldsymbol{\delta}\mathbf{h}_j$ and observing that, by virtue of the EVD properties, $\mathbf{R}_{\mathbf{zz}}^{-1} - \sigma_v^{-2}\,\mathbf{U}_n\,\mathbf{U}_n^H = \mathbf{U}_s\,\boldsymbol{\Lambda}_s^{-1}\,\mathbf{U}_s^H$, the perturbation term (B.23) can be rewritten as

$$\begin{aligned}\boldsymbol{\delta}\mathbf{f}_{j,\text{WL-SUB-CE}}^{(2)} &= \underbrace{\left[(\mathbf{h}_j^H\,\mathbf{R}_{\mathbf{zz}}^{-1}\,\mathbf{h}_j)^{-1}\,\mathbf{U}_s\,\boldsymbol{\Lambda}_s^{-1}\,\mathbf{U}_s^H - 2\,\mathbf{f}_{j,\text{WL-MOE}}\,\mathbf{f}_{j,\text{WL-MOE}}^H\right]}_{\boldsymbol{\Delta}_{j,\text{WL}}\in\mathbb{C}^{2N\times 2N}}\,\boldsymbol{\delta}\mathbf{h}_j \\ &= \boldsymbol{\Delta}_{j,\text{WL}}\,\boldsymbol{\delta}\mathbf{h}_j\,. \end{aligned} \tag{B.25}$$

B.6 Proof of Lemma 4.5

For a sufficiently large sample size K, when the EVD is applied to $\widehat{\mathbf{R}}_{\mathbf{zz}} \overset{\triangle}{=} \mathbf{R}_{\mathbf{zz}} + \boldsymbol{\delta}\mathbf{R}_{\mathbf{zz}}$, where $\boldsymbol{\delta}\mathbf{R}_{\mathbf{zz}} = \mathbf{h}_j\,\widehat{\mathbf{r}}_{\mathbf{q}_j b_j}^H + \widehat{\mathbf{r}}_{\mathbf{q}_j b_j}\mathbf{h}_j^H \in \mathbb{C}^{2N\times 2N}$, with $\widehat{\mathbf{r}}_{\mathbf{q}_j b_j} \overset{\triangle}{=} \frac{1}{K}\sum_{k=0}^{K-1}\mathbf{q}_j(k)\,b_j(k)$, the matrix $\widehat{\mathbf{U}}_n$ can be decomposed [76, 77] as $\widehat{\mathbf{U}}_n = \mathbf{U}_n + \boldsymbol{\delta}\mathbf{U}_n$ and the perturbation in the estimated noise subspace has the following form $\boldsymbol{\delta}\mathbf{U}_n \approx -\mathbf{U}_s\,\boldsymbol{\Omega}_{\text{WL}}^{-1}\,\mathbf{U}_s^H\boldsymbol{\delta}\mathbf{R}_{\mathbf{zz}}\,\mathbf{U}_n$, with $\boldsymbol{\Omega}_{\text{WL}} \overset{\triangle}{=} \text{diag}(\lambda_1,\lambda_2,\ldots,\lambda_J) \in \mathbb{R}^{J\times J}$. By substituting the expression of $\boldsymbol{\delta}\mathbf{R}_{\mathbf{zz}}$ and noticing that $\mathbf{U}_n^H\,\mathbf{h}_j = \mathbf{0}_{2N-J}$, one obtains

$$\boldsymbol{\delta}\mathbf{U}_n \approx -\mathbf{U}_s\,\boldsymbol{\Omega}_{\text{WL}}^{-1}\,\mathbf{U}_s^H\,\mathbf{h}_j\,\widehat{\mathbf{r}}_{\mathbf{q}_j b_j}^H\,\mathbf{U}_n\,. \tag{B.26}$$

The perturbation $\boldsymbol{\delta}\mathbf{U}_n$ implies an error in the channel estimate $\widehat{\boldsymbol{\varrho}}_j$ given by (4.107), which assumes the form $\widehat{\boldsymbol{\varrho}}_j = \boldsymbol{\varrho}_j + \boldsymbol{\delta}\boldsymbol{\varrho}_j$, where $\boldsymbol{\delta}\boldsymbol{\varrho}_j$ represents the CE error. Remembering that $\widehat{\mathbf{h}}_j = \mathbf{h}_j + \boldsymbol{\delta}\mathbf{h}_j = \widetilde{\alpha}_j\,\boldsymbol{\mathcal{C}}_j\,\mathbf{T}_j\,\widehat{\boldsymbol{\varrho}}_j$ is the estimate of the signature $\mathbf{h}_j = \widetilde{\alpha}_j\,\boldsymbol{\mathcal{C}}_j\,\mathbf{T}_j\,\boldsymbol{\varrho}_j$, one easily gets $\boldsymbol{\delta}\mathbf{h}_j = \widetilde{\alpha}_j\,\boldsymbol{\mathcal{C}}_j\,\mathbf{T}_j\,\boldsymbol{\delta}\boldsymbol{\varrho}_j$. According to (4.47), the channel vector $\boldsymbol{\varrho}_j$ is the unique eigenvector corresponding to the null eigenvalue of $\mathbf{T}_j^H\,\boldsymbol{\mathcal{Q}}_{j,\text{WL}}\,\mathbf{T}_j \in \mathbb{C}^{2L_j\times 2L_j}$, with $\boldsymbol{\mathcal{Q}}_{j,\text{WL}} \overset{\triangle}{=} \boldsymbol{\mathcal{C}}_j^H\,\mathbf{U}_n\,\mathbf{U}_n^H\,\boldsymbol{\mathcal{C}}_j \in \mathbb{C}^{2L_j\times 2L_j}$. The sample estimate $\widehat{\boldsymbol{\mathcal{Q}}}_{j,\text{WL}} = \boldsymbol{\mathcal{C}}_j^H\,\widehat{\mathbf{U}}_n\,\widehat{\mathbf{U}}_n^H\,\boldsymbol{\mathcal{C}}_j$ of matrix $\boldsymbol{\mathcal{Q}}_{j,\text{WL}}$ can be decomposed as $\widehat{\boldsymbol{\mathcal{Q}}}_{j,\text{WL}} = \boldsymbol{\mathcal{Q}}_{j,\text{WL}} + \boldsymbol{\delta}\boldsymbol{\mathcal{Q}}_{j,\text{WL}}$ where, accounting for (B.26), the perturbation $\boldsymbol{\delta}\boldsymbol{\mathcal{Q}}_{j,\text{WL}}$ has the form

$$\begin{aligned}\boldsymbol{\delta}\boldsymbol{\mathcal{Q}}_{j,\text{WL}} &\approx \boldsymbol{\mathcal{C}}_j^H\,\boldsymbol{\delta}\mathbf{U}_n\,\mathbf{U}_n^H\,\boldsymbol{\mathcal{C}}_j + \boldsymbol{\mathcal{C}}_j^H\,\mathbf{U}_n\,\boldsymbol{\delta}\mathbf{U}_n^H\,\boldsymbol{\mathcal{C}}_j \\ &= -\boldsymbol{\mathcal{C}}_j^H\,\mathbf{U}_s\,\boldsymbol{\Omega}_{\text{WL}}^{-1}\,\mathbf{U}_s^H\,\mathbf{h}_j\,\widehat{\mathbf{r}}_{\mathbf{q}_j b_j}^H\,\mathbf{U}_n\,\mathbf{U}_n^H\,\boldsymbol{\mathcal{C}}_j \\ &\quad - \boldsymbol{\mathcal{C}}_j^H\,\mathbf{U}_n\,\mathbf{U}_n^H\,\widehat{\mathbf{r}}_{\mathbf{q}_j b_j}\,\mathbf{h}_j^H\,\mathbf{U}_s\,\boldsymbol{\Omega}_{\text{WL}}^{-1}\,\mathbf{U}_s^H\,\boldsymbol{\mathcal{C}}_j\,. \end{aligned} \tag{B.27}$$

Based on (4.47), one has $\mathbf{T}_j^H\,\widehat{\boldsymbol{\mathcal{Q}}}_{j,\text{WL}}\,\mathbf{T}_j\,\widehat{\boldsymbol{\varrho}}_j = \mathbf{T}_j^H\,(\boldsymbol{\mathcal{Q}}_{j,\text{WL}} + \boldsymbol{\delta}\boldsymbol{\mathcal{Q}}_{j,\text{WL}})\,\mathbf{T}_j\,(\boldsymbol{\varrho}_j + \boldsymbol{\delta}\boldsymbol{\varrho}_j) \approx \mathbf{T}_j^H\,\boldsymbol{\mathcal{Q}}_{j,\text{WL}}\,\mathbf{T}_j\,\boldsymbol{\delta}\boldsymbol{\varrho}_j + \mathbf{T}_j^H\,\boldsymbol{\delta}\boldsymbol{\mathcal{Q}}_{j,\text{WL}}\,\mathbf{T}_j\,\boldsymbol{\varrho}_j \approx \mathbf{0}_{2L_j}$, which implies

that $\mathbf{T}_j^H\,\boldsymbol{\mathcal{Q}}_{j,\mathrm{WL}}\,\mathbf{T}_j\,\boldsymbol{\delta\varrho}_j \approx -\mathbf{T}_j^H\,\boldsymbol{\delta\mathcal{Q}}_{j,\mathrm{WL}}\,\mathbf{T}_j\,\boldsymbol{\varrho}_j$, whose minimal-norm least-squares solution [24] is given by

$$\begin{aligned}\boldsymbol{\delta\varrho}_j &\approx -(\mathbf{T}_j^H\,\boldsymbol{\mathcal{Q}}_{j,\mathrm{WL}}\,\mathbf{T}_j)^{\dagger}\,\mathbf{T}_j^H\,\boldsymbol{\delta\mathcal{Q}}_{j,\mathrm{WL}}\,\mathbf{T}_j\,\boldsymbol{\varrho}_j \\ &= -\mathbf{T}_j^H\,\boldsymbol{\mathcal{Q}}_{j,\mathrm{WL}}^{\dagger}\,\mathbf{T}_j\,\mathbf{T}_j^H\,\boldsymbol{\delta\mathcal{Q}}_{j,\mathrm{WL}}\,\mathbf{T}_j\,\boldsymbol{\varrho}_j = -\mathbf{T}_j^H\,\boldsymbol{\mathcal{Q}}_{j,\mathrm{WL}}^{\dagger}\,\boldsymbol{\delta\mathcal{Q}}_{j,\mathrm{WL}}\,\mathbf{T}_j\,\boldsymbol{\varrho}_j\,,\end{aligned} \tag{B.28}$$

since $\mathbf{T}_j$ is unitary. Substituting (B.27) in (B.28) and observing that, due to (4.47), $\mathbf{U}_n^H\,\boldsymbol{\mathcal{C}}_j\,\mathbf{T}_j\,\boldsymbol{\varrho}_j = \mathbf{0}_{2N-J}$, one has $\boldsymbol{\delta\varrho}_j \approx \mathbf{T}_j^H\,\boldsymbol{\mathcal{Q}}_{j,\mathrm{WL}}^{\dagger}\,\boldsymbol{\mathcal{C}}_j^H\,\mathbf{U}_n\,\mathbf{U}_n^H\,\widehat{\mathbf{r}}_{\mathbf{q}_j b_j}\,(\mathbf{h}_j^H\,\mathbf{U}_s\,\boldsymbol{\Omega}_{\mathrm{WL}}^{-1}\,\mathbf{U}_s^H\,\boldsymbol{\mathcal{C}}_j\,\mathbf{T}_j\,\boldsymbol{\varrho}_j)$, from which we finally have $\boldsymbol{\delta}\mathbf{h}_j = \widetilde{\alpha}_j\,\boldsymbol{\mathcal{C}}_j\,\mathbf{T}_j\,\boldsymbol{\delta\varrho}_j = (\mathbf{h}_j^H\,\mathbf{U}_s\,\boldsymbol{\Omega}_{\mathrm{WL}}^{-1}\,\mathbf{U}_s^H\,\mathbf{h}_j)\,\boldsymbol{\mathcal{C}}_j\,\boldsymbol{\mathcal{Q}}_{j,\mathrm{WL}}^{\dagger}\,\boldsymbol{\mathcal{C}}_j^H\,\mathbf{U}_n\,\mathbf{U}_n^H\,\widehat{\mathbf{r}}_{\mathbf{q}_j b_j}$.

B.7 Evaluation of *trace*$(\boldsymbol{\Sigma}_{j,\boldsymbol{WL}}^{H}\,\mathbf{R}_{\mathbf{zz}}\,\boldsymbol{\Sigma}_{j,\boldsymbol{WL}}\,\mathbf{R}_{\mathbf{q}_j\mathbf{q}_j})$

Initially, we will proceed in a unified manner by treating the SMI and SUB cases jointly. Since $\boldsymbol{\Sigma}_{j,\mathrm{WL}} = -\boldsymbol{\Gamma}_{j,\mathrm{WL}} + \boldsymbol{\Delta}_{j,\mathrm{WL}}\,\boldsymbol{\Pi}_{j,\mathrm{WL}}$ (see Lemma 4.6), using the linearity property of the trace operator and observing that $\boldsymbol{\Delta}_{j,\mathrm{WL}}$ is Hermitian (see Lemma 4.4), we can write

$$\begin{aligned}&\mathrm{trace}(\boldsymbol{\Sigma}_{j,\mathrm{WL}}^{H}\,\mathbf{R}_{\mathbf{zz}}\,\boldsymbol{\Sigma}_{j,\mathrm{WL}}\,\mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}) = \mathrm{trace}(\boldsymbol{\Gamma}_{j,\mathrm{WL}}^{H}\,\mathbf{R}_{\mathbf{zz}}\,\boldsymbol{\Gamma}_{j,\mathrm{WL}}\,\mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}) \\ &-\mathrm{trace}(\boldsymbol{\Pi}_{j,\mathrm{WL}}^{H}\,\boldsymbol{\Delta}_{j,\mathrm{WL}}\,\mathbf{R}_{\mathbf{zz}}\,\boldsymbol{\Gamma}_{j,\mathrm{WL}}\,\mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}) - \mathrm{trace}(\boldsymbol{\Gamma}_{j,\mathrm{WL}}^{H}\,\mathbf{R}_{\mathbf{zz}}\,\boldsymbol{\Delta}_{j,\mathrm{WL}}\,\boldsymbol{\Pi}_{j,\mathrm{WL}}\,\mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}) \\ &+\mathrm{trace}(\boldsymbol{\Pi}_{j,\mathrm{WL}}^{H}\,\boldsymbol{\Delta}_{j,\mathrm{WL}}\,\mathbf{R}_{\mathbf{zz}}\,\boldsymbol{\Delta}_{j,\mathrm{WL}}\,\boldsymbol{\Pi}_{j,\mathrm{WL}}\,\mathbf{R}_{\mathbf{q}_j\mathbf{q}_j})\,.\end{aligned} \tag{B.29}$$

By invoking the properties of the trace operator, it follows that

$$\begin{aligned}&\mathrm{trace}(\boldsymbol{\Gamma}_{j,\mathrm{WL}}^{H}\,\mathbf{R}_{\mathbf{zz}}\,\boldsymbol{\Delta}_{j,\mathrm{WL}}\,\boldsymbol{\Pi}_{j,\mathrm{WL}}\,\mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}) \\ &\qquad = \mathrm{trace}^{*}(\mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}\,\boldsymbol{\Pi}_{j,\mathrm{WL}}^{H}\,\boldsymbol{\Delta}_{j,\mathrm{WL}}\,\mathbf{R}_{\mathbf{zz}}\,\boldsymbol{\Gamma}_{j,\mathrm{WL}}) \\ &\qquad = \mathrm{trace}^{*}(\boldsymbol{\Pi}_{j,\mathrm{WL}}^{H}\,\boldsymbol{\Delta}_{j,\mathrm{WL}}\,\mathbf{R}_{\mathbf{zz}}\,\boldsymbol{\Gamma}_{j,\mathrm{WL}}\,\mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}),\end{aligned} \tag{B.30}$$

which shows that the third summand in (B.29) is the conjugate version of the second one. Moreover, remembering that $\mathbf{R}_{\mathbf{zz}} = \mathbf{h}_j\,\mathbf{h}_j^H + \mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}$ and $\mathbf{U}_n^H\,\mathbf{h}_j = \mathbf{0}_{2N-J}$, and accounting for the expressions of $\boldsymbol{\Gamma}_{j,\mathrm{WL}}$ (see Lemma 4.4) and $\boldsymbol{\Pi}_{j,\mathrm{WL}}$ (see Lemma 4.5), it can be directly verified that $\mathbf{R}_{\mathbf{zz}}\,\boldsymbol{\Gamma}_{j,\mathrm{WL}} = \mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}\,\boldsymbol{\Gamma}_{j,\mathrm{WL}}$ and $\boldsymbol{\Pi}_{j,\mathrm{WL}}\,\mathbf{R}_{\mathbf{q}_j\mathbf{q}_j} = \boldsymbol{\Pi}_{j,\mathrm{WL}}\,\mathbf{R}_{\mathbf{zz}}$. Thus, the first summand in (B.29) becomes $\mathrm{trace}(\boldsymbol{\Gamma}_{j,\mathrm{WL}}^{H}\,\mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}\,\boldsymbol{\Gamma}_{j,\mathrm{WL}}\,\mathbf{R}_{\mathbf{q}_j\mathbf{q}_j})$, whereas the fourth one reduces to $\mathrm{trace}[(\boldsymbol{\Delta}_{j,\mathrm{WL}}\,\mathbf{R}_{\mathbf{zz}}\,\boldsymbol{\Delta}_{j,\mathrm{WL}})\,(\boldsymbol{\Pi}_{j,\mathrm{WL}}\,\mathbf{R}_{\mathbf{zz}}\,\boldsymbol{\Pi}_{j,\mathrm{WL}}^{H})]$, where

we have also used the properties of the trace operator again. This last trace can be further explicated by replacing $\mathbf{\Pi}_{j,\mathrm{WL}}$ with its expression given in Lemma 4.5: in particular, using $\mathbf{R_{zz}} = \mathbf{U}_s\,\mathbf{\Lambda}_s\,\mathbf{U}_s^H + \sigma_v^2\,\mathbf{U}_n\,\mathbf{U}_n^H$, remembering that $\mathbf{U}_n^H\,\mathbf{U}_n = \mathbf{I}_{2N-J}$ and $\mathbf{U}_n^H\,\mathbf{U}_s = \mathbf{O}_{(2N-J)\times J}$, and observing that, on the basis of the Moore-Penrose conditions [24], $\boldsymbol{\mathcal{Q}}_{j,\mathrm{WL}}^{\dagger}\,\boldsymbol{\mathcal{Q}}_{j,\mathrm{WL}}\,\boldsymbol{\mathcal{Q}}_{j,\mathrm{WL}}^{\dagger} = \boldsymbol{\mathcal{Q}}_{j,\mathrm{WL}}^{\dagger}$, one has

$$\mathbf{\Pi}_{j,\mathrm{WL}}\,\mathbf{R_{zz}}\,\mathbf{\Pi}_{j,\mathrm{WL}}^H = \sigma_v^2\,(\mathbf{h}_j^H\,\mathbf{U}_s\,\mathbf{\Omega}_{\mathrm{WL}}^{-1}\,\mathbf{U}_s^H\,\mathbf{h}_j)^2\,\boldsymbol{\mathcal{C}}_j\,\boldsymbol{\mathcal{Q}}_{j,\mathrm{WL}}^{\dagger}\,\boldsymbol{\mathcal{C}}_j^H. \tag{B.31}$$

Consequently, taking in account (B.30)-(B.31) and substituting the expression of $\mathbf{\Pi}_{j,\mathrm{WL}}$ in the second summand of (B.29), we get

$$\begin{aligned}
&\mathrm{trace}(\mathbf{\Sigma}_{j,\mathrm{WL}}^H\,\mathbf{R_{zz}}\,\mathbf{\Sigma}_{j,\mathrm{WL}}\,\mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}) = \mathrm{trace}(\mathbf{\Gamma}_{j,\mathrm{WL}}^H\,\mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}\,\mathbf{\Gamma}_{j,\mathrm{WL}}\,\mathbf{R}_{\mathbf{q}_j\mathbf{q}_j})\\
&-2\,(\mathbf{h}_j^H\,\mathbf{U}_s\,\mathbf{\Omega}_{\mathrm{WL}}^{-1}\,\mathbf{U}_s^H\,\mathbf{h}_j)\,\mathrm{Re}[\mathrm{trace}(\mathbf{U}_n\,\mathbf{U}_n^H\,\boldsymbol{\mathcal{C}}_j\,\boldsymbol{\mathcal{Q}}_{j,\mathrm{WL}}^{\dagger}\,\boldsymbol{\mathcal{C}}_j^H\,\mathbf{\Delta}_{j,\mathrm{WL}}\,\mathbf{R_{zz}}\,\mathbf{\Gamma}_{j,\mathrm{WL}}\,\mathbf{R}_{\mathbf{q}_j\mathbf{q}_j})]\\
&+\sigma_v^2\,(\mathbf{h}_j^H\,\mathbf{U}_s\,\mathbf{\Omega}_{\mathrm{WL}}^{-1}\,\mathbf{U}_s^H\,\mathbf{h}_j)^2\,\mathrm{trace}[(\mathbf{\Delta}_{j,\mathrm{WL}}\,\mathbf{R_{zz}}\,\mathbf{\Delta}_{j,\mathrm{WL}})\,(\boldsymbol{\mathcal{C}}_j\,\boldsymbol{\mathcal{Q}}_{j,\mathrm{WL}}^{\dagger}\,\boldsymbol{\mathcal{C}}_j^H)]\,.
\end{aligned} \tag{B.32}$$

At this point, we have to consider the SMI and SUB cases separately.

Let us start from the SMI case, for which $\mathbf{\Gamma}_{j,\mathrm{WL}} = \mathbf{P}_{j,\mathrm{WL}}\,\mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}^{-1}$ and $\mathbf{\Delta}_{j,\mathrm{WL}} = (\mathbf{h}_j^H\mathbf{R_{zz}}^{-1}\mathbf{h}_j)^{-1}\mathbf{R_{zz}}^{-1} - 2\,\mathbf{f}_j\mathbf{f}_j^H$. Recalling the equation (4.76) in the subsection 4.5.1, we know that $\mathrm{trace}(\mathbf{\Gamma}_{j,\mathrm{WL}}^H\mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}\,\mathbf{\Gamma}_{j,\mathrm{WL}}\,\mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}) = 2\,N - 1$. As regards the second summand in (B.32), we observe that $\mathbf{\Delta}_{j,\mathrm{WL}}\,\mathbf{R_{zz}}\,\mathbf{\Gamma}_{j,\mathrm{WL}}\,\mathbf{R}_{\mathbf{q}_j\mathbf{q}_j} = (\mathbf{h}_j^H\mathbf{R_{zz}}^{-1}\mathbf{h}_j)^{-1}\,(\mathbf{P}_{j,\mathrm{WL}} - \mathbf{f}_j\,\mathbf{h}_j^H)\,\mathbf{P}_{j,\mathrm{WL}} = (\mathbf{h}_j^H\mathbf{R_{zz}}^{-1}\mathbf{h}_j)^{-1}\,\mathbf{P}_{j,\mathrm{WL}}$, where the second equality follows by noticing that $\mathbf{h}_j^H\,\mathbf{P}_{j,\mathrm{WL}} = \mathbf{0}_{2N}^T$ and $\mathbf{P}_{j,\mathrm{WL}}^2 = \mathbf{P}_{j,\mathrm{WL}}$. Henceforth, observing that $\mathbf{P}_{j,\mathrm{WL}}\,\mathbf{U}_n = \mathbf{U}_n$ and using the trace properties, the second summand in (B.32) becomes

$$\begin{aligned}
&-2\,\zeta_{j,\mathrm{WL}}\,\mathrm{Re}[\mathrm{trace}(\mathbf{U}_n\,\mathbf{U}_n^H\,\boldsymbol{\mathcal{C}}_j\,\boldsymbol{\mathcal{Q}}_{j,\mathrm{WL}}^{\dagger}\,\boldsymbol{\mathcal{C}}_j^H\,\mathbf{P}_{j,\mathrm{WL}})]\\
&\quad = -2\,\zeta_{j,\mathrm{WL}}\,\mathrm{Re}[\mathrm{trace}(\mathbf{P}_{j,\mathrm{WL}}\,\mathbf{U}_n\,\mathbf{U}_n^H\,\boldsymbol{\mathcal{C}}_j\,\boldsymbol{\mathcal{Q}}_{j,\mathrm{WL}}^{\dagger}\,\boldsymbol{\mathcal{C}}_j^H)]\\
&\quad = -2\,\zeta_{j,\mathrm{WL}}\,\mathrm{Re}[\mathrm{trace}(\boldsymbol{\mathcal{Q}}_{j,\mathrm{WL}}^{\dagger}\,\underbrace{\boldsymbol{\mathcal{C}}_j^H\,\mathbf{U}_n\,\mathbf{U}_n^H\,\boldsymbol{\mathcal{C}}_j}_{\boldsymbol{\mathcal{Q}}_{j,\mathrm{WL}}})] = -2\,\zeta_{j,\mathrm{WL}}\,(2\,L_j - 1),
\end{aligned} \tag{B.33}$$

with $\zeta_{j,\mathrm{WL}} \stackrel{\triangle}{=} (\mathbf{h}_j^H\,\mathbf{R_{zz}}^{-1}\,\mathbf{h}_j)^{-1}\,\mathbf{h}_j^H\,\mathbf{U}_s\,\mathbf{\Omega}_{\mathrm{WL}}^{-1}\,\mathbf{U}_s^H\,\mathbf{h}_j > 0$, where the last equality comes from the fact that $\boldsymbol{\mathcal{Q}}_{j,\mathrm{WL}}^{\dagger}\,\boldsymbol{\mathcal{Q}}_{j,\mathrm{WL}}$ is the orthogonal projector onto the

subspace $\mathcal{R}(\boldsymbol{\mathcal{Q}}^H_{j,\mathrm{WL}}) \equiv \mathcal{R}(\boldsymbol{\mathcal{Q}}_{j,\mathrm{WL}})$ (Moore definition of the generalized inverse [24]) and, hence,[1]

$$\mathrm{trace}(\boldsymbol{\mathcal{Q}}^{\dagger}_{j,\mathrm{WL}}\,\boldsymbol{\mathcal{Q}}_{j,\mathrm{WL}}) = \mathrm{rank}(\boldsymbol{\mathcal{Q}}_{j,\mathrm{WL}}) = \mathrm{rank}(\mathbf{U}^H_n\,\boldsymbol{\mathcal{C}}_j) = 2\,L_j - 1 \quad \text{(B.34)}$$

where the last equality is justified because $\mathrm{rank}(\mathbf{U}^H_n\,\boldsymbol{\mathcal{C}}_j) = 2\,L_j - 1$ by virtue of condition **(c3)**. Considering the third summand in (B.32), we note that $\boldsymbol{\Delta}_{j,\mathrm{WL}}\,\mathbf{R}_{\mathbf{zz}}\,\boldsymbol{\Delta}_{j,\mathrm{WL}} = \mathbf{R}^{-1}_{\mathbf{zz}}(\mathbf{R}_{\mathbf{zz}}\,\boldsymbol{\Delta}_{j,\mathrm{WL}})^2 = (\mathbf{h}^H_j\mathbf{R}^{-1}_{\mathbf{zz}}\mathbf{h}_j)^{-2}\,\mathbf{R}^{-1}_{\mathbf{zz}}\,(\mathbf{P}^H_{j,\mathrm{WL}} - \mathbf{h}_j\,\mathbf{f}^H_j)^2$ which, using the facts that $(\mathbf{P}^H_{j,\mathrm{WL}})^2 = \mathbf{P}^H_{j,\mathrm{WL}}$, $\mathbf{P}^H_{j,\mathrm{WL}}\,\mathbf{h}_j = \mathbf{0}_{2N}$, $\mathbf{f}^H_j\,\mathbf{P}^H_{j,\mathrm{WL}} = \mathbf{0}^T_{2N}$ and $\mathbf{P}^H_{j,\mathrm{WL}} + \mathbf{h}_j\,\mathbf{f}^H_j = \mathbf{I}_{2N}$, ends up to $\boldsymbol{\Delta}_{j,\mathrm{WL}}\,\mathbf{R}_{\mathbf{zz}}\,\boldsymbol{\Delta}_{j,\mathrm{WL}} = (\mathbf{h}^H_j\mathbf{R}^{-1}_{\mathbf{zz}}\mathbf{h}_j)^{-2}\,\mathbf{R}^{-1}_{\mathbf{zz}}$. Consequently, the third summand in (B.32) assumes the form

$$\sigma^2_v\,\zeta^2_{j,\mathrm{WL}}\,\mathrm{trace}(\mathbf{R}^{-1}_{\mathbf{zz}}\,\boldsymbol{\mathcal{C}}_j\,\boldsymbol{\mathcal{Q}}^{\dagger}_{j,\mathrm{WL}}\,\boldsymbol{\mathcal{C}}^H_j). \quad \text{(B.35)}$$

By substituting (B.35), (B.33) and (4.76) in (B.32), we have proven (4.127):

$$\begin{aligned}\mathrm{trace}(\boldsymbol{\Sigma}^H_{j,\mathrm{WL}}\,\mathbf{R}_{\mathbf{zz}}\,\boldsymbol{\Sigma}_{j,\mathrm{WL}}\,\mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}) &= (2\,N-1) - 2\,\zeta_{j,\mathrm{WL}}(2\,L_j - 1)\\ &+ \zeta^2_{j,\mathrm{WL}}\,\sigma^2_v\,\mathrm{trace}(\mathbf{R}^{-1}_{\mathbf{zz}}\,\boldsymbol{\mathcal{C}}_j\,\boldsymbol{\mathcal{Q}}^{\dagger}_{j,\mathrm{WL}}\,\boldsymbol{\mathcal{C}}^H_j). \quad \text{(B.36)}\end{aligned}$$

Let us consider now the SUB case, wherein $\boldsymbol{\Delta}_{j,\mathrm{WL}} = (\mathbf{h}^H_j\mathbf{R}^{-1}_{\mathbf{zz}}\mathbf{h}_j)^{-1}\mathbf{U}_s\boldsymbol{\Lambda}^{-1}_s\mathbf{U}^H_s - 2\,\mathbf{f}_j\mathbf{f}^H_j$ and $\boldsymbol{\Gamma}_{j,\mathrm{WL}} = \mathbf{P}_{j,\mathrm{WL}}\,\mathbf{R}^{-1}_{\mathbf{q}_j\mathbf{q}_j} - \gamma_{j,\mathrm{WL}}\,\mathbf{U}_n\mathbf{U}^H_n$, with $\gamma_{j,\mathrm{WL}} \stackrel{\triangle}{=} \sigma^{-2}_v + (\mathbf{h}^H_j\,\mathbf{R}^{-1}_{\mathbf{zz}}\,\mathbf{h}_j)^{-1}\mathbf{h}^H_j\mathbf{U}_s\,\boldsymbol{\Omega}^{-1}_{\mathrm{WL}}\,\mathbf{U}^H_s\mathbf{R}^{-1}_{\mathbf{zz}}\,\mathbf{h}_j$. In subsection 4.5.1, we have shown equation (4.78) that we report here for simplicity

$$\mathrm{trace}(\boldsymbol{\Gamma}^H_{j,\mathrm{WL}}\mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}\,\boldsymbol{\Gamma}_{j,\mathrm{WL}}\,\mathbf{R}_{\mathbf{q}_j\mathbf{q}_j}) = (J-1) + (2\,N-J)\,|1-\gamma_{j,\mathrm{WL}}\,\sigma^2_v|^2. \quad \text{(B.37)}$$

As to the second summand in (B.32), since $\mathbf{R}_{\mathbf{zz}} = \mathbf{h}_j\,\mathbf{h}^H_j + \mathbf{R}_{\mathbf{q}_j\mathbf{q}_j} = \mathbf{U}_s\,\boldsymbol{\Lambda}_s\,\mathbf{U}^H_s + \sigma^2_v\,\mathbf{U}_n\,\mathbf{U}^H_n$, with $\mathbf{U}^H_n\,\mathbf{h}_j = \mathbf{0}_{2N-J}$, $\mathbf{U}^H_n\,\mathbf{U}_n = \mathbf{I}_{2N-J}$, $\mathbf{U}^H_s\,\mathbf{U}_s = \mathbf{I}_J$, $\mathbf{U}^H_n\,\mathbf{U}_s = \mathbf{O}_{(2N-J)\times J}$ and $\mathbf{U}_n\,\mathbf{U}^H_n + \mathbf{U}_s\,\mathbf{U}^H_s = \mathbf{I}_{2N}$, we obtain that $\boldsymbol{\Gamma}_{j,\mathrm{WL}}\,\mathbf{R}_{\mathbf{q}_j\mathbf{q}_j} = \mathbf{P}_{j,\mathrm{WL}} - \gamma_{j,\mathrm{WL}}\,\mathbf{U}_n\mathbf{U}^H_n\mathbf{R}_{\mathbf{zz}} = \mathbf{P}_{j,\mathrm{WL}} - \gamma_{j,\mathrm{WL}}\,\sigma^2_v\,\mathbf{U}_n\mathbf{U}^H_n$ and $\boldsymbol{\Delta}_{j,\mathrm{WL}}\,\mathbf{R}_{\mathbf{zz}} = (\mathbf{h}^H_j\mathbf{R}^{-1}_{\mathbf{zz}}\mathbf{h}_j)^{-1}\,(\mathbf{P}_{j,\mathrm{WL}} - \mathbf{f}_j\,\mathbf{h}^H_j - \mathbf{U}_n\mathbf{U}^H_n)$. Consequently, we get

$$\boldsymbol{\Delta}_{j,\mathrm{WL}}\,\mathbf{R}_{\mathbf{zz}}\,\boldsymbol{\Gamma}_{j,\mathrm{WL}}\,\mathbf{R}_{\mathbf{q}_j\mathbf{q}_j} = (\mathbf{h}^H_j\mathbf{R}^{-1}_{\mathbf{zz}}\mathbf{h}_j)^{-1}\,(\mathbf{P}_{j,\mathrm{WL}} - \mathbf{U}_n\,\mathbf{U}^H_n)\,, \quad \text{(B.38)}$$

where we have used the facts that $\mathbf{P}^2_{j,\mathrm{WL}} = \mathbf{P}_{j,\mathrm{WL}}$, $\mathbf{h}^H_j\,\mathbf{P}_{j,\mathrm{WL}} = \mathbf{0}^T_{2N}$, $\mathbf{U}^H_n\,\mathbf{P}_{j,\mathrm{WL}} = \mathbf{U}^H_n$, $\mathbf{h}^H_j\,\mathbf{U}_n = \mathbf{0}^T_{2N-J}$ and $\mathbf{P}_{j,\mathrm{WL}}\,\mathbf{U}_n = \mathbf{U}_n$. Therefore, observing again that $\mathbf{P}_{j,\mathrm{WL}}\,\mathbf{U}_n = \mathbf{U}_n$, $\mathbf{U}^H_n\,\mathbf{U}_n = \mathbf{I}_{2N-J}$ and using the trace

[1] If χ is an eigenvalue of the orthogonal projector $\boldsymbol{\mathcal{Q}}^{\dagger}_{j,\mathrm{WL}}\,\boldsymbol{\mathcal{Q}}_{j,\mathrm{WL}}$, then $\chi \in \{0,1\}$.

properties, the second summand in (B.32) simplifies to

$$\begin{aligned}-2\,\zeta_{j,\mathrm{WL}}\mathrm{Re}[&\mathrm{trace}(\mathbf{U}_n\,\mathbf{U}_n^H\,\boldsymbol{\mathcal{C}}_j\,\boldsymbol{\mathcal{Q}}_{j,\mathrm{WL}}^{\dagger}\,\boldsymbol{\mathcal{C}}_j^H\,\mathbf{P}_{j,\mathrm{WL}})\\ &-\mathrm{trace}(\mathbf{U}_n\,\mathbf{U}_n^H\,\boldsymbol{\mathcal{C}}_j\,\boldsymbol{\mathcal{Q}}_{j,\mathrm{WL}}^{\dagger}\,\boldsymbol{\mathcal{C}}_j^H\,\mathbf{U}_n\,\mathbf{U}_n^H)]=\\ -2\,\zeta_{j,\mathrm{WL}}&\left\{\mathrm{Re}\left[\mathrm{trace}(\boldsymbol{\mathcal{Q}}_{j,\mathrm{WL}}^{\dagger}\,\boldsymbol{\mathcal{Q}}_{j,\mathrm{WL}})-\mathrm{trace}(\boldsymbol{\mathcal{Q}}_{j,\mathrm{WL}}^{\dagger}\,\boldsymbol{\mathcal{Q}}_{j,\mathrm{WL}})\right]\right\}=0.\end{aligned}\quad\text{(B.39)}$$

With reference to the third summand in (B.32), we note that $\boldsymbol{\Delta}_{j,\mathrm{WL}}\,\mathbf{R}_{\mathbf{zz}}\,\boldsymbol{\Delta}_{j,\mathrm{WL}}=\mathbf{R}_{\mathbf{zz}}^{-1}(\mathbf{R}_{\mathbf{zz}}\,\boldsymbol{\Delta}_{j,\mathrm{WL}})^2=(\mathbf{h}_j^H\mathbf{R}_{\mathbf{zz}}^{-1}\mathbf{h}_j)^{-2}\,\mathbf{R}_{\mathbf{zz}}^{-1}\,(\mathbf{P}_{j,\mathrm{WL}}^H-\mathbf{h}_j\,\mathbf{f}_j^H-\mathbf{U}_n\mathbf{U}_n^H)^2$ which, exploiting the EVD $\mathbf{R}_{\mathbf{zz}}=\mathbf{U}_s\,\boldsymbol{\Lambda}_s\,\mathbf{U}_s^H+\sigma_v^2\,\mathbf{U}_n\,\mathbf{U}_n^H$ and its related properties (as done for the second summand), and using the facts that $(\mathbf{P}_{j,\mathrm{WL}}^H)^2=\mathbf{P}_{j,\mathrm{WL}}^H$, $\mathbf{f}_j^H\,\mathbf{P}_{j,\mathrm{WL}}^H=\mathbf{0}_{2N}^T$, $\mathbf{U}_n^H\,\mathbf{P}_{j,\mathrm{WL}}^H=\mathbf{U}_n^H$, $\mathbf{P}_{j,\mathrm{WL}}^H\,\mathbf{h}_j=\mathbf{0}_{2N}$, $\mathbf{U}_n^H\,\mathbf{h}_j=\mathbf{0}_{2N-J}$, $\mathbf{P}_{j,\mathrm{WL}}^H\,\mathbf{U}_n=\mathbf{U}_n$, $\mathbf{f}_j^H\,\mathbf{U}_n=\mathbf{0}_{2N-J}^T$ and $\mathbf{P}_{j,\mathrm{WL}}^H+\mathbf{h}_j\,\mathbf{f}_j^H=\mathbf{I}_{2N}$, boils down to $\boldsymbol{\Delta}_{j,\mathrm{WL}}^H\,\mathbf{R}_{\mathbf{zz}}\,\boldsymbol{\Delta}_{j,\mathrm{WL}}=(\mathbf{h}_j^H\mathbf{R}_{\mathbf{zz}}^{-1}\mathbf{h}_j)^{-2}\,(\mathbf{R}_{\mathbf{zz}}^{-1}-\sigma_v^{-2}\,\mathbf{U}_n\,\mathbf{U}_n^H)$. Consequently, the third summand in (B.32) assumes the form

$$\begin{aligned}\sigma_v^2\,\zeta_{j,\mathrm{WL}}^2\,\mathrm{trace}(\mathbf{R}_{\mathbf{zz}}^{-1}\,\boldsymbol{\mathcal{C}}_j\,\boldsymbol{\mathcal{Q}}_{j,\mathrm{WL}}^{\dagger}\,\boldsymbol{\mathcal{C}}_j^H)-\zeta_{j,\mathrm{WL}}^2\,&\underbrace{\mathrm{trace}(\mathbf{U}_n\,\mathbf{U}_n^H\,\boldsymbol{\mathcal{C}}_j\,\boldsymbol{\mathcal{Q}}_{j,\mathrm{WL}}^{\dagger}\,\boldsymbol{\mathcal{C}}_j^H)}_{\mathrm{trace}(\boldsymbol{\mathcal{Q}}_{j,\mathrm{WL}}^{\dagger}\,\boldsymbol{\mathcal{Q}}_{j,\mathrm{WL}})}\\ &=\sigma_v^2\,\zeta_{j,\mathrm{WL}}^2\,\mathrm{trace}(\mathbf{R}_{\mathbf{zz}}^{-1}\,\boldsymbol{\mathcal{C}}_j\,\boldsymbol{\mathcal{Q}}_{j,\mathrm{WL}}^{\dagger}\,\boldsymbol{\mathcal{C}}_j^H)-\zeta_{j,\mathrm{WL}}^2(2\,L_j-1).\end{aligned}\quad\text{(B.40)}$$

Thus, by substituting (B.40), (B.39) and (B.37) in (B.32), we have proven (4.128), too:

$$\begin{aligned}\mathrm{trace}(\boldsymbol{\Sigma}_{j,\mathrm{WL}}^H\,\mathbf{R}_{\mathbf{zz}}\,\boldsymbol{\Sigma}_{j,\mathrm{WL}}\,\mathbf{R}_{\mathbf{q}_j\mathbf{q}_j})&=(J-1)+(2\,N-J)|1-\gamma_{j,\mathrm{WL}}\,\sigma_v^2|^2\\ &-\zeta_{j,\mathrm{WL}}^2(2\,L_j-1)+\zeta_{j,\mathrm{WL}}^2\,\sigma_v^2\,\mathrm{trace}(\mathbf{R}_{\mathbf{zz}}^{-1}\,\boldsymbol{\mathcal{C}}_j\,\boldsymbol{\mathcal{Q}}_{j,\mathrm{WL}}^{\dagger}\,\boldsymbol{\mathcal{C}}_j^H).\end{aligned}\quad\text{(B.41)}$$

Appendix C

Equalization Techniques for MC-CDMA Systems

C.1 Proof of Theorem 5.1

Let us consider the case when the channel transfer function $G(z)$ has $0 \leq M_z \leq L$ distinct zeros on the subcarriers $z_{m_1} = e^{i\frac{2\pi}{N}m_1}, z_{m_2} = e^{i\frac{2\pi}{N}m_2}, \ldots, z_{m_{M_z}} = e^{i\frac{2\pi}{N}m_{M_z}}$, with $m_1 \neq m_2 \neq \cdots \neq m_{M_z} \in \{0, 1, \ldots, N-1\}$. In this case, one has

$$\gamma_{\text{cp}}(m_1) = \gamma_{\text{cp}}(m_2) = \cdots = \gamma_{\text{cp}}(m_{M_z}) = 0 \tag{C.1}$$

and, thus, the diagonal matrix $\mathbf{\Gamma}_{\text{cp}}$ is singular with

$$\text{rank}(\mathbf{\Gamma}_{\text{cp}}) = N - M_z. \tag{C.2}$$

In its turn, this implies that $\boldsymbol{\mathcal{G}}_{\text{cp}}$ may be rank deficient even if the code vectors $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_J$ are linearly independent, i.e., $\mathbf{C}$ is full-column rank. Indeed, under the assumptions that $J \leq N$ and $\text{rank}(\mathbf{C}) = J$, the matrix $\mathbf{\Gamma}_{\text{cp}}\,\mathbf{C}$ is full-column rank iff [24]

$$\mathcal{N}(\mathbf{\Gamma}_{\text{cp}}) \cap \mathcal{R}(\mathbf{C}) = \mathbf{0}_N. \tag{C.3}$$

The null space of $\mathbf{\Gamma}_{\text{cp}}$ can be readily characterized: an arbitrary vector $\boldsymbol{\mu} \in \mathbb{C}^N$ belongs to $\mathcal{N}(\mathbf{\Gamma}_{\text{cp}})$ iff there exists a vector $\boldsymbol{\beta} \in \mathbb{C}^{M_z}$ such that $\boldsymbol{\mu} = \mathbf{S}_z\,\boldsymbol{\beta}$. Hence, an arbitrary vector $\boldsymbol{\mu} \in \mathcal{N}(\mathbf{\Gamma}_{\text{cp}})$ also belongs to the subspace $\mathcal{R}(\mathbf{C})$ iff there exists a vector $\boldsymbol{\alpha} \in \mathbb{C}^J$ such that $\mathbf{S}_z\,\boldsymbol{\beta} = \mathbf{C}\,\boldsymbol{\alpha}$. As a consequence,

condition $\mathcal{N}(\mathbf{\Gamma}_{\text{cp}}) \cap \mathcal{R}(\mathbf{C}) = \mathbf{0}_N$ holds iff the system of equations $\mathbf{C}\,\boldsymbol{\alpha} - \mathbf{S}_z\,\boldsymbol{\beta} = \mathbf{0}_N$ admits the unique solution $\boldsymbol{\alpha} = \mathbf{0}_J$ and $\boldsymbol{\beta} = \mathbf{0}_{M_z}$. It can be seen [56] that this happens iff the matrix $[\mathbf{C}, \mathbf{S}_z] \in \mathbb{C}^{N\times(J+M_z)}$ turns out to be full-column rank.

C.2 Proof of Lemma 5.1

Preliminarily, observe that rank$(\overline{\mathbf{C}}) = J$ iff the null spaces of $\mathbf{C}$ and $\mathbf{C}^*$ intersect only trivially, that is, $\mathcal{N}(\mathbf{C}) \cap \mathcal{N}(\mathbf{C}^*) = \{\mathbf{0}_J\}$. An arbitrary *nonzero* vector $\boldsymbol{\alpha} \in \mathbb{C}^J$ belongs to $\mathcal{N}(\mathbf{C})$ iff $\mathbf{C}\,\boldsymbol{\alpha} = \mathbf{0}_N$, from which, by conjugating, one obtains $\mathbf{C}^*\boldsymbol{\alpha}^* = \mathbf{0}_N$. The last two systems of equations show that $\boldsymbol{\alpha} \in \mathcal{N}(\mathbf{C})$ iff $\boldsymbol{\alpha}^* \in \mathcal{N}(\mathbf{C}^*)$. Consequently, an arbitrary vector $\boldsymbol{\alpha} \neq \mathbf{0}_J$ belongs to $\mathcal{N}(\mathbf{C}) \cap \mathcal{N}(\mathbf{C}^*)$ iff there exists a nonzero vector $\boldsymbol{\beta} \in \mathbb{C}^J$ belonging to $\mathcal{N}(\mathbf{C})$ such that $\boldsymbol{\beta}^* = \boldsymbol{\alpha}$.

Bibliography

[1] P. Schreier and L. Scharf, "Second-order analysis of improper complex random vectors and processes," *IEEE Trans. Signal Processing*, vol. 51, no. 3, pp. 714–725, Mar. 2003.

[2] B. Picinbono and P. Chevalier, "Widely linear estimation with complex data," *IEEE Trans. Signal Processing*, vol. 43, no. 8, pp. 2030–2033, Aug. 1995.

[3] G. Gelli, L. Paura, and A. R. P. Ragozini, "Blind widely linear multiuser detection," *IEEE Commun. Lett.*, vol. 4, no. 6, pp. 187–189, June 2000.

[4] D. N. Godard, "Self-recovering equalitation and carrier tracking in two dimensional data communication systems," *IEEE Trans. Commun.*, vol. COMM-28, pp. 1867–1875, Nov. 1980.

[5] C. R. Johnson, P. Schniter, T. J. Endres, J. D. Behm, D. R. Brown, and R. A. Casas, "Blind equalization using the constant modulus criterion: a review," *Proc. IEEE*, vol. 86, pp. 1927–1950, Oct. 1998.

[6] H. H. Zeng, L. Tong, and C. R. Johnson, "An analysis of constant modulus receivers," *IEEE Transactions on Signal Processing*, vol. 47, pp. 2990–2999, Nov. 1999.

[7] F. D. Neeser and J. L. Massey, "Proper complex random processes with applications to information theory," *IEEE Trans. Inform. Theory*, pp. 1293–1302, July 1993.

[8] A. S. Cacciapuoti, G. Gelli, L. Paura, and F. Verde, "Widely-linear fractionally-spaced blind equalization of frequency-selective channels," in 14*th Proceed. of Europ. Sign. Proc. Conf. (EUSIPCO)*, Florence, Italy, Sept. 2006.

[9] ——, "Design and analysis of widely-linear constant modulus equalizers," 2008, to be submitted.

[10] S. Verdú, "Minimum probability of error for asynchronous gaussian multiple-access channels," *IEEE Trans. Inform. Theory*, vol. 32, no. 1, pp. 85–96, Jan. 1986.

[11] S. Hara and R. Prasad, "Overview of multicarrier cdma," *IEEE Commun. Mag.*, pp. 126–133, Dec. 1997.

[12] A. S. Cacciapuoti, G. Gelli, L. Paura, and F. Verde, "Finite-sample performance analysis of widely-linear multiuser receivers for ds-cdma systems," *IEEE Trans. Signal Processing*, vol. 56, no. 4, pp. 1572–1588, Apr. 2008.

[13] ——, "Finite-sample performance analysis of widely-linear multiuser receivers for ds-cdma systems," in *IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Helsinki, Finland, June 2007.

[14] ——, "Widely-linear versus linear blind multiuser detection with subspace-based channel estimation: finite sample-size effects," *IEEE Trans. Signal Processing*, 2008, to appear on.

[15] M. Honig, U. Madhow, and S. Verdù, "Blind adaptive multiuser detection," *IEEE Trans. Inform. Theory*, vol. 41, no. 4, pp. 944–960, July 1995.

[16] J. Mendel, "Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications," *Proc. IEEE*, vol. 79, no. 3, pp. 278–305, Mar. 1991.

[17] A. S. Cacciapuoti, G. Gelli, and F. Verde, "Universal code design for widely-linear multiuser detection in downlink mc-cdma," in *IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Helsinki, Finland, June 2007.

[18] ——, "Fir zero-forcing multiuser detection and code designs for downlink mc-cdma," *IEEE Trans. Signal Processing*, vol. 56, no. 4, pp. 1572–1588, Apr. 2008.

[19] A. Goldsmith, *Wireless Communications*. Cambridge: Cambridge University Press, 2005.

[20] R. E. Ziemer and R. L. Peterson, *Introduction to Digital Communications, Second Edition*. New Jersey: Prentice Hall, 2001.

[21] S. Barbarossa, *Multiantenna Wireless Communication Systems*. Boston, Mass: Artech House, 2005.

[22] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge: Cambridge University Press, 2005.

[23] K. Pahlavan and P. Krishnamurthy, *Principles of wireless networks*. New Jersey: Prentice Hall, 2002.

[24] A. Ben-Israel and T. N. E. Greville, *Generalized Inverses*. New York: Springer-Verlag, 2002.

[25] J. G. Proakis, *Digital Communications*. New York: McGraw-Hill, 2001.

[26] B. Picinbono, "On circularity," *IEEE Trans. Signal Processing*, pp. 3473–3482, Dec. 1994.

[27] Y. Li and Z. Ding, "Global convergence of fractionally spaced godard (cma) adaptive equalizers," *IEEE Trans. Signal Processing*, vol. 44, pp. 818–826, Apr. 1996.

[28] D. Liu and L. Tong, "An analysis of constant modulus algorithm for array signal processing," *Eurasip Journal on Signal Processing*, vol. 73, pp. 81–104, Jan. 1999.

[29] C. Papadias, "On the existence of undesirable global minima of godard equalizers," in *Proc. of Int. Conf. on Acoustic, Speech, and Signal Processing*, vol. 12, Munich, Germany, Apr. 1997, pp. 3941–3944.

[30] A. S. Cacciapuoti and F. Verde, "On the misbehavior of costant modulus equalizars for improper modulations," *IEEE Signal Processing Lett.*, vol. 14, no. 8, pp. 513–516, Aug. 2007.

[31] A. Shah, S. Biracree, R. Casas, T. Endres, S. Hulyalkar, T. Schaffer, and C. Strolle, "Global convergence of a single-axis constant modulus algorithm," in *Proc. of the Tenth IEEE Workshop on Statistical Signal and Array Processing*, Pennsylvania, USA, Aug. 2000, pp. 645–649.

[32] W. H. Gerstacker, R. Schober, and A. Lampe, "Receivers with widely linear processing for frequency-selective channels," *IEEE Trans. Commun.*, pp. 1512–1523, Sept. 2003.

[33] E. Serpedin and G. B. Giannakis, "A simple proof of a known blind channel identificability result," *IEEE Trans. Signal Processing*, vol. 47, pp. 591–593, Feb. 1999.

[34] C. Johnson and B. Anderson, "Godard blind equalizer error surface characteristics: white, zero-mean, binary source case," *Int. J. Adaptive Contr. Signal Processing*, vol. 9, pp. 301–324, July/Aug 1995.

[35] J. Leblanc, I. Fijalkow, and C. Johnson, "Cma fractionally spaced equalizers: stationary points and stability under iid and temporally correlated sources," *Int. J. Adaptive Contr. Signal Processing*, vol. vol. 12, pp. pp. 135–155, Mar. 1998.

[36] H. H. Zeng, L. Tong, and C. R. Johnson, "Relationships between the constant modulus and wiener receivers," *IEEE Trans. Inform. Theory*, vol. vol. 44, pp. pp. 1523–1538, July 1998.

[37] D. Darsena, G. Gelli, L. Paura, and F. Verde, "Subspace-based blind channel identification of siso-fir systems with improper random inputs," *Eurasip Journal on Signal Processing, Special Issue on Signal Processing in Communications*, vol. 84, no. 11, pp. 2021–2039, Nov. 2004.

[38] W. Gardner, *Introduction to Random Processes*. New York: McGraw-Hill, 1990.

[39] M. Rupp and S. C. Douglas, "A posteriori analysis of adaptive blind equalizers," in *Proc. of* 32*nd Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, Nov. 1998, pp. 369–373.

[40] R. L. Pickholtz, D. L. Schilling, and L. B. Milstein, "Theory of spread-spectrum communications-a tutorial," *IEEE Trans. Commun.*, vol. 30, pp. 855–883, May 1982.

[41] R. Kohno, R. Meidan, and L. B. Milstein, "Spread spectrum access methods for wireless communications," *IEEE Commun. Mag.*, Jan. 1995.

[42] M. Honig and M. K. Tsatsanis, "Adaptive techniques for multiuser cdma receivers," *IEEE Signal Processing Mag.*, vol. 17, no. 3, pp. 49–61, May 2000.

[43] V. Veeravalli and A. Mantravadi, "The coding-spreading tradeoff in cdma systems," *IEEE J. Select. Areas Commun.*, vol. 20, no. 2, pp. 396–408, Feb. 2002.

[44] X. Wang and H. Poor, "Blind equalization and multiuser detection in dispersive cdma channels," *IEEE Trans. Commun.*, pp. pp. 91–103, Jan. 1998.

[45] S. Verdú, *Multiuser detection*. Cambridge University Press, 1998.

[46] U. Madhow and M. Honig, "Mmse interference suppression for direct-sequence spread-spectrum cdma," *IEEE Trans. Commun.*, vol. 42, no. 12, pp. 3178–3188, Dec. 1994.

[47] S. V. Schell, "A separability theorem for $2m$ conjugate-symmetric signals impinging on an m-element sensor array," *IEEE Trans. Signal Processing*, vol. 45, no. 3, pp. 789–792, Mar. 1997.

[48] S. Verd´u, "Optimum multiuser signal detection," Ph.D. dissertation, University of Illinois, Urbana-Champaign, 1984.

[49] R. Lupas and S. Verdú, "Linear multiuser detectors for synchronous code-division multiple access channels," *IEEE Trans. Inform. Theory*, vol. 35, no. 1, pp. 123–136, Jan. 1989.

[50] S. Buzzi, M. Lops, and A. M. Tulino, "A new family of mmse multiuser receivers for interference suppression in ds-cdma systems employing bpsk modulation," *IEEE Trans. Commun.*, vol. 49, no. 1, pp. 154–167, Jan. 2001.

[51] R. Schober, W. H. Gerstacker, and L. Lampe, "On suboptimum receivers for ds-cdma with bpsk modulation," *Signal Processing*, vol. 85, no. 7, pp. 1149–1163, 2005.

[52] J.-J. Jeon, J. G. Andrews, and K.-M. Sung, "The blind widely linear minimum output energy algorithm for ds-cdma systems," *IEEE Trans. Signal Processing*, vol. 54, no. 5, pp. 1926–1931, May 2006.

[53] S. Buzzi, M. Lops, and A. M. Tulino, "A generalized minimum-mean-output-energy strategy for cdma systems with improper mai," *IEEE Trans. Inform. Theory*, vol. 48, no. 3, pp. 761–767, Mar. 2002.

[54] A. M. Tulino and S. Verdú, "Asymptotic analysis of improved linear receivers for bpsk-cdma subject to fading," *IEEE J. Select. Areas Commun.*, vol. 19, no. 8, pp. 1544–1555, Aug. 2001.

[55] Y. Yoon and K.-M. Kim, "An efficient blind multiuser detection for improper ds/cdma signals," *IEEE Trans. Veh. Technol.*, vol. 55, no. 2, pp. 572–582, Mar. 2006.

[56] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge: Cambridge University Press, 1990.

[57] A. Høst-Madsen, X. Wang, and S. Bahng, "Asymptotic analysis of blind multiuser detection with blind channel estimation," *IEEE Trans. Signal Processing*, vol. 52, no. 6, pp. 1722–1738, June 2004.

[58] H. V. Poor and X. Wang, "Code-aided interference suppression for ds/cdma communications – part i: Interference suppression capability," *IEEE Trans. Commun.*, vol. 45, no. 9, pp. 1101–1111, Sept. 1997.

[59] H. V. Poor and S. Verdù, "Probability of error in mmse multiuser detection," *IEEE Trans. Inform. Theory*, vol. 43, no. 5, pp. 868–871, May 1997.

[60] E. Moulines, P. Duhamel, J.-F. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel fir filters," *IEEE Trans. Signal Processing*, vol. 43, no. 2, pp. 516–525, Feb. 1995.

[61] H. Liu and G. Xu, "A subspace method for signature waveform estimation in synchronous cdma systems," *IEEE Trans. Commun.*, vol. 44, no. 10, pp. 1346–1354, Oct. 1996.

[62] P. Loubaton and E. Moulines, "On blind multiuser forward link channel estimation by the subspace method: identifiability results," *IEEE Trans. Signal Processing*, pp. 2366–2376, Aug. 2000.

[63] X. Wang and H. Poor, "Blind multiuser detection: a subspace approach," *IEEE Trans. Inform. Theory*, vol. 44, no. 2, pp. 677–690, Mar. 1998.

[64] S. Lasaulce, P. Loubaton, and E. Moulines, "A semi-blind channel estimation technique based on second-order blind method for cdma systems," *IEEE Trans. Signal Processing*, vol. 51, no. 7, pp. 1894–1904, July 2003.

[65] X. Li and H. H. Fan, "Qr factorization based blind channel identification and equalization with second-order statistics," *IEEE Trans. Signal Processing*, vol. 48, no. 1, pp. 60–69, Jan. 2000.

[66] K. Zarifi and A. B. Gershman, "Blind subspace-based signature waveform estimation in bpsk-modulated ds-cdma systems with circular noise," *IEEE Trans. Signal Processing*, pp. 3592–3602, Sept. 2006.

[67] H. Akaike, "A new look at the statistical model identification," *IEEE Trans. Automat. Contr.*, vol. 19, no. 6, pp. 716–723, June 1974.

[68] J. Rissanen, "Modeling by shortest data description," Automatica, vol. 14, pp. 465–471, 1978.

[69] X. Li and H. H. Fan, "Blind channel identification: subspace tracking method without rank estimation," *IEEE Trans. Signal Processing*, vol. 49, no. 10, pp. 2372–2382, Oct. 2001.

[70] D. N. C. Tse and S. V. Hanly, "Linear multiuser receivers: effective interference, effective bandwidth and user capacity," *IEEE Trans. Inform. Theory*, vol. vol. 45, no. 2, pp. 641–657, Mar. 1999.

[71] A. Lampe and M. Breiling, "Asymptotic analysis of widely linear mmse multiuser detection – complex vs real modulation," in *Proc. of the Information Theory Workshop (ITW) 2001*, Cairns, Australia, Sept. 2001, pp. 55–57.

[72] S. Buzzi and M. Lops, "Performance analysis for the improved linear multiuser detectors in bpsk-modulated ds-cdma systems," *IEEE Trans. Commun.*, vol. 51, no. 1, pp. 37–42, Jan. 2003.

[73] P. J. Schreier, L. L. Scharf, and C. T. Mullis, "Detection and estimation of improper complex random signals," *IEEE Trans. Inform. Theory*, vol. 51, no. 1, pp. 306–312, Jan. 2005.

[74] R. Schober, W. H. Gerstacker, and L. H.-J. Lampe, "Data-aided and blind stochastic gradient algorithms for widely linear mmse mai suppression for ds-cdma," *IEEE Trans. Signal Processing*, vol. 52, no. 3, pp. 746–756, Mar. 2004.

[75] M. Wax and Y. Anu, "Performance analysis of the minimum variance beamformer," *IEEE Trans. Signal Processing*, vol. 44, no. 4, pp. 928–937, Apr. 1996.

[76] Z. Xu and X. Wang, "Large-sample performance of blind and group-blind multiuser detectors: A perturbation perspective," *IEEE Trans. Inform. Theory*, vol. 50, no. 10, pp. 2389–2401, Oct. 2004.

[77] F. Li, H. Liu, and R. J. Vaccaro, "Performance analysis for doa estimation algorithms: unification, simplification, and observations," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 29, no. 4, pp. 1170–1184, Oct. 1993.

[78] D. D. Feldman and L. J. Griffiths, "A projection approach for robust adaptive beamforming," *IEEE Trans. Signal Processing*, vol. 42, no. 4, pp. 867–876, Apr. 1994.

[79] M. C. Dogan and J. M. Mendel, "Applications of cumulants to array processing – part i: Aperture extension and array calibration," *IEEE Trans. Signal Processing*, vol. 43, no. 5, pp. 1200–1216, May 1995.

[80] D. D. Feldman, "An analysis of the projection method for robust adaptive beamforming," *IEEE Trans. Antennas Propagat.*, vol. 44, no. 7, pp. 1023–1030, July 1996.

[81] A. Høst-Madsen and X. Wang, "Performance of blind and group-blind multiuser detection," *IEEE Trans. Inform. Theory*, vol. 48, no. 7, pp. 1849–1872, July 2002.

[82] M. Wax and Y. Anu, "Performance analysis of the minimum variance beamformer in presence of steerin vector errors," *IEEE Trans. Signal Processing*, vol. 44, no. 4, pp. 938–947, Apr. 1996.

[83] N. Yee, J.-P. Linnartz, and G. Fettweis, "Multicarrier cdma in indoor wireless radio networks," in *Proc. of 4th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC '93)*, vol. 12, Yokohama, Japan, Sept. 1993, pp. 468–472.

[84] Z. Wang and G. B. Giannakis, "Wireless multicarrier communications – where fourier meets shannon," *IEEE Signal Processing Mag.*, vol. 82, pp. 29–48, May 2000.

[85] X. Yue and H. Fan, "Near-far resistance of multicarrier cdma communication systems," *IEEE Signal Processing Lett.*, pp. 212–215, Feb. 2004.

[86] B. Muquet, Z. Wang, G. B. Giannakis, M. de Courville, and P. Duhamel, "Cyclic prefixing or zero padding for wireless multicarrier trasmissions?" *IEEE Trans. Commun.*, pp. 2136–2148, Dec. 2002.

[87] G. B. Giannakis, P. A. Anghel, and Z. Wang, "Generalized multicarrier cdma: unification and linear equalization," *Eurasip Journal on Applied Signal Processing*, pp. 743–756, May 2005.

[88] A. Kapur and M. K. Varanasi, "Multiuser detection for overloaded cdma systems," *IEEE Trans. Inform. Theory*, vol. 49, no. 7, pp. 1728–1742, July 2003.

[89] A. Scaglione, G. B. Giannakis, and S. Barbarossa, "Lagrange/vandermonde mui eliminating user codes for quasi-synchronous cdma in unknown multipath," *IEEE Trans. Signal Processing*, pp. 2057–2073, July 2000.

[90] ——, "Redundant filterbank precoders and equalizers – part i & ii," *IEEE Trans. Signal Processing*, pp. 1988–2006, July 1999.

[91] G. Marsaglia and G. P. H. Styan, "Equalities and inequalities for ranks of matrices," *Linear and Multilinear Algebra*, pp. 269–292, Feb. 1974.

[92] Z. Ding and G. Li, "Single-channel blind equalization for gsm cellular systems," *IEEE J. Select. Areas Commun.*, pp. 1493–1505, Oct. 1998.

[93] F. Verde, "Subspace-based blind multiuser detection for quasi-synchronous mc-cdma systems," *IEEE Signal Processing Lett.*, vol. 84, pp. 621–624, July 2004.

[94] M. K. Tsatsanis and Z. D. Xu, "Performance analysis of minimum variance cdma receivers," *IEEE Trans. Signal Processing*, vol. 46, no. 11, pp. pp. 3014–3022, Nov. 1998.

[95] G. Gelli and F. Verde, "Blind subspace-based channel identification for quasi-synchronous mc-cdma systems employing improper data symbols," in 14*th Proceed. of Europ. Sign. Proc. Conf. (EUSIPCO)*, Florence, Italy, Sept. 2006.