Blind direct multiuser detection for uplink MC-CDMA: performance analysis and robust implementation

Giacinto Gelli, Luigi Paura, and Francesco Verde *

Abstract

In this paper, we consider the problem of blind (i.e., without training sequences) linear mitigation of multiple-access interference in the uplink of quasi-synchronous multicarrier code-division multiple-access (MC-CDMA) systems. In the first part of the paper, we present the analytical performance assessment of the recently proposed blind two-stage multiuser detector, whose synthesis requires only the knowledge of the spreading code of the desired user. The analysis allows one to evaluate the actual performance when the receiver's parameters are estimated by resorting to a finite data record. Based on this analysis, in the second part of the paper, we propose to improve the performance of the two-stage detector by adding a quadratic constraint in the first stage synthesis, which exploits the knowledge of the spreading codes of the active users within the cell of interest. It is shown analytically that incorporation of such a quadratic constraint improves the receiver robustness against errors in the estimated statistics of the received data, although it slightly reduces the interference suppression capabilities of the two-stage detector. The effectiveness of the proposed receiver is further corroborated by computer simulation results.

Index Terms. Blind multiuser detection, linearly and quadratically constrained optimization, quasi-synchronous systems, multicarrier CDMA systems.

1 Introduction

The wideband direct-sequence code-division multiple-access (DS-CDMA) technique has emerged in recent years as the preferred air interface for providing voice and multimedia services in third-generation mobile communications. However, the use of DS-CDMA technology does not seem to be realistic [1] for very high data-rate multimedia services (at speeds of the order of several hundred megabits per second), due to the severe multipath-induced interchip and intersymbol interference, as well as because of synchronization difficulties. In order to alleviate the previous drawbacks, a great bulk of research activities has focused on the multicarrier CDMA technology [2], which integrates the advantages of multicarrier transmission systems, such as orthogonal frequency-division multiplexing (OFDM), with those of DS-CDMA. As discussed in [2], multicarrier CDMA systems can be categorized in two major types, according to whether the code spreading is performed

*The authors are with Dipartimento di Ingegneria Elettronica e delle Telecomunicazioni, Università degli Studi di Napoli Federico II, via Claudio 21, I-80125 Napoli, Italy. Tel.: +39-0817683154, Fax : +39-0817683149, E-mail: {gelli, paura,f.verde}@unina.it

in the time or frequency domain. The multicarrier CDMA system considered in this paper, generally referred to as MC-CDMA and originally proposed in [3], is based on frequency-domain spreading, which consists of copying each information symbol over the N subcarriers and multiplying it by a user-specific code. Besides representing an inherent form of frequency diversity, transmission over the N subcarriers allows one to cope with interchip and intersymbol interference more effectively than in DS-CDMA systems, by lowering the data-rate by serial-to-parallel conversion and introducing a cyclic prefix (CP) in the transmitted data. Additionally, since the symbol rate on each subcarrier is much lower than the chip-rate in a DS-CDMA system with comparable processing gain, the synchronization task is easier in MC-CDMA and, therefore, it is reasonable to consider a quasi-synchronous (QS) uplink [4, 5], with a beneficial impact on system performance and capacity.

Early papers on MC-CDMA reception [3, 6] deal with synchronous downlink transmission, wherein the receiver can be implemented by means of simple *diversity-combining strategies* [7], such as orthogonal restoring combining (ORC), equal gain combining (EGC), maximal ratio combining (MRC), or minimum mean-square error combining (MMSEC) (see also [8]). In addition to the knowledge of the spreading code and timing of the user to be demodulated, the ORC, MRC and MMSEC receivers require also the knowledge of the corresponding channel impulse response. When employed in the asynchronous uplink channel, MC-CDMA with these simple diversity-combining strategies can still perform better [9] than both DS-CDMA with a comparable value of the processing gain and RAKE reception, and multicarrier CDMA schemes with time-domain spreading. However, due to the presence of severe multiple-access interference (MAI), diversity-combining schemes tend to exhibit exceedingly large values of the bit-error rate (BER) floor in certain scenarios, even for a QS uplink [8]. To drastically improve the performance in this case, more sophisticated reception strategies, such as multiuser detection (MUD) techniques, are needed. Among these, the use of a linear MMSE receiver was originally proposed in [6, 10] to mitigate MAI in the synchronous downlink of a MC-CDMA system; in the asynchronous uplink scenario, this detector significantly outperforms [11] all the diversity-combining schemes, requiring the same *a priori* information (i.e., code, timing, and channel of each user to be demodulated) with a slightly increased complexity. A fractionally-spaced version of the MMSE (FS-MMSE) receiver, which does not require timing information, is proposed in [11], at the price of a further increased complexity over the MMSE detector, while still requiring the knowledge of the desired channel impulse response.

Most of the above-mentioned diversity-combining and MMSE MUD techniques rely on channel estimation, which can be performed by resorting to bandwidth-consuming training sequences. To avoid waste of resources, a subspace-based *blind* (i.e., without requiring training sequences) version of the linear MMSE receiver for a QS MC-CDMA system is proposed in [12], where the channel of the desired user is estimated on the basis of the eigenstructure properties of the received autocorrelation matrix; such a receiver belongs to the class of

indirect blind MUD techniques, where the channel is first estimated and then the estimate is plugged into the corresponding non-blind detector. By extending some of the concepts originally proposed in [14], which have proven fruitful in the area of joint multiuser detection and equalization in asynchronous DS-CDMA systems, a *direct* MUD technique is proposed in [13], where the detector's parameters are extracted from the received data without performing an explicit channel identification. This receiver consists of two stages: the former performs a suitably pre-filtering of the received signal, in order to mitigate MAI; the latter exploits the constant modulus (CM) property of the transmitted symbol sequence to recover the desired signal. Since the direct two-stage receiver requires the only knowledge of the code of the desired user, it is a blind and delay-independent MUD technique.

In this paper, with reference to the QS uplink of a MC-CDMA system, we first provide the analytical performance assessment of the direct blind MUD two-stage receiver proposed in [13], aimed at evaluating the performance degradation, in terms of signal-to-interference-plus-noise ratio (SINR) at the output of the first stage, when the receiver is implemented by using a finite data record. The analysis allows one to identify sufficient conditions assuring that the second stage, based on the CM, converges to the extraction of the desired symbol. The analysis, moreover, allows one to derive a new optimization criterion, aimed at improving the robustness of the two-stage receiver when it is implemented by using very short data records. The new criterion is based on the assumption, which is reasonable in the uplink, that the base station receiver has knowledge not only of the desired spreading code, but also of the spreading codes of a group of users, e.g., the users within its cell. This same assumption, considered in the context of DS-CDMA systems, leads to the synthesis of the so-called *group-blind receivers* [15, 16]; although in principle these receivers could be extended to the MC-CDMA case, they would fall into the class of indirect methods, wherein channel identification is first performed for *all* the known users, by a costly eigendecomposition; moreover they would require oversampling the received signal and/or employing an array of sensor at the receiver. Since our method, instead, is a direct one, it does not require any explicit eigenstructure-based channel estimation step; moreover, it does not require oversampling and/or multiple sensors at the receiver, hence it is inherently simpler.

The paper is organized as follows. Section 2 introduces the basic signal model of the considered QS-MC-CDMA system. Section 3 briefly reviews the two-stage approach proposed in [13] and presents the performance analysis in terms of SINR at the output of both stages. Section 4 proposes and analyzes the robust version of the two-stage receiver. Section 5 is devoted to the numerical performance analysis carried out by means of Monte Carlo computer simulations. Finally, conclusions are drawn in Section 6.

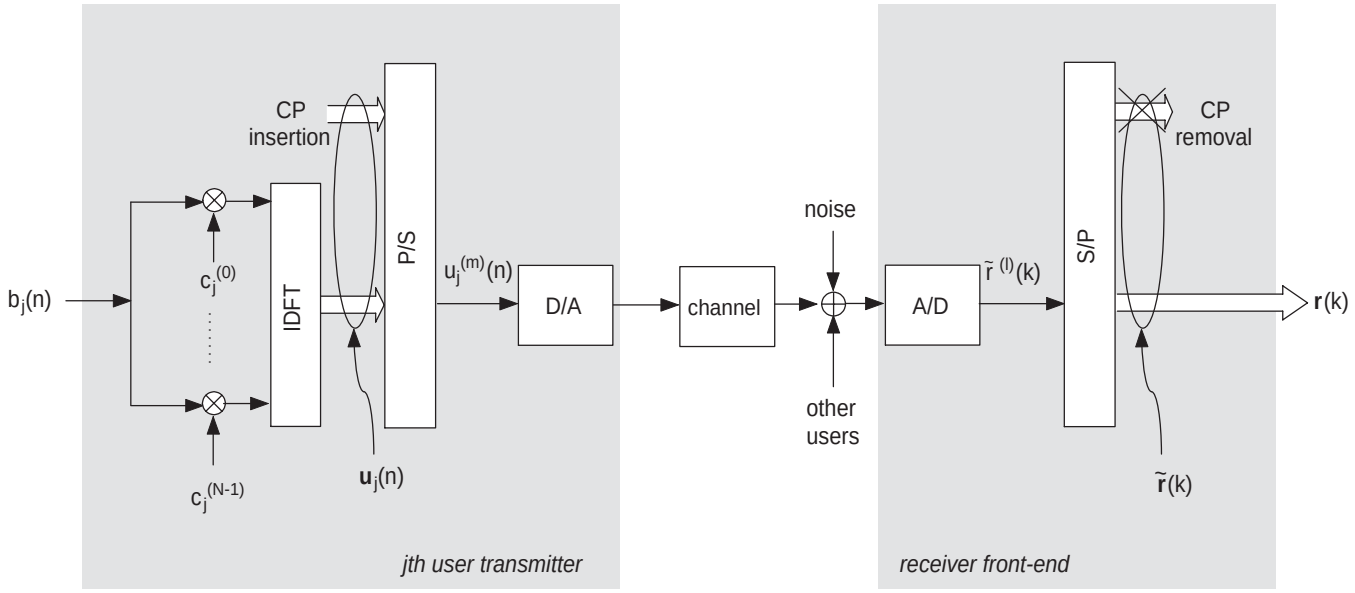

Figure 1: The considered MC-CDMA system.

2 The quasi-synchronous MC-CDMA uplink model

In the rest of the paper, we will use the following notations. Upper- and lower-case bold letters denote matrices and vectors, respectively; the superscripts $*$, T, H, -1 and $\dagger$ denote the conjugate, the transpose, the conjugate transpose, the inverse and the Moore-Penrose inverse of a matrix, respectively; $\mathbb{C}$, $\mathbb{R}$ and $\mathbb{Z}$ are the fields of complex, real and integer numbers, respectively; $\mathbb{C}_r^n$ and $\mathbb{R}_r^n$ [$\mathbb{C}^n$ and $\mathbb{R}^n$] denote the vector-spaces of all n-column random [deterministic] vectors with complex and real coordinates, respectively; similarly, $\mathbb{C}_r^{n\times m}$ and $\mathbb{R}_r^{n\times m}$ [$\mathbb{C}^{n\times m}$ and $\mathbb{R}^{n\times m}$] denote the vector-spaces of all the $n\times m$ random [deterministic] matrices with complex and real elements, respectively; $\mathbf{0}_n$, $\mathbf{O}_{n\times m}$ and $\boldsymbol{I}_n$ denote the n-column zero vector, the $n\times m$ zero and $n\times n$ identity matrices, respectively; $\mathrm{trace}(\boldsymbol{A})$ denotes the trace of a square matrix $\boldsymbol{A}$; $\mathrm{rank}(\boldsymbol{A})$ and $\mathcal{R}(\boldsymbol{A})$ denote the rank and the column space of any matrix $\boldsymbol{A}$; $\langle \boldsymbol{A}, \boldsymbol{B}\rangle \stackrel{\triangle}{=} \mathrm{trace}(\boldsymbol{A}\,\boldsymbol{B}^H)$ will denote the inner product in $\mathbb{C}^{n\times m}$ and $\|\boldsymbol{A}\| \stackrel{\triangle}{=} \sqrt{\mathrm{trace}(\boldsymbol{A}\boldsymbol{A}^H)}$ the induced (Frobenius) norm; $\boldsymbol{A} = \mathrm{diag}\,[\boldsymbol{A}_{11}, \boldsymbol{A}_{22}, \ldots, \boldsymbol{A}_{nn}]$ is the block diagonal matrix wherein $\{\boldsymbol{A}_{ii}\}_{i=1}^n$ are diagonal matrices; the subscript c stands for continuous-time signal and $\mathrm{E}[\cdot]$ denotes statistical averaging; and, finally, $\star$ and $i \stackrel{\triangle}{=} \sqrt{-1}$ denote (linear) convolution and imaginary unit, respectively.

Let us consider (see Fig. 1) the baseband-equivalent of a MC-CDMA uplink with N subcarriers. 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 $\boldsymbol{c}_j \stackrel{\triangle}{=} [c_j^{(0)}, c_j^{(1)}, \ldots, c_j^{(N-1)}]^T \in \mathbb{C}^N$; the resulting N-length sequence is subject to the inverse discrete Fourier transform (IDFT), producing thus the N-dimensional block $\widetilde{\boldsymbol{u}}_j(n) = \boldsymbol{W}_{\mathrm{IDFT}}\,\boldsymbol{c}_j\,b_j(n)$,

where $\boldsymbol{W}_{\rm IDFT} \in \mathbb{C}^{N\times N}$ denotes the *unitary symmetric* IDFT matrix, with (ξ,η)th entry $\{\boldsymbol{W}_{\rm IDFT}\}_{(\xi,\eta)} \stackrel{\triangle}{=} \frac{1}{\sqrt{N}}\cdot e^{i\,\frac{2\pi}{N}\,\xi\,\eta}$, for $\xi,\eta\in\{0,1,\ldots,N-1\}$. After computing the IDFT, a cyclic prefix (CP) of length $L_{\rm cp}\ll N$, consisting of a replica of the last $L_{\rm cp}$ symbols of $\widetilde{\boldsymbol{u}}_j(n)$, is inserted at the beginning of $\widetilde{\boldsymbol{u}}_j(n)$, obtaining thus the vector $\boldsymbol{u}_j(n)=\boldsymbol{T}_{\rm cp}\,\boldsymbol{W}_{\rm IDFT}\,\boldsymbol{c}_j\,b_j(n)$, where $P\stackrel{\triangle}{=}L_{\rm cp}+N$ and $\boldsymbol{T}_{\rm cp}\stackrel{\triangle}{=}[\boldsymbol{I}_{\rm cp}^T,\boldsymbol{I}_N]^T\in\mathbb{R}^{P\times N}$, with $\boldsymbol{I}_{\rm cp}\in\mathbb{R}^{L_{\rm cp}\times N}$ obtained by drawing out the last $L_{\rm cp}$ rows of the identity matrix $\boldsymbol{I}_N$. The block $\boldsymbol{u}_j(n)$ is subject to parallel-to-serial (P/S) conversion, and the resulting sequence[1] $\{u_j^{(m)}(n)\}_{m=0}^{P-1}$ feeds a digital-to-analog (D/A) converter with impulse response $\psi_c(t)$, operating at rate $1/T_c=P/T_s$, where T_s and T_c denote the symbol and the sampling period, respectively. The continuous-time signal at the D/A output is therefore given by

$$u_{c,j}(t)=\sum_{n=-\infty}^{+\infty}\sum_{m=0}^{P-1}u_j^{(m)}(n)\,\psi_c(t-n\,T_s-m\,T_c-\tau_j)\,, \tag{1}$$

where $\tau_j=d_j\,T_c+\beta_j$, with $d_j\in\{0,1,\ldots,P-1\}$ and $\beta_j\in[\,0,T_c)$, represents the transmission delay of the jth user. The signal (1) is transmitted over a multipath channel modeled as a linear time-invariant[2] system with impulse response $h_{c,j}(t)$. Denoting with $\phi_c(t)$ the impulse response of the receiving filter and assuming that ideal carrier-frequency recovery is carried out at the receiver, the (overall) received baseband signal in the uplink channel (i.e., mobile to base station) can be expressed as

$$r_c(t)=\sum_{j=1}^{J}\sum_{n=-\infty}^{+\infty}\sum_{m=0}^{P-1}u_j^{(m)}(n)\,g_{c,j}(t-n\,T_s-m\,T_c-\tau_j)+v_c(t)\,, \tag{2}$$

where J is the number of users picked up by the base-station receiver, $g_{c,j}(t)\stackrel{\triangle}{=}\psi_c(t)\star h_{c,j}(t)\star\phi_c(t)$ is the impulse response (including transmitting filter, physical channel and receiving filter) of the *composite* channel of the jth user, and $v_c(t)$ represents the additive noise at the output of the receiving filter. The following assumptions will be considered throughout the paper: **A1)** the information symbols $b_j(n)$ are mutually independent zero-mean and independent identically-distributed (iid) sequences, with equal variance $\sigma_b^2\stackrel{\triangle}{=}\mathrm{E}[|b_j(k)|^2]$; **A2)** the additive noise $v_c(t)$ is a zero-mean wide-sense stationary complex proper process, which is independent of the sequences $b_j(n)$, for $j\in\{1,2,\ldots,J\}$; **A3)** the composite channel impulse response $g_{c,j}(t)$ of the jth user spans L_j sampling periods, i.e., $g_{c,j}(t)\equiv 0$, for $t\notin[\,0,L_j\,T_c)$, with L_j within one symbol interval[3], that is, $L_j<P$. To demodulate the kth block ($k\in\mathbb{Z}$), the received signal $r_c(t)$ is sampled at the time epochs

[1] To avoid notational complications, we denote with $u_j^{(m)}(n)$ the $(m+1)$th component of vector $\boldsymbol{u}_j(n)$, for $m=0,1,\ldots,P-1$.

[2] This assumption is common in high data-rate multicarrier systems (envisioned to support broadband multimedia services) in which the channel time-selectivity can be neglected for several consecutive symbol intervals [17].

[3] If the channel support L_j exceeds one symbol period, the information symbol $b_j(n)$ needs first to be serial-to-parallel (S/P) converted before spreading over the frequency domain [8]. However, for the sake of clearness, we prefer to assume that $L_j<P$, since the proposed method can be straightforwardly extended to take into account such a S/P conversion of the original information stream.

$t_{k,\ell} \triangleq k\,T_s + \ell\,T_c$, with $\ell \in \{0, 1, \ldots, P-1\}$, yielding [see equation (2)]

$$\widetilde{r}^{(\ell)}(k) \triangleq r_c(t_{k,\ell}) = \sum_{j=1}^{J} \sum_{n=-\infty}^{+\infty} \sum_{m=0}^{P-1} u_j^{(m)}(n)\, g_j[(k-n)\,P + (\ell - m) - d_j] + \widetilde{v}^{(\ell)}(k)\,, \tag{3}$$

where $g_j(k) \triangleq g_{c,j}(k\,T_c - \beta_j)$ and $\widetilde{v}^{(\ell)}(k) \triangleq v_c(t_{k,\ell})$. Observe that the channel frequency-selectivity introduces two impairments in the demodulation of the kth block of each user: (i) the interblock interference (IBI), which is represented by the terms with $n \neq k$ in (3); (ii) the intercarrier interference (ICI), which is generated by the terms with $m \neq \ell$ in (3). However, as a consequence of assumption A3, the discrete-time channel $g_j(k)$ turns out to be a causal finite-impulse response filter of order L_j, i.e., $g_j(k) \equiv 0$ for $k \notin \{\,0, 1, \ldots, L_j\,\}$. Therefore, only the terms in (3) with $n \in \{k-2, k-1, k\}$ contribute to the IBI of each user, that is, for $\ell \in \{0, 1, \ldots, P-1\}$,

$$\widetilde{r}^{(\ell)}(k) = \sum_{j=1}^{J} \sum_{p=0}^{2} \sum_{q=\ell-P+1}^{\ell} g_j^{(q)}(p)\, u_j^{(\ell-q)}(k-p) + \widetilde{v}^{(\ell)}(k)\,, \tag{4}$$

where, for mathematical convenience, we have defined the (fictitious) subchannels $g_j^{(q)}(p) \triangleq g_j(p\,P + q - d_j)$. A more compact matrix-vector model can be obtained by collecting the P different samples $\{\widetilde{r}^{(\ell)}(k)\}_{\ell=0}^{P-1}$ in the vector $\widetilde{\boldsymbol{r}}(k) \triangleq [\widetilde{r}^{(0)}(k), \widetilde{r}^{(1)}(k), \ldots, \widetilde{r}^{(P-1)}(k)]^T \in \mathbb{C}^P$, obtaining thus

$$\widetilde{\boldsymbol{r}}(k) = \sum_{j=1}^{J} \sum_{p=0}^{2} \widetilde{\boldsymbol{G}}_j(p)\, \boldsymbol{T}_{\text{cp}}\, \boldsymbol{W}_{\text{IDFT}}\, \boldsymbol{c}_j\, b_j(k-p) + \widetilde{\boldsymbol{v}}(k)\,, \tag{5}$$

where $\widetilde{\boldsymbol{v}}(k) \triangleq [\tilde{v}^{(0)}(k), \tilde{v}^{(1)}(k), \ldots, \tilde{v}^{(P-1)}(k)]^T \in \mathbb{C}^P$ is the noise vector and we have defined the *Toeplitz matrices* (see [18])

$$\widetilde{\boldsymbol{G}}_j(p) = \sum_{h=0}^{P-1} g_j^{(h)}(p)\, \widetilde{\boldsymbol{F}}^{\,h} + \sum_{h=1}^{P-1} g_j^{(-h)}(p)\, \widetilde{\boldsymbol{B}}^{\,h}\,, \tag{6}$$

for $p \in \{0, 1, 2\}$ and $j \in \{1, 2, \ldots, J\}$, with $\widetilde{\boldsymbol{F}}^{\,h}$ and $\widetilde{\boldsymbol{B}}^{\,h}$ denoting the hth power of the Toeplitz *forward shift* $\widetilde{\boldsymbol{F}} \in \mathbb{R}^{P\times P}$ and *backward shift* $\widetilde{\boldsymbol{B}} \in \mathbb{R}^{P\times P}$ matrices, where the first column of $\widetilde{\boldsymbol{F}}$ and the first row of $\widetilde{\boldsymbol{B}}$ are given by $[0, 1, 0, \ldots, 0]^T$ and $[0, 1, 0, \ldots, 0]$, respectively, with $\widetilde{\boldsymbol{F}}^{\,0} \triangleq \boldsymbol{I}_P$. Note that, for $h \in \{1, 2, \ldots, P-1\}$, the first h rows of $\widetilde{\boldsymbol{F}}^{\,h}$ and the last h rows of $\widetilde{\boldsymbol{B}}^{\,h}$ are identically zero.

In the sequel, we assume that, without loss of generality, the *desired user* is the first one ($j = 1$) and that, with reference to the uplink of a QS-MC-CDMA system [4, 5, 12], the first J_{in} out of J users are within the cell of interest (referred to as *in-cell users*) and attempt to synchronize[4] their transmissions by resorting to a local

[4]It should be observed that, due to oscillator drifts, GPS uncertainties, and the relative motion among the mobiles and the base station, the in-cell users signals received by the base station are still asynchronous, even though their asynchronisms are contained within a small number of sampling intervals.

reference clock (obtained, e.g., with the help of a GPS device) or to a pilot signal transmitted by the base station, whereas the remaining $J_{\rm out} \triangleq J - J_{\rm in}$ users are outside the cell of interest (referred to as *out-of-cell users*). Moreover, according to [17, 19], we reasonably assume that: **A4)** the CP length $L_{\rm cp}$ satisfies the inequality $L_{\rm cp} \geq \max_{j\in\{1,2,\ldots,J_{\rm in}\}} \big[L_j + d_j + 1\big]$; under this assumption, it can be shown that, by exploiting the structure of the matrices $\widetilde{\boldsymbol{G}}_j(p)$, for $p \in \{0,1,2\}$, the IBI contribution for *each* in-cell user can be completely discarded by dropping the first $L_{\rm cp}$ components of $\widetilde{\boldsymbol{r}}(k)$. This operation can be accomplished in matrix form by defining the matrix $\boldsymbol{R}_{\rm cp} \triangleq [\mathbf{O}_{N\times L_{\rm cp}}, \boldsymbol{I}_N] \in \mathbb{R}^{N\times P}$ and forming at the receiver the product $\boldsymbol{r}(k) \triangleq \boldsymbol{R}_{\rm cp}\,\widetilde{\boldsymbol{r}}(k) \in \mathbb{C}^N$. According to A4, it results that $\boldsymbol{R}_{\rm cp}\,\widetilde{\boldsymbol{G}}_j(1) = \mathbf{O}_{N\times P}$, for $j \in \{1,2,\ldots,J_{\rm in}\}$, which, in its turn, implies that, after CP removal, the received signal is given by

$$\boldsymbol{r}(k) = \underbrace{\boldsymbol{G}_1(0)\,\boldsymbol{W}_{\rm IDFT}\,\boldsymbol{c}_1\,b_1(k)}_{\text{desired user}} + \underbrace{\sum_{j=2}^{J_{\rm in}} \boldsymbol{G}_j(0)\,\boldsymbol{W}_{\rm IDFT}\,\boldsymbol{c}_j\,b_j(k)}_{\text{in-cell MAI}} + \\ + \underbrace{\sum_{j=J_{\rm in}+1}^{J}\sum_{p=0}^{2} \boldsymbol{G}_j(p)\,\boldsymbol{W}_{\rm IDFT}\,\boldsymbol{c}_j\,b_j(k-p)}_{\text{out-of-cell MAI}} + \underbrace{\boldsymbol{v}(k)}_{\text{noise}}, \quad (7)$$

where $\boldsymbol{G}_j(p) \triangleq \boldsymbol{R}_{\rm cp}\,\widetilde{\boldsymbol{G}}_j(p)\,\boldsymbol{T}_{\rm cp} \in \mathbb{C}^{N\times N}$, for $p \in \{0,1,2\}$ and $j \in \{1,2,\ldots,J\}$, and $\boldsymbol{v}(k) \triangleq \boldsymbol{R}_{\rm cp}\,\widetilde{\boldsymbol{v}}(k) \in \mathbb{C}^N$. Moreover, the *signatures* $\boldsymbol{G}_j(0)\,\boldsymbol{W}_{\rm IDFT}\,\boldsymbol{c}_j$ of the in-cell users can be parameterized as (see [12] for details)

$$\boldsymbol{G}_j(0)\,\boldsymbol{W}_{\rm IDFT}\,\boldsymbol{c}_j = \boldsymbol{\mathcal{C}}_j\,\boldsymbol{Q}_j\,\boldsymbol{g}_j\,, \quad \text{for} \quad j \in \{1,2,\ldots,J_{\rm in}\}\,, \quad (8)$$

where $\boldsymbol{\mathcal{C}}_j \triangleq \sqrt{N}\cdot\boldsymbol{W}_{\rm IDFT}\,\boldsymbol{C}_j\,\boldsymbol{W}_{\rm DFT} \in \mathbb{C}^{N\times N}$ is the nonsingular *code matrix*, with $\boldsymbol{W}_{\rm DFT} \triangleq \boldsymbol{W}_{\rm IDFT}^H$ denoting the discrete Fourier transform (DFT) unitary symmetric matrix and $\boldsymbol{C}_j \triangleq {\rm diag}[c_j^{(0)}, c_j^{(1)}, \ldots, c_j^{(N-1)}] \in \mathbb{C}^{N\times N}$, and we have defined the full-column rank matrix $\boldsymbol{Q}_j \triangleq [\mathbf{O}_{d_j\times(L_j+1)}^T, \boldsymbol{I}_{L_j+1}, \mathbf{O}_{(N-L_j-d_j-1)\times(L_j+1)}^T]^T \in \mathbb{R}^{N\times(L_j+1)}$, which accounts for the *unknown* delay d_j, and the vector $\boldsymbol{g}_j \triangleq [g_j(0), g_j(1), \ldots, g_j(L_j)]^T \in \mathbb{C}^{L_j+1}$, which collects the *unknown* channel coefficients. Finally, by substituting (8) in (7), we obtain

$$\boldsymbol{r}(k) = \boldsymbol{\mathcal{C}}_1\,\boldsymbol{Q}_1\,\boldsymbol{g}_1\,b_1(k) + \boldsymbol{d}(k)\,, \quad (9)$$

where the N-column vector

$$\boldsymbol{d}(k) \triangleq \boldsymbol{\mathcal{H}}_{\rm in}\,\boldsymbol{b}_{\rm in}(k) + \boldsymbol{\mathcal{H}}_{\rm out}\,\boldsymbol{b}_{\rm out}(k) + \boldsymbol{v}(k) \quad (10)$$

represents the overall disturbance (MAI plus noise), the vectors $\boldsymbol{b}_{\rm in}(k) \triangleq [b_2(k), b_3(k), \ldots, b_{J_{\rm in}}(k)]^T \in \mathbb{C}^{J_{\rm in}-1}$, and $\boldsymbol{b}_{\rm out}(k) \triangleq [b_{J_{\rm in}+1}(k), b_{J_{\rm in}+1}(k-1), b_{J_{\rm in}+1}(k-2), \ldots, b_J(k), b_J(k-1), b_J(k-2)]^T \in \mathbb{C}^{3J_{\rm out}}$, and the matrices $\boldsymbol{\mathcal{H}}_{\rm in} \triangleq [\boldsymbol{\mathcal{H}}_2, \boldsymbol{\mathcal{H}}_3, \ldots, \boldsymbol{\mathcal{H}}_{J_{\rm in}}] \in \mathbb{C}^{N\times(J_{\rm in}-1)}$ and $\boldsymbol{\mathcal{H}}_{\rm out} \triangleq [\boldsymbol{\mathcal{H}}_{J_{\rm in}+1}, \boldsymbol{\mathcal{H}}_{J_{\rm in}+2}, \ldots, \boldsymbol{\mathcal{H}}_J] \in \mathbb{C}^{N\times(3J_{\rm out})}$ collect all the

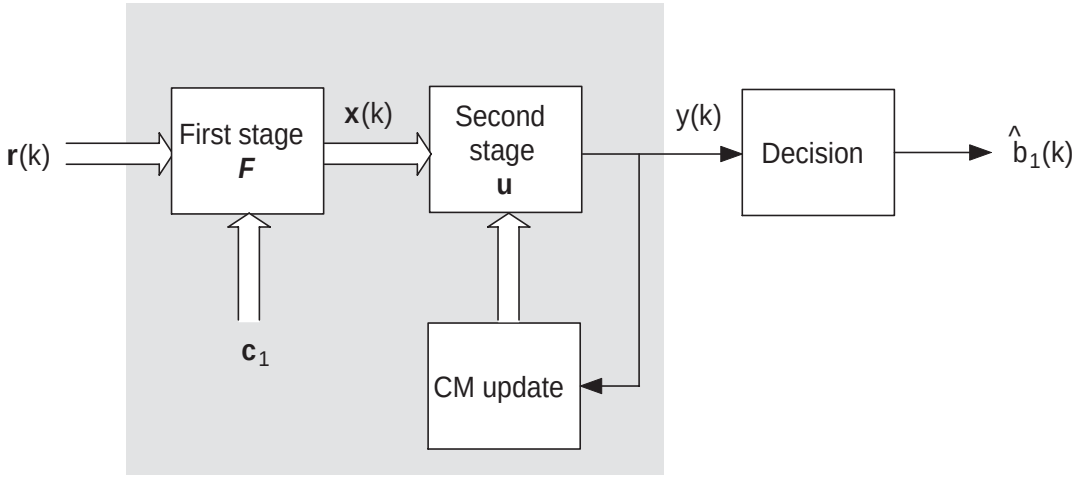

Figure 2: The two-stage receiver structure.

interfering symbols and signatures of the in-cell and out-of-cell users, respectively, with $\boldsymbol{\mathcal{H}}_j \stackrel{\triangle}{=} \boldsymbol{\mathcal{C}}_j \, \boldsymbol{Q}_j \, \boldsymbol{g}_j \in \mathbb{C}^N$, for $j = 2, 3, \ldots, J_{\text{in}}$, whereas $\boldsymbol{\mathcal{H}}_j \stackrel{\triangle}{=} [\boldsymbol{G}_j(0) \, \boldsymbol{W}_{\text{IDFT}} \, \boldsymbol{c}_j, \boldsymbol{G}_j(1) \, \boldsymbol{W}_{\text{IDFT}} \, \boldsymbol{c}_j, \boldsymbol{G}_j(2) \, \boldsymbol{W}_{\text{IDFT}} \, \boldsymbol{c}_j] \in \mathbb{C}^{N \times 3}$, for $j = J_{\text{in}} + 1, J_{\text{in}} + 2, \ldots, J$.

Some comments are now in order about model (9). First, observe that, since the out-of-cell users are quasi-synchronous with respect to a *different* base station, the CP removal does not assure the complete elimination of their IBI. Moreover, note that assumption A4 requires *only* upper bounds (rather than the exact knowledge) on the channel orders and delays of the in-cell users. This is a reasonable assumption in the considered scenario since: (i) in general, depending on the transmitted signal parameters (carrier frequency and bandwidth) and application (indoor or outdoor), the maximum channel multipath spread is known; (ii) for QS cellular systems, the delays of the in-cell users are confined to a small uncertainty interval, whose support can be typically predicted [19].

3 Performance analysis of the blind two-stage receiver

This section provides a detailed analysis of the two-stage detector (see Fig. 2) recently proposed in [13]. In particular, our analysis consists of two steps: firstly, we present an analysis of the SINR at the output of the first stage, when the receiver's parameters are computed from K samples of the received vector $\boldsymbol{r}(k)$; secondly, we investigate the relationship between the potential for "interference capture" of the CM-based second stage and the SINR at the output of the first stage. To put the basis, we briefly review in Subsection 3.1 the two-stage approach of [13].

3.1 The blind two-stage receiver

In the framework of linear blind and delay-independent MUD, the problem of detecting the desired user symbol $b_1(k)$ consists of synthesizing, without requiring knowledge of the timings and channel impulse responses of all the active users (included the desired one), a linear filter $\boldsymbol{f} \in \mathbb{C}^N$, whose output $y(k) = \boldsymbol{f}^H \boldsymbol{r}(k)$ represents a soft estimate of $b_1(k)$. The two-stage detector (see Fig. 2) proposed in [13] for QS-MC-CDMA systems is based on factorizing the overall receiver weight vector as $\boldsymbol{f} = \boldsymbol{\mathcal{F}}\,\boldsymbol{u}$, where the weight vector $\boldsymbol{u} \in \mathbb{C}^{L_{\rm cp}}$ in the second stage is determined according to the well-known CM criterion (see, e.g., [20])

$$\boldsymbol{u}_{\rm opt} = \arg\min_{\boldsymbol{u}} \mathrm{E}[(\gamma - |\boldsymbol{u}^H\,\boldsymbol{x}(k)|^2)^2]\,, \tag{11}$$

with $\gamma \triangleq \mathrm{E}[|b_1(k)|^4]/\sigma_b^2$ being the second-order dispersion coefficient of the desired symbol sequence $b_1(k)$; whereas the output of the first stage $\boldsymbol{x}(k)$ is a linear transformation of $\boldsymbol{r}(k)$, that is, $\boldsymbol{x}(k) = \boldsymbol{\mathcal{F}}^H \boldsymbol{r}(k)$, which, accounting for (9), can be expressed by means of the coincise vector model

$$\boldsymbol{x}(k) = \boldsymbol{\mathcal{F}}^H \boldsymbol{\mathcal{C}}_1\,\boldsymbol{Q}_1\,\boldsymbol{g}_1\,b_1(k) + \boldsymbol{\mathcal{F}}^H \boldsymbol{d}(k)\,. \tag{12}$$

Moreover, by observing that, under assumption A4, for the in-cell users, $\boldsymbol{Q}_j$ can be factorized as

$$\boldsymbol{Q}_j = \boldsymbol{\Pi}\,\boldsymbol{Q}_{j,2}\,, \quad \text{for} \quad j \in \{1, 2, \ldots, J_{\rm in}\}\,, \tag{13}$$

where the full-column rank matrix $\boldsymbol{\Pi} \triangleq [\boldsymbol{I}_{L_{\rm cp}}, \mathbf{O}^T_{(N-L_{\rm cp})\times L_{\rm cp}}]^T \in \mathbb{R}^{N\times L_{\rm cp}}$ is completely known at the receiving side, whereas the matrix $\boldsymbol{Q}_{j,2} \triangleq [\mathbf{O}^T_{d_j\times(L_j+1)}, \boldsymbol{I}_{L_j+1}, \mathbf{O}^T_{(L_{\rm cp}-L_j-d_j-1)\times(L_j+1)}]^T \in \mathbb{R}^{L_{\rm cp}\times(L_j+1)}$ is unknown (it depends on L_j and d_j), equation (12) can be rewritten as

$$\boldsymbol{x}(k) = \boldsymbol{\mathcal{F}}^H \boldsymbol{\mathcal{C}}_1\,\boldsymbol{\Pi}\,\boldsymbol{Q}_{1,2}\,\boldsymbol{g}_1\,b_1(k) + \boldsymbol{\mathcal{F}}^H \boldsymbol{d}(k) = \boldsymbol{\mathcal{F}}^H \boldsymbol{\Upsilon}_1\,\overline{\boldsymbol{g}}_1\,b_1(k) + \boldsymbol{\mathcal{F}}^H \boldsymbol{d}(k) \tag{14}$$

where $\boldsymbol{\Upsilon}_j \triangleq \boldsymbol{\mathcal{C}}_j\,\boldsymbol{\Pi} \in \mathbb{C}^{N\times L_{\rm cp}}$ is a known matrix and $\overline{\boldsymbol{g}}_j \triangleq \boldsymbol{Q}_{j,2}\,\boldsymbol{g}_j \in \mathbb{C}^{L_{\rm cp}}$ is the unknown signature of the jth in-cell user, for $j \in \{1, 2, \ldots, J_{\rm in}\}$.

A careful choice of $\boldsymbol{\mathcal{F}} \in \mathbb{C}^{N\times L_{\rm cp}}$ must assure MAI-plus-noise mitigation at the input of the second stage, so as to avoid the interference capture phenomenon [21] typical of the CM criterion. Such a choice is pursued in [13, 14] by solving the following linearly constrained optimization problem

$$\boldsymbol{\mathcal{F}}_{\rm opt} = \arg\min_{\boldsymbol{\mathcal{F}}} \mathrm{E}[\|\boldsymbol{x}(k)\|^2]\,, \quad \text{subject to} \quad \boldsymbol{\mathcal{F}}^H \boldsymbol{\Upsilon}_1 = \boldsymbol{I}_{L_{\rm cp}}\,, \tag{15}$$

where the *linear* matrix constraint is aimed at preserving the desired symbol $b_1(k)$ and does not require neither channel nor timing knowledge. The solution of (15) can be canonically decomposed [14] as

$$\boldsymbol{\mathcal{F}}_{\rm opt} = \boldsymbol{\mathcal{F}}^{(0)}_{\rm opt} - \boldsymbol{B}_1\,\boldsymbol{\mathcal{F}}^{(a)}_{\rm opt}\,, \tag{16}$$

where $\boldsymbol{\mathcal{F}}_{\text{opt}}^{(0)} \triangleq (\boldsymbol{\Upsilon}_1^{\dagger})^H = \boldsymbol{\Upsilon}_1 (\boldsymbol{\Upsilon}_1^H \boldsymbol{\Upsilon}_1)^{-1}$, the matrix $\boldsymbol{B}_1 \in \mathbb{C}^{N \times (N-L_{\text{cp}})}$ satisfies the following relations $\boldsymbol{B}_1^H \boldsymbol{\Upsilon}_1 = \mathbf{O}_{(N-L_{\text{cp}}) \times L_{\text{cp}}}$ and $\boldsymbol{B}_1^H \boldsymbol{B}_1 = \boldsymbol{I}_{N-L_{\text{cp}}}$, and

$$\boldsymbol{\mathcal{F}}_{\text{opt}}^{(a)} = (\boldsymbol{B}_1^H \boldsymbol{R}_{\boldsymbol{rr}} \boldsymbol{B}_1)^{-1} \boldsymbol{B}_1^H \boldsymbol{R}_{\boldsymbol{rr}} \boldsymbol{\mathcal{F}}_{\text{opt}}^{(0)} = (\boldsymbol{B}_1^H \boldsymbol{R}_{\boldsymbol{dd}} \boldsymbol{B}_1)^{-1} \boldsymbol{B}_1^H \boldsymbol{R}_{\boldsymbol{dd}} \boldsymbol{\mathcal{F}}_{\text{opt}}^{(0)} , \tag{17}$$

with $\boldsymbol{R}_{\boldsymbol{rr}} \triangleq \text{E}[\boldsymbol{r}(k)\, \boldsymbol{r}^H(k)] \in \mathbb{C}^{N \times N}$ and $\boldsymbol{R}_{\boldsymbol{dd}} \triangleq \text{E}[\boldsymbol{d}(k)\, \boldsymbol{d}^H(k)] \in \mathbb{C}^{N \times N}$ being the statistical correlation matrices of $\boldsymbol{r}(k)$ and $\boldsymbol{d}(k)$, respectively. We will refer to the receiver based on (16)–(17) as to the *optimal* two-stage receiver. Observe that while $\boldsymbol{\mathcal{F}}_{\text{opt}}^{(0)}$ depends only on the desired code and, thus, it can be evaluated off-line, $\boldsymbol{\mathcal{F}}_{\text{opt}}^{(a)}$ must be estimated from the received data, by resorting to a consistent estimate of $\boldsymbol{R}_{\boldsymbol{rr}}$. In this case, if one resorts to batch algorithms, the computational complexity of the first stage is basically dominated by the matrix inversion in (17), which is of order $\mathcal{O}[(N - L_{\text{cp}})^3]$. On the other hand, reasoning as in [14], the matrix $\boldsymbol{\mathcal{F}}_{\text{opt}}^{(a)}$ can also be estimated by means of a simple and effective recursion, similar to the well-known RLS algorithm, with a complexity per symbol interval of order only $\mathcal{O}[(N - L_{\text{cp}})^2]$.

The disturbance suppression capability of the optimal first stage (16) can be analyzed by following the guidelines given in [14], under the assumption that the noise $\boldsymbol{v}(k)$ is white with variance σ_v^2, i.e., $\boldsymbol{R}_{\boldsymbol{vv}} \triangleq \text{E}[\boldsymbol{v}(k)\, \boldsymbol{v}^H(k)] = \sigma_v^2\, \boldsymbol{I}_N$. It can be shown that, in the high signal-to-noise ratio (SNR) region, i.e., as $\sigma_v^2/\sigma_b^2 \rightarrow 0$, the filtering matrix $\boldsymbol{\mathcal{F}}_{\text{opt}}$ is able to achieve *perfect* disturbance cancellation if and only if: C1) $\mathcal{R}(\boldsymbol{\mathcal{C}}_1\, \boldsymbol{Q}_{1,1}) \cap \mathcal{R}(\boldsymbol{\mathcal{G}}) = \{\mathbf{0}_N\}$ or, equivalently, $\text{rank}(\boldsymbol{B}_1^H \boldsymbol{\mathcal{G}}) = \text{rank}(\boldsymbol{\mathcal{G}})$, where $\boldsymbol{\mathcal{G}} \triangleq [\boldsymbol{\mathcal{G}}_{\text{in}}, \boldsymbol{\mathcal{G}}_{\text{out}}] \in \mathbb{C}^{N \times D}$, with $D \triangleq J_{\text{in}} + 3\, J_{\text{out}} - 1$ representing the total number of MAI signatures [see the signal model (9)–(10)]. In this case, the first stage behaves as a *blind zero-forcing* detector. By using straightforward rank inequalities, it can be easily seen that the equality $\text{rank}(\boldsymbol{B}_1^H \boldsymbol{\mathcal{G}}) = \text{rank}(\boldsymbol{\mathcal{G}})$ requires that $N - L_{\text{cp}} \geq D$, that is, the number of *degrees of freedom* $N - L_{\text{cp}}$ for disturbance suppression must be greater than or equal to D.

3.2 Ideal performance analysis

A different measure of MAI-plus-noise suppression capability achieved by the first stage, which can be more directly related to the interference capture phenomenon [21] of the second stage, is the SINR at the output of the first stage, which, for an arbitrary $\boldsymbol{\mathcal{F}} \in \mathbb{C}^{N \times L_{\text{cp}}}$, is defined, on the basis of (12), as

$$\text{SINR}^{(\text{I})}(\boldsymbol{\mathcal{F}}) \triangleq \frac{\text{E}\left[\|\boldsymbol{\mathcal{F}}^H \boldsymbol{\Upsilon}_1\, \overline{\boldsymbol{g}}_1\, b_1(k)\|^2\right]}{\text{E}\left[\|\boldsymbol{\mathcal{F}}^H \boldsymbol{d}(k)\|^2\right]} = \frac{\sigma_b^2\, \|\boldsymbol{\mathcal{F}}^H \boldsymbol{\Upsilon}_1 \overline{\boldsymbol{g}}_1\|^2}{\text{trace}(\boldsymbol{\mathcal{F}}^H \boldsymbol{R}_{\boldsymbol{dd}}\, \boldsymbol{\mathcal{F}})} . \tag{18}$$

Since, from (14), one has $\text{E}[\|\boldsymbol{x}(k)\|^2] = \sigma_b^2 \|\boldsymbol{\mathcal{F}}^H \boldsymbol{\Upsilon}_1\, \overline{\boldsymbol{g}}_1\|^2 + \text{trace}(\boldsymbol{\mathcal{F}}^H \boldsymbol{R}_{\boldsymbol{dd}}\, \boldsymbol{\mathcal{F}})$, (18) can also be written as

$$\text{SINR}^{(\text{I})}(\boldsymbol{\mathcal{F}}) = \frac{\sigma_b^2\, \|\boldsymbol{\mathcal{F}}^H \boldsymbol{\Upsilon}_1\, \overline{\boldsymbol{g}}_1\|^2}{\text{E}[\|\boldsymbol{x}(k)\|^2] - \sigma_b^2 \|\boldsymbol{\mathcal{F}}^H \boldsymbol{\Upsilon}_1\, \overline{\boldsymbol{g}}_1\|^2} . \tag{19}$$

Therefore, maximizing $\text{SINR}^{(\text{I})}(\boldsymbol{\mathcal{F}})$ with the constraint $\boldsymbol{\mathcal{F}}^H \boldsymbol{\Upsilon}_1 = \boldsymbol{I}_{L_{\text{cp}}}$ amounts to minimizing $\text{E}[\|\boldsymbol{x}(k)\|^2]$ with the same constraint: hence, the maximum value of the (constrained) SINR at the output of the first stage can be obtained by substituting (16) in (18), or in (19), and is given by

$$\text{SINR}^{(\text{I})}_{\text{max}} \triangleq \text{SINR}^{(\text{I})}(\boldsymbol{\mathcal{F}}_{\text{opt}}) = \frac{\sigma_b^2 \, \|\overline{\boldsymbol{g}}_1\|^2}{\mathcal{P}_d} \, , \tag{20}$$

where $\mathcal{P}_d \triangleq \text{trace}\left[(\boldsymbol{\mathcal{F}}^{(0)}_{\text{opt}})^H \boldsymbol{R}_{\boldsymbol{dd}} \boldsymbol{\mathcal{F}}^{(0)}_{\text{opt}}\right] - \text{trace}\left[(\boldsymbol{\mathcal{F}}^{(0)}_{\text{opt}})^H \boldsymbol{R}_{\boldsymbol{dd}} \boldsymbol{B}_1 \, (\, \boldsymbol{B}_1^H \boldsymbol{R}_{\boldsymbol{dd}} \boldsymbol{B}_1 \,)^{-1} \boldsymbol{B}_1^H \boldsymbol{R}_{\boldsymbol{dd}} \boldsymbol{\mathcal{F}}^{(0)}_{\text{opt}}\right]$ represents the residual disturbance power at the output of the first stage.

Let us now focus attention on the interference capture of the CM-based filter employed in the second stage. To this end, we initially observe that, accounting for (14), the output of the second stage can be written as

$$y(k) = \boldsymbol{u}^H \boldsymbol{x}(k) = \boldsymbol{u}^H \boldsymbol{\mathcal{F}}^H \boldsymbol{\Upsilon}_1 \overline{\boldsymbol{g}}_1 \, b_1(k) + \boldsymbol{u}^H \boldsymbol{\mathcal{F}}^H \boldsymbol{d}(k) \tag{21}$$

and, thus, for a given $\boldsymbol{\mathcal{F}} \in \mathbb{C}^{N \times L_{\text{cp}}}$ and an arbitrary $\boldsymbol{u} \in \mathbb{C}^{L_{\text{cp}}}$, the SINR at the output of the second stage can be defined as

$$\text{SINR}^{(\text{II})}(\boldsymbol{u}) \triangleq \frac{\text{E}\left[|\boldsymbol{u}^H \boldsymbol{\mathcal{F}}^H \boldsymbol{\Upsilon}_1 \, \overline{\boldsymbol{g}}_1 \, b_1(k)|^2\right]}{\text{E}\left[|\boldsymbol{u}^H \boldsymbol{\mathcal{F}}^H \boldsymbol{d}(k)|^2\right]} = \frac{\sigma_b^2 \, |\boldsymbol{u}^H \boldsymbol{\mathcal{F}}^H \boldsymbol{\Upsilon}_1 \, \overline{\boldsymbol{g}}_1|^2}{\boldsymbol{u}^H \boldsymbol{\mathcal{F}}^H \boldsymbol{R}_{\boldsymbol{dd}} \boldsymbol{\mathcal{F}} \, \boldsymbol{u}} \, . \tag{22}$$

Since a closed-form expression for the solution of the minimization problem (11) is not available, the interference capture behavior of CM-based filters is typically studied by assuming that the gradient descent (GD) algorithm is employed to minimize the CM cost function. Along this line, Schniter and Johnson have derived in [21] a sufficient condition, expressed in terms of SINR, which assures that, in a noiseless multiuser scenario, the GD-based minimization of the CM cost function safely extracts the desired symbol. In the following, we recall this result (we refer to [21] for further details), particularizing it to our framework.

Theorem 1. Assume that, in addition to A1, the sequences $\{b_j(n)\}_{j=1}^J$ are proper, i.e., $\text{E}[b_j^2(n)] = 0$, for any $n \in \mathbb{N}$, and sub-Gaussian, i.e., with normalized kurtosis $\kappa_b \triangleq \gamma / \sigma_b^2 < 2$. Let $\boldsymbol{u}_0$ denote the initial value of the CM weight vector $\boldsymbol{u}$, if: C2) $\boldsymbol{u}_0^H \boldsymbol{\mathcal{F}}^H \boldsymbol{R}_{\boldsymbol{rr}} \boldsymbol{\mathcal{F}} \, \boldsymbol{u}_0 = (2\,\gamma)/(\kappa_b + 2)$ and C3) $\text{SINR}^{(\text{II})}(\boldsymbol{u}_0) > 1 + \sqrt{2}$, the GD minimization of the CM cost function, initialized with $\boldsymbol{u}_0$, will converge, in the absence of noise, to a solution extracting the desired symbol $b_1(k)$.

In practice, Theorem 1 represents a sufficient condition assuring that, in the high SNR region, the desired symbol is extracted, provided that conditions C2 and C3 are fulfilled. As pointed out in [21], the gain condition C2 is not critical if the value of $\text{SINR}^{(\text{II})}(\boldsymbol{u}_0)$ is far enough from its critical value $1 + \sqrt{2}$; in this case, extraction of the desired symbol is guaranteed also for a value of $\boldsymbol{u}_0^H \boldsymbol{\mathcal{F}}^H \boldsymbol{R}_{\boldsymbol{rr}} \boldsymbol{\mathcal{F}} \, \boldsymbol{u}_0$ lying in a bounded interval around $(2\,\gamma)/(\kappa_b + 2)$. Note that, for a given filtering matrix $\boldsymbol{\mathcal{F}}$, condition C2 can be blindly satisfied by suitably

scaling the initial weight vector $\boldsymbol{u}_0$; for this reason, in the sequel, we will essentially concentrate on condition C3.

The last step of our analysis is to relate condition C3 to the SINR at the output of the first stage, that is, to express $\text{SINR}^{(\text{II})}(\boldsymbol{u})$ as a function of $\text{SINR}^{(\text{I})}(\boldsymbol{\mathcal{F}})$. To this aim, we restrict our attention to the subset of matrices $\boldsymbol{\mathcal{F}}$ that satisfy the constraint $\boldsymbol{\mathcal{F}}^H \boldsymbol{\Upsilon}_1 = \boldsymbol{I}_{L_{\text{cp}}}$; in this case, one has

$$\text{SINR}^{(\text{I})}(\boldsymbol{\mathcal{F}}) = \frac{\sigma_b^2 \, \|\overline{\boldsymbol{g}}_1\|^2}{\text{trace}(\boldsymbol{\mathcal{F}}^H \boldsymbol{R}_{\boldsymbol{dd}} \boldsymbol{\mathcal{F}})} \quad \text{and} \quad \text{SINR}^{(\text{II})}(\boldsymbol{u}) = \frac{\sigma_b^2 \, |\boldsymbol{u}^H \overline{\boldsymbol{g}}_1|^2}{\boldsymbol{u}^H \boldsymbol{\mathcal{F}}^H \boldsymbol{R}_{\boldsymbol{dd}} \boldsymbol{\mathcal{F}} \, \boldsymbol{u}} \, . \tag{23}$$

The denominator of $\text{SINR}^{(\text{II})}(\boldsymbol{u})$ in (23) cannot be explicitly expressed in terms of $\text{trace}(\boldsymbol{\mathcal{F}}^H \boldsymbol{R}_{\boldsymbol{dd}} \boldsymbol{\mathcal{F}})$: in general, let λ_{max} denote the maximum eigenvalue of $\boldsymbol{\mathcal{F}}^H \boldsymbol{R}_{\boldsymbol{dd}} \boldsymbol{\mathcal{F}}$, one is able to derive [18] only the following bound $\boldsymbol{u}^H \boldsymbol{\mathcal{F}}^H \boldsymbol{R}_{\boldsymbol{dd}} \boldsymbol{\mathcal{F}} \, \boldsymbol{u} \leq \lambda_{\text{max}} \, \|\boldsymbol{u}\|^2$, which, utilized in (23), leads to

$$\text{SINR}^{(\text{II})}(\boldsymbol{u}) \geq \frac{\sigma_b^2 \, \|\overline{\boldsymbol{g}}_1\|^2 \, |\rho(\boldsymbol{u})|^2}{\lambda_{\text{max}}} \, , \tag{24}$$

where $\rho(\boldsymbol{u}) \stackrel{\triangle}{=} (\boldsymbol{u}^H \overline{\boldsymbol{g}}_1)/(\|\boldsymbol{u}\| \cdot \|\overline{\boldsymbol{g}}_1\|)$ represents the correlation coefficient between the weight vector $\boldsymbol{u}$ and the desired signature $\overline{\boldsymbol{g}}_1$. Accounting for (24) and observing that $\lambda_{\text{max}} \leq \text{trace}(\boldsymbol{\mathcal{F}}^H \boldsymbol{R}_{\boldsymbol{dd}} \boldsymbol{\mathcal{F}})$, the SINR at the output of the second stage can be related to the SINR at the output of the first stage as follows

$$\text{SINR}^{(\text{II})}(\boldsymbol{u}) \geq |\rho(\boldsymbol{u})|^2 \, \text{SINR}^{(\text{I})}(\boldsymbol{\mathcal{F}}) \, , \tag{25}$$

which shows that, for an arbitrary $\boldsymbol{u} \in \mathbb{C}^{L_{\text{cp}}}$, the *minimum value* of the SINR at the output of the second stage is proportional to the SINR at the output of the first stage. By using the lower-bound (25), condition C3 can be translated into an equivalent condition over the SINR at the output of the first stage; indeed, condition C3 is verified if

$$\text{SINR}^{(\text{I})}(\boldsymbol{\mathcal{F}}) > \frac{1+\sqrt{2}}{|\rho(\boldsymbol{u}_0)|^2} \, , \quad \text{with} \quad |\rho(\boldsymbol{u}_0)| \neq 0 \, . \tag{26}$$

It is worthwhile to note that, under condition C1, the proposed first stage behaves as a blind zero-forcing detector in the high SNR region, i.e., $\text{SINR}^{(\text{I})}(\boldsymbol{\mathcal{F}}_{\text{opt}}) \to \infty$ and, thus, the sufficient condition (26) is certainly fulfilled by using the optimal two-stage receiver.

3.3 Performance analysis for finite sample-size

The aim of this subsection is to investigate the SINR degradation when the first stage is synthesized by using the sample correlation matrix of $\boldsymbol{r}(k)$, estimated over K symbol intervals, i.e., when the adaptive part (17) of the filtering matrix is evaluated as

$$\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}^{(a)} = (\boldsymbol{B}_1^H \widehat{\boldsymbol{R}}_{\boldsymbol{rr}} \boldsymbol{B}_1)^{-1} \boldsymbol{B}_1^H \widehat{\boldsymbol{R}}_{\boldsymbol{rr}} \boldsymbol{\mathcal{F}}_{\text{opt}}^{(0)} \, . \tag{27}$$

with

$$\widehat{\boldsymbol{R}}_{\boldsymbol{rr}} \triangleq \frac{1}{K} \sum_{k=0}^{K-1} \boldsymbol{r}(k)\, \boldsymbol{r}^H(k)\,. \tag{28}$$

In this case, since the overall matrix $\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}} = \boldsymbol{\mathcal{F}}_{\text{opt}}^{(0)} - \boldsymbol{B}_1\, \widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}^{(a)} \in \mathbb{C}_r^{N \times L_{\text{cp}}}$ is random, the expectations in (18) must be evaluated also with respect to $\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}$. To this end, let us rewrite (18), with $\boldsymbol{\mathcal{F}} = \widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}$, as

$$\text{SINR}^{(\text{I})}(\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}) \triangleq \frac{\text{E}_{b_1(k),\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}} \left[\| \widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}^H \boldsymbol{\Upsilon}_1\, \overline{\boldsymbol{g}}_1\, b_1(k) \|^2 \right]}{\text{E}_{\boldsymbol{d}(k),\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}} \left[\| \widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}^H\, \boldsymbol{d}(k) \|^2 \right]} = \frac{\sigma_b^2\, \| \overline{\boldsymbol{g}}_1 \|^2}{\text{E}_{\boldsymbol{d}(k),\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}} \left[\| \widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}^H\, \boldsymbol{d}(k) \|^2 \right]} \tag{29}$$

where the last equality accounts for the constraint $\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}^H \boldsymbol{\Upsilon}_1 = \boldsymbol{I}_{L_{\text{cp}}}$. The starting point of the analysis is to find a simple expression for the adaptive matrix $\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}^{(a)}$, which is more suited to our purposes. By substituting (9) in (28), one has

$$\widehat{\boldsymbol{R}}_{\boldsymbol{rr}} = \widehat{\sigma}_b^2\, \boldsymbol{\Upsilon}_1\, \overline{\boldsymbol{g}}_1\, \overline{\boldsymbol{g}}_1^H \boldsymbol{\Upsilon}_1^H + \boldsymbol{\Upsilon}_1\, \widehat{\boldsymbol{R}}^H + \widehat{\boldsymbol{R}}\, \boldsymbol{\Upsilon}_1^H + \widehat{\boldsymbol{R}}_{\boldsymbol{dd}}\,, \tag{30}$$

where $\widehat{\sigma}_b^2 \triangleq K^{-1} \sum_{k=0}^{K-1} |b_1(k)|^2$, $\widehat{\boldsymbol{R}} \triangleq K^{-1} \sum_{k=0}^{K-1} b_1^*(k)\, \boldsymbol{d}(k)\, \overline{\boldsymbol{g}}_1^H$ and $\widehat{\boldsymbol{R}}_{\boldsymbol{dd}} \triangleq K^{-1} \sum_{k=0}^{K-1} \boldsymbol{d}(k)\, \boldsymbol{d}^H(k)$ represent sample estimates of the symbol variance σ_b^2, the cross-correlation matrix between the disturbance vector $\boldsymbol{d}(k)$ and the desired vector (at the output of the first stage) $\overline{\boldsymbol{g}}_1\, b_1(k)$ and the correlation matrix of $\boldsymbol{d}(k)$, respectively. Obviously, for a finite K, the sample cross-correlation matrix $\widehat{\boldsymbol{R}}$ is nonzero even if the disturbance $\boldsymbol{d}(k)$ is statistically independent of the desired symbol $b_1(k)$ (see assumptions A1 and A2). By substituting (30) in (27), we obtain, after tedious but straightforward matrix algebra,

$$\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}^{(a)} = (\boldsymbol{B}_1^H \widehat{\boldsymbol{R}}_{\boldsymbol{dd}}\, \boldsymbol{B}_1)^{-1} \boldsymbol{B}_1^H \widehat{\boldsymbol{R}}_{\boldsymbol{dd}}\, \boldsymbol{\mathcal{F}}_{\text{opt}}^{(0)} + (\boldsymbol{B}_1^H \widehat{\boldsymbol{R}}_{\boldsymbol{dd}}\, \boldsymbol{B}_1)^{-1} \boldsymbol{B}_1^H \widehat{\boldsymbol{R}}\,, \tag{31}$$

which evidences that the estimate of $\boldsymbol{\mathcal{F}}_{\text{opt}}^{(a)}$ is composed of two terms: the former represents an estimate of the optimal matrix $\boldsymbol{\mathcal{F}}_{\text{opt}}^{(a)}$ given by (17), while the latter is the perturbation resulting from the nonzero sample cross-correlation matrix $\widehat{\boldsymbol{R}}$. To simplify the analysis, following [22], we resort in (31) to the approximation

$$\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}^{(a)} \approx \boldsymbol{\mathcal{F}}_{\text{opt}}^{(a)} + (\boldsymbol{B}_1^H \boldsymbol{R}_{\boldsymbol{dd}}\, \boldsymbol{B}_1)^{-1} \boldsymbol{B}_1^H \widehat{\boldsymbol{R}}\,, \tag{32}$$

that is, we replace the sample correlation matrix $\widehat{\boldsymbol{R}}_{\boldsymbol{dd}}$ with the exact one $\boldsymbol{R}_{\boldsymbol{dd}}$. As noted in [22] and confirmed by simulation results not reported here, this approximation is rather poor for very low values of the sample-size, i.e., for $K \approx N - L_{\text{cp}}$, whereas, for moderate to large values of the sample-size, e.g., $K \geq 3\,(N - L_{\text{cp}})$, the effect on the SINR of replacing $\widehat{\boldsymbol{R}}_{\boldsymbol{dd}}$ with $\boldsymbol{R}_{\boldsymbol{dd}}$ is marginal, since the matrix $\widehat{\boldsymbol{R}}$ is the principal cause of the SINR degradation.

In Appendix A it is shown that, by invoking assumptions A1 and A2, it results that

$$\text{SINR}^{(\text{I})}(\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}) = \frac{\text{SINR}_{\text{max}}^{(\text{I})}}{1 + \dfrac{N - L_{\text{cp}}}{K}\, \text{SINR}_{\text{max}}^{(\text{I})}}\,, \tag{33}$$

where $\text{SINR}_{\text{max}}^{(\text{I})}$ is given by (20). Under the assumption that the noise $\boldsymbol{v}(k)$ is white with variance σ_v^2, it is interesting to note that, as $\sigma_v^2/\sigma_b^2 \rightarrow 0$ and under condition C1, it results that $\text{SINR}_{\text{max}}^{(\text{I})} \rightarrow \infty$ and, thus, expression (33) becomes

$$\lim_{\sigma_v^2/\sigma_b^2 \rightarrow 0} \text{SINR}^{(\text{I})}(\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}) = \frac{K}{N - L_{\text{cp}}} \,, \tag{34}$$

which shows that, due to the effect of the finite sample-size K, the SINR saturates to a fixed value even when $\sigma_v^2/\sigma_b^2 \rightarrow 0$. In this case, by using (26), we observe that the second stage can safely extract the desired symbol $b_1(k)$ if the sample-size K satisfies the inequality

$$K > (1 + \sqrt{2}) \, \frac{N - L_{\text{cp}}}{|\rho(\boldsymbol{u}_0)|^2} \,, \quad \text{with} \quad |\rho(\boldsymbol{u}_0)| \neq 0 \,. \tag{35}$$

Relation (35) allows one to derive two interesting conclusions. Firstly, the minimum sample-size K_{min} required to avoid the interference capture in the second stage increases linearly with the number of degrees of freedom $N - L_{\text{cp}}$ for disturbance suppression which, in its turn, increases linearly with the total number D of the MAI signatures, in order to fulfill condition C1: this ultimately implies that K_{min} increases linearly with D. Secondly, in the case of a finite sample-size, the initial weight vector $\boldsymbol{u}_0$ plays an important role in determining the overall performance of the two-stage receiver. In fact, if $\boldsymbol{u}_0$ is mistakenly chosen so as to be nearly orthogonal to the unknown signature $\overline{\boldsymbol{g}}_1$, i.e., $|\rho(\boldsymbol{u}_0)| \approx 0$, the extraction of the desired symbol requires an exceedingly large sample-size. Therefore, in setting the initial vector $\boldsymbol{u}_0$, one has to find in principle an approximation that is close to $\varrho \, \overline{\boldsymbol{g}}_1$, with $\varrho \in \mathbb{C}$, across all possible scenarios of interest. In practice, one can only resort to some reasonable *ad hoc* choices. In macrocellular system, typical multipath intensity profiles show [7] that most of the average power is concentrated within the first sampling interval: in this scenario, a reasonable approximation [23] of the channel vector $\boldsymbol{g}_1$ is given by $\widehat{\boldsymbol{g}}_1 = [1, a, \ldots, a]^T$, with $a \stackrel{\triangle}{=} 1/L_{1,\text{max}}$, where $L_{1,\text{max}}$ represents a known upper-bound of the desired channel order L_1. In our case, accounting for the structure of the *composite* channel vector $\overline{\boldsymbol{g}}_1$, we have chosen in Section 5 the following initialization for the second stage

$$\boldsymbol{u}_0 = [\underbrace{1, 1, \ldots, 1}_{d_{1,\text{max}}+1}, \; \underbrace{a, \ldots, a}_{L_{\text{cp}} - d_{1,\text{max}} - 1} \,]^T \,, \tag{36}$$

where $d_{1,\text{max}}$ denotes a known upper-bound of the desired transmission delay d_1. This choice was verified by computer simulations to lead to acceptable values of $|\rho(\boldsymbol{u}_0)|$.

4 Robust version of the blind two-stage receiver

The analysis carried out in Subsection 3.3 shows that the SINR degradation at the output of the first stage due to the finite sample-size is basically imputable to the effect of the sample cross-correlation matrix $\widehat{\boldsymbol{R}}$ between the

disturbance vector $\boldsymbol{d}(k)$ and the desired vector $\overline{\boldsymbol{g}}_1\, b_1(k)$; moreover, this degradation increases as the number of degrees of freedom $N - L_{\text{cp}}$ increases. A simple and effective way to reduce the SINR degradation is thus to suitably reduce the degrees of freedom, which is equivalent to adding constraints to the optimization problem (15). On the other hand, for a fixed disturbance suppression level, reducing the number of the degrees of freedom entails a reduction of the total number of MAI signatures that the two-stage receiver is able to handle. In this section, our goal is to add an appropriate constraint in the synthesis of the first stage in order to gain robustness against finite sample-size effects, without significantly compromising its MAI suppression capability.

4.1 The blind robust receiver

Let us start from considering the sample power $\widehat{P}_{\text{out}} \triangleq K^{-1} \sum_{k=0}^{K-1} \boldsymbol{x}^H(k)\, \boldsymbol{x}(k)$ at the output of the first stage which, accounting for (14) and (30), can be expressed as

$$\widehat{P}_{\text{out}} = \text{trace}(\boldsymbol{\mathcal{F}}^H \widehat{\boldsymbol{R}}_{\boldsymbol{rr}}\, \boldsymbol{\mathcal{F}}) = \widehat{\sigma}_b^2\, \|\boldsymbol{\mathcal{F}}^H \boldsymbol{\Upsilon}_1\, \overline{\boldsymbol{g}}_1\|^2 + 2\,\text{Re}\{\text{trace}(\boldsymbol{\mathcal{F}}^H \widehat{\boldsymbol{R}}\, \boldsymbol{\Upsilon}_1^H \boldsymbol{\mathcal{F}})\} + \text{trace}(\boldsymbol{\mathcal{F}}^H \widehat{\boldsymbol{R}}_{\boldsymbol{dd}}\, \boldsymbol{\mathcal{F}})\,. \tag{37}$$

Observe that, accounting for (10), matrix $\widehat{\boldsymbol{R}}$ can be explicitly written as

$$\widehat{\boldsymbol{R}} = \boldsymbol{\mathcal{H}}_{\text{in}}\, \widehat{\boldsymbol{R}}_{\text{in}} + \widehat{\boldsymbol{\Xi}}\,, \tag{38}$$

where $\widehat{\boldsymbol{R}}_{\text{in}} \triangleq K^{-1} \sum_{k=0}^{K-1} \boldsymbol{b}_{\text{in}}(k)\, b_1^*(k)\, \overline{\boldsymbol{g}}_1^H$ is the sample cross-correlation matrix between the symbol vector $\boldsymbol{b}_{\text{in}}(k)$ of the interfering in-cell users and the desired vector $b_1(k)\, \overline{\boldsymbol{g}}_1$, whereas $\widehat{\boldsymbol{\Xi}} \triangleq K^{-1} \sum_{k=0}^{K-1} \boldsymbol{\jmath}(k)\, b_1^*(k)\, \overline{\boldsymbol{g}}_1^H$, with $\boldsymbol{\jmath}(k) \triangleq \boldsymbol{\mathcal{H}}_{\text{out}}\, \boldsymbol{b}_{\text{out}}(k) + \boldsymbol{v}(k)$, represents the sample cross-correlation matrix between the residual disturbance $\boldsymbol{\jmath}(k)$ (out-of-cell MAI plus noise) and the desired vector. It is important to observe that, since the spreading codes of all the in-cell users are available at the base station, the matrix $\boldsymbol{\mathcal{H}}_{\text{in}}$ is partially known at the receiving side; in fact, taking into account parameterization (8) and equation (13), it results that[5]

$$\boldsymbol{\mathcal{H}}_{\text{in}} = \boldsymbol{\mathcal{Q}}_{\text{in}}\, \boldsymbol{\mathcal{G}}_{\text{in}} \tag{39}$$

where $\boldsymbol{\mathcal{Q}}_{\text{in}} \triangleq [\boldsymbol{\Upsilon}_2, \boldsymbol{\Upsilon}_3, \ldots, \boldsymbol{\Upsilon}_{J_{\text{in}}}] \in \mathbb{C}^{N \times (J_{\text{in}}-1) L_{\text{cp}}}$ and $\boldsymbol{\mathcal{G}}_{\text{in}} \triangleq \text{diag}[\overline{\boldsymbol{g}}_2, \overline{\boldsymbol{g}}_3, \ldots, \overline{\boldsymbol{g}}_{J_{\text{in}}}] \in \mathbb{C}^{(J_{\text{in}}-1) L_{\text{cp}} \times (J_{\text{in}}-1)}$. By substituting (38) and (39) in (37), and imposing the linear constraints $\boldsymbol{\mathcal{F}}^H \boldsymbol{\Upsilon}_1 = \boldsymbol{I}_{L_{\text{cp}}}$, one obtains

$$\widehat{P}_{\text{out}} = \widehat{\sigma}_b^2\, \|\overline{\boldsymbol{g}}_1\|^2 + 2\,\text{Re}\{\text{trace}(\boldsymbol{\mathcal{F}}^H \boldsymbol{\mathcal{Q}}_{\text{in}}\, \boldsymbol{\mathcal{G}}_{\text{in}}\, \widehat{\boldsymbol{R}}_{\text{in}})\} + 2\,\text{Re}\{\text{trace}(\boldsymbol{\mathcal{F}}^H\, \widehat{\boldsymbol{\Xi}})\} + \text{trace}(\boldsymbol{\mathcal{F}}^H \widehat{\boldsymbol{R}}_{\boldsymbol{dd}}\, \boldsymbol{\mathcal{F}})\,. \tag{40}$$

This relation suggests a simple strategy to exploit the knowledge of the matrix $\boldsymbol{\mathcal{Q}}_{\text{in}}$ for partially mitigating the sample cross-correlation between the disturbance and the desired vector. To this end, observe that, by invoking

[5]Note that, since the CP length L_{cp} is typically chosen of order of $0.25\, N$ to limit the amount of introduced redundancy, the matrix $\boldsymbol{\mathcal{Q}}_{\text{in}}$ turns out to be "wide", that is, $(J_{\text{in}} - 1)\, L_{\text{cp}} > N$, in many cases of practical interest, i.e., when $J_{\text{in}} > 5$.

the Cauchy-Schwarz inequality, one has (see [18])

$$\mathrm{Re}^2\{\mathrm{trace}(\boldsymbol{\mathcal{F}}^H \boldsymbol{\mathcal{Q}}_{\mathrm{in}}\, \boldsymbol{\mathcal{G}}_{\mathrm{in}}\, \widehat{\boldsymbol{R}}_{\mathrm{in}})\} = \mathrm{Re}^2\{\langle \boldsymbol{\mathcal{F}}^H \boldsymbol{\mathcal{Q}}_{\mathrm{in}}, (\boldsymbol{\mathcal{G}}_{\mathrm{in}}\, \widehat{\boldsymbol{R}}_{\mathrm{in}})^H \rangle\} \leq \|\boldsymbol{\mathcal{F}}^H \boldsymbol{\mathcal{Q}}_{\mathrm{in}}\|^2 \cdot \|\boldsymbol{\mathcal{G}}_{\mathrm{in}}\, \widehat{\boldsymbol{R}}_{\mathrm{in}}\|^2 \,, \tag{41}$$

from which it results that, by imposing that $\boldsymbol{\mathcal{F}}$ satisfies the *quadratic constraint* $\|\boldsymbol{\mathcal{F}}^H \boldsymbol{\mathcal{Q}}_{\mathrm{in}}\|^2 \leq \epsilon_0$, with ϵ_0 being a nonnegative number, the squared modulus of the contribution to the output power $\widehat{P}_{\mathrm{out}}$ due to the sample cross-correlation between the in-cell MAI and the desired vector is at most equal to $\epsilon_0\, \|\boldsymbol{\mathcal{G}}_{\mathrm{in}}\, \widehat{\boldsymbol{R}}_{\mathrm{in}}\|^2$. This means that the magnitude of the second term in (40) can be *deterministically* bounded by appropriately choosing the value of ϵ_0. Based on this consideration, we propose to modify (15) and to choose the filtering matrix $\boldsymbol{\mathcal{F}}$ so as to satisfy the following optimization problem with a linear equality constraint and a quadratic inequality constraint:

$$\widehat{\boldsymbol{\mathcal{F}}}_{\mathrm{rob}} = \arg\min_{\boldsymbol{\mathcal{F}}} \mathrm{trace}(\boldsymbol{\mathcal{F}}^H \widehat{\boldsymbol{R}}_{\boldsymbol{rr}}\, \boldsymbol{\mathcal{F}}) \quad \text{subject to} \quad \begin{cases} \boldsymbol{\mathcal{F}}^H \boldsymbol{\Upsilon}_1 = \boldsymbol{I}_{L_{\mathrm{cp}}} \,, \\ \|\boldsymbol{\mathcal{F}}^H \boldsymbol{\mathcal{Q}}_{\mathrm{in}}\|^2 \leq \epsilon_0 \,. \end{cases} \tag{42}$$

Similarly to (16), the linear equality constraint $\boldsymbol{\mathcal{F}}^H \boldsymbol{\Upsilon}_1 = \boldsymbol{I}_{L_{\mathrm{cp}}}$ gives to the solution of (42) the canonical structure

$$\widehat{\boldsymbol{\mathcal{F}}}_{\mathrm{rob}} = \boldsymbol{\mathcal{F}}_{\mathrm{opt}}^{(0)} - \boldsymbol{B}_1\, \widehat{\boldsymbol{\mathcal{F}}}_{\mathrm{rob}}^{(a)} \,, \tag{43}$$

where the matrix $\widehat{\boldsymbol{\mathcal{F}}}_{\mathrm{rob}}^{(a)} \in \mathbb{C}_r^{(N-L_{\mathrm{cp}})\times L_{\mathrm{cp}}}$ turns out to be the solution of the following quadratically constrained optimization problem

$$\widehat{\boldsymbol{\mathcal{F}}}_{\mathrm{rob}}^{(a)} = \arg\min_{\boldsymbol{\mathcal{F}}^{(a)}} \mathrm{trace}\left\{ [\boldsymbol{\mathcal{F}}_{\mathrm{opt}}^{(0)} - \boldsymbol{B}_1\, \boldsymbol{\mathcal{F}}^{(a)}]^H \widehat{\boldsymbol{R}}_{\boldsymbol{rr}}\, [\boldsymbol{\mathcal{F}}_{\mathrm{opt}}^{(0)} - \boldsymbol{B}_1\, \boldsymbol{\mathcal{F}}^{(a)}] \right\} \tag{44}$$

subject to $\|[\boldsymbol{\mathcal{F}}_{\mathrm{opt}}^{(0)} - \boldsymbol{B}_1\, \boldsymbol{\mathcal{F}}^{(a)}]^H \boldsymbol{\mathcal{Q}}_{\mathrm{in}}\|^2 \leq \epsilon_0$, whose solution is given by (see Appendix B)

$$\widehat{\boldsymbol{\mathcal{F}}}_{\mathrm{rob}}^{(a)} = [\boldsymbol{B}_1^H (\widehat{\boldsymbol{R}}_{\boldsymbol{rr}} + \mu_0\, \boldsymbol{\mathcal{Q}}_{\mathrm{in}}\, \boldsymbol{\mathcal{Q}}_{\mathrm{in}}^H)\, \boldsymbol{B}_1]^{-1} \boldsymbol{B}_1^H (\widehat{\boldsymbol{R}}_{\boldsymbol{rr}} + \mu_0\, \boldsymbol{\mathcal{Q}}_{\mathrm{in}}\, \boldsymbol{\mathcal{Q}}_{\mathrm{in}}^H)\, \boldsymbol{\mathcal{F}}_{\mathrm{opt}}^{(0)} \,, \tag{45}$$

where $\mu_0 \geq 0$ is the Lagrange multiplier, which is chosen so as to satisfy the equation

$$\|[\boldsymbol{\mathcal{F}}_{\mathrm{opt}}^{(0)} - \boldsymbol{B}_1\, \widehat{\boldsymbol{\mathcal{F}}}_{\mathrm{rob}}^{(a)}]^H \boldsymbol{\mathcal{Q}}_{\mathrm{in}}\|^2 = \epsilon_0 \,. \tag{46}$$

It should be observed that, unlike linearly and quadratically constrained minimum power beamforming techniques [24], which are well-known reception strategies in the context of array processing, the amount of loading induced by the quadratic constraint in (40) is not diagonal, that is, $\boldsymbol{\mathcal{Q}}_{\mathrm{in}}\, \boldsymbol{\mathcal{Q}}_{\mathrm{in}}^H \neq \boldsymbol{I}_N$, and depends on the spreading codes of the in-cell active users. When $\mu_0 = 0$, matrix (45) degenerates into the adaptive matrix $\boldsymbol{\mathcal{F}}_{\mathrm{opt}}^{(a)}$ given by (27): this corresponds to the case where $\epsilon_0 \to \infty$, that is, when the quadratic constraint is inactive. On the other hand, the value of the Lagrange multiplier μ_0 cannot be chosen arbitrarily large or, equivalently, the constraint value ϵ_0 cannot be chosen arbitrarily small. In fact, in order to assure that the constrained optimization

problem (44) admits a solution, the constraint value ϵ_0 must satisfy the condition

$$\min_{\boldsymbol{\mathcal{F}}^{(a)}} \|[\boldsymbol{\mathcal{F}}_{\text{opt}}^{(0)} - \boldsymbol{B}_1 \; \boldsymbol{\mathcal{F}}^{(a)}]^H \boldsymbol{\mathcal{Q}}_{\text{in}}\|^2 \leq \epsilon_0 \; . \tag{47}$$

Appendix B shows that, when the matrix $\boldsymbol{\mathcal{Q}}_{\text{in}}$ is full-row rank[6], a reasonable choice for the constraint value is $\epsilon_0 \geq \text{trace}[(\boldsymbol{\mathcal{F}}_{\text{opt}}^{(0)})^H \boldsymbol{\mathcal{Q}}_{\text{in}} \, \boldsymbol{\mathcal{Q}}_{\text{in}}^H \, \boldsymbol{\mathcal{F}}_{\text{opt}}^{(0)}]$. Unfortunately, the optimal value of μ_0 is related to ϵ_0 by means of the trascendental equation (46) and, thus, it can be evaluated only numerically [24, 25]. This can be accomplished by observing that equation (46) can be equivalently written as

$$g(\mu_0) \stackrel{\triangle}{=} \text{trace}[(\widehat{\boldsymbol{\mathcal{F}}}_{\text{rob}}^{(a)})^H \boldsymbol{B}_1^H \boldsymbol{\mathcal{Q}}_{\text{in}} \, \boldsymbol{\mathcal{Q}}_{\text{in}}^H \boldsymbol{B}_1 \, \widehat{\boldsymbol{\mathcal{F}}}_{\text{rob}}^{(a)}] - 2 \, \text{Re}\{\text{trace}[(\widehat{\boldsymbol{\mathcal{F}}}_{\text{rob}}^{(a)})^H \boldsymbol{B}_1^H \boldsymbol{\mathcal{Q}}_{\text{in}} \, \boldsymbol{\mathcal{Q}}_{\text{in}}^H \boldsymbol{\mathcal{F}}_{\text{opt}}^{(0)}]\} = \beta_0 \; , \tag{48}$$

where $\beta_0 \stackrel{\triangle}{=} \epsilon_0 - \text{trace}[(\boldsymbol{\mathcal{F}}_{\text{opt}}^{(0)})^H \boldsymbol{\mathcal{Q}}_{\text{in}} \, \boldsymbol{\mathcal{Q}}_{\text{in}}^H \, \boldsymbol{\mathcal{F}}_{\text{opt}}^{(0)}]$. Assuming that (48) is not satisfied when $\mu_0 = 0$, that is, $g(0) > \beta_0$, the following iterative procedure can be used to determine the optimal value of the Lagrange multiplier μ_0: starting with $\mu_0^{(0)} = 0$, let $\mu_0^{(1)} = \mu_0^{(0)} + \Delta\,\mu_0, \ldots, \mu_0^{(\ell)} = \mu_0^{(\ell-1)} + \Delta\,\mu_0$, where $\Delta\,\mu_0$ is a small positive number. At the ℓth step, compute $g(\mu_0^{(\ell)})$ and compare it with the threshold β_0: if $f(\mu_0^{(\ell)}) \leq \beta_0$, then choose $\mu_0^{(\ell)}$ as the optimal value of the Lagrange multiplier μ_0, else perform the $(\ell+1)$th iteration and repeat the procedure.

A final remark is now in order about the computational load of the robust version of the first stage. For a given value of the Lagrange multiplier μ_0, the synthesis of the robust filtering matrix $\widehat{\boldsymbol{\mathcal{F}}}_{\text{rob}}^{(a)}$ in (45) involves essentially the same computational complexity required to estimate in batch-mode the optimal matrix $\boldsymbol{\mathcal{F}}_{\text{opt}}^{(a)}$ in (17); furthermore, reasoning as in [14], one can estimate $\widehat{\boldsymbol{\mathcal{F}}}_{\text{rob}}^{(a)}$ adaptively by means of RLS-based algorithms, with computational requirements per symbol interval of order of $\mathcal{O}[(N - L_{\text{cp}})^2]$. When the above-mentioned iterative procedure is used to determine the optimal value of μ_0, such a quadratic complexity must be multiplied by the number of iterations involved.

4.2 Performance analysis for finite sample-size

In this subsection we provide a first-order analysis of the SINR at the output of the first stage synthesized by using the robust filtering matrix (45): this analysis is aimed at showing the SINR enhancement provided by using the quadratic constraint in (42) as well as the impact of this constraint on the number of degrees of freedom for disturbance suppression.

Accounting for (18) and reasoning as in Subsection 3.3 and in Appendix A [see, in particular, equations (29) and (58)], the SINR at the output of the first stage can be written as

$$\text{SINR}^{(\text{I})}(\widehat{\boldsymbol{\mathcal{F}}}_{\text{rob}}) = \frac{\sigma_b^2 \, \|\overline{\boldsymbol{g}}_1\|^2}{\text{E}_{\widehat{\boldsymbol{\mathcal{F}}}_{\text{rob}}} \left\{ \text{trace}[\widehat{\boldsymbol{\mathcal{F}}}_{\text{rob}}^H \, \boldsymbol{R}_{\boldsymbol{dd}} \, \widehat{\boldsymbol{\mathcal{F}}}_{\text{rob}}] \right\}} \; . \tag{49}$$

[6]Note that this assumption is very mild and it is fulfilled by the spreading codes commonly used in practice, e.g., Walsh-Hadamard spreading sequences (used in our computer simulations).

Appendix C shows that, for $\mu_0 \, \| \, (\boldsymbol{B}_1^H \boldsymbol{R}_{\boldsymbol{dd}} \, \boldsymbol{B}_1)^{-1} \boldsymbol{B}_1^H \boldsymbol{\mathcal{Q}}_{\text{in}} \, \boldsymbol{\mathcal{Q}}_{\text{in}}^H \, \boldsymbol{B}_1 \| \ll 1$, the robust filtering matrix $\widehat{\boldsymbol{\mathcal{F}}}_{\text{rob}}$ is approximately related to $\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}$ by means of the simple expression

$$\widehat{\boldsymbol{\mathcal{F}}}_{\text{rob}} \approx (\boldsymbol{I}_N - \mu_0 \, \boldsymbol{\Psi} \, \boldsymbol{\mathcal{Q}}_{\text{in}} \, \boldsymbol{\mathcal{Q}}_{\text{in}}^H) \, \widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}} \, , \tag{50}$$

where $\boldsymbol{\Psi} \triangleq \boldsymbol{B}_1 \, (\boldsymbol{B}_1^H \boldsymbol{R}_{\boldsymbol{dd}} \, \boldsymbol{B}_1)^{-1} \boldsymbol{B}_1^H$. By using this approximation, under assumptions A1 and A2, the SINR at the output of the first stage of the robust receiver is given by (see Appendix C)

$$\text{SINR}^{(\text{I})}(\widehat{\boldsymbol{\mathcal{F}}}_{\text{rob}}) = \frac{\text{SINR}_{\text{max}}^{(\text{I})}}{1 + \dfrac{N - L_{\text{cp}} - \Omega(\mu_0)}{K} \, \text{SINR}_{\text{max}}^{(\text{I})}} \, , \tag{51}$$

where $\text{SINR}_{\text{max}}^{(\text{I})}$ is given by (20) and $\Omega(\mu_0) \triangleq -(a + K \, b) \, \mu_0^2 + 2 \, c \, \mu_0$ is a quadratic function of the Lagrange multiplier μ_0, with

$$a \triangleq \text{trace}[(\boldsymbol{\mathcal{Q}}_{\text{in}}^H \, \boldsymbol{\Psi} \, \boldsymbol{\mathcal{Q}}_{\text{in}})^2] \geq 0 \, , \tag{52}$$

$$b \triangleq \sigma_b^{-2} \, \|\overline{\boldsymbol{g}}_1\|^{-2} \, \text{trace}(\boldsymbol{\mathcal{F}}_{\text{opt}}^H \boldsymbol{\mathcal{Q}}_{\text{in}} \, \boldsymbol{\mathcal{Q}}_{\text{in}}^H \boldsymbol{\Psi} \, \boldsymbol{\mathcal{Q}}_{\text{in}} \, \boldsymbol{\mathcal{Q}}_{\text{in}}^H \, \boldsymbol{\mathcal{F}}_{\text{opt}}) \geq 0 \, , \tag{53}$$

$$c \triangleq \text{trace}(\boldsymbol{\mathcal{Q}}_{\text{in}}^H \, \boldsymbol{\Psi} \, \boldsymbol{\mathcal{Q}}_{\text{in}}) \geq 0 \, . \tag{54}$$

Note that (51) is similar to (33), except for the presence of $\Omega(\mu_0)$. It can be shown that, under the assumption $\mu_0 \, \| \, (\boldsymbol{B}_1^H \boldsymbol{R}_{\boldsymbol{dd}} \, \boldsymbol{B}_1)^{-1} \boldsymbol{B}_1^H \boldsymbol{\mathcal{Q}}_{\text{in}} \, \boldsymbol{\mathcal{Q}}_{\text{in}}^H \, \boldsymbol{B}_1 \| \ll 1$, the function $\Omega(\mu_0)$ is positive for all practical values of the sample size K, that is, $\mu_0 < 2 \, c/(a + K \, b)$ and, moreover, its maximum value is much smaller than $N - L_{\text{cp}}$.

Some remarks are now in order. As previously claimed, imposing the quadratic constraint in (42) improves robustness against estimation errors in the estimate of the correlation matrix $\boldsymbol{R}_{\boldsymbol{rr}}$: in fact, from (51), it turns out that, for (reasonable) finite values of the sample-size K, $\text{SINR}^{(\text{I})}(\widehat{\boldsymbol{\mathcal{F}}}_{\text{rob}}) > \text{SINR}^{(\text{I})}(\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}})$. By comparing (33) to (51), this favourable behavior is basically due to the presence of the term $\Omega(\mu_0)$ in (51), which reduces the degrees of freedom of the first stage. Clearly, the beneficial effect of the quadratic constraint disappears as K approaches infinity: in this case, it results that $\lim_{K \to \infty} \text{SINR}^{(\text{I})}(\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}) = \text{SINR}_{\text{max}}^{(\text{I})}$, whereas

$$\lim_{K \to \infty} \text{SINR}^{(\text{I})}(\widehat{\boldsymbol{\mathcal{F}}}_{\text{rob}}) = \frac{\text{SINR}_{\text{max}}^{(\text{I})}}{1 + \mu_0^2 \, \text{SINR}_{\text{max}}^{(\text{I})} \, \text{trace}(\boldsymbol{\mathcal{F}}_{\text{opt}}^H \boldsymbol{\mathcal{Q}}_{\text{in}} \, \boldsymbol{\mathcal{Q}}_{\text{in}}^H \boldsymbol{\Psi} \, \boldsymbol{\mathcal{Q}}_{\text{in}} \, \boldsymbol{\mathcal{Q}}_{\text{in}}^H \, \boldsymbol{\mathcal{F}}_{\text{opt}})} \, , \tag{55}$$

which is slightly smaller than $\text{SINR}_{\text{max}}^{(\text{I})}$. On the other hand, by comparing (33) with (51), it is apparent that, for small values of $\text{SINR}_{\text{max}}^{(\text{I})}$, i.e., when $\text{SINR}_{\text{max}}^{(\text{I})} \ll K/(N - L_{\text{cp}})$, one has $\text{SINR}^{(\text{I})}(\widehat{\boldsymbol{\mathcal{F}}}_{\text{rob}}) \approx \text{SINR}^{(\text{I})}(\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}) \approx \text{SINR}_{\text{max}}^{(\text{I})}$ and, thus, adopting the quadratic constraint in (42) is practically useless: this typically happens in the low SNR region and/or when condition C1 is near to be violated.

Our analysis is conservative: indeed, it applies only to very small values of $\Omega(\mu_0)$ (compared to $N - L_{\text{cp}}$). However, it should be noted that even a small decrease of $N - L_{\text{cp}}$ in the denominator of (33) can lead to a non-negligible increase of $\text{SINR}^{(\text{I})}(\widehat{\boldsymbol{\mathcal{F}}}_{\text{rob}})$ with respect to $\text{SINR}^{(\text{I})}(\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}})$. In fact, let us consider a small perturbation

$0 < \Omega(\mu_0) \ll N - L_{\text{cp}}$ of the degrees of freedom $N - L_{\text{cp}}$, accounting for (33) and (51), it turns out that

$$\frac{\text{SINR}^{(\text{I})}(\widehat{\boldsymbol{\mathcal{F}}}_{\text{rob}}) - \text{SINR}^{(\text{I})}(\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}})}{\text{SINR}^{(\text{I})}(\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}})} \approx \frac{\text{SINR}^{(\text{I})}(\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}})}{K}\,\Omega(\mu_0)\;, \tag{56}$$

which shows that the relative SINR variation is greater than $\Omega(\mu_0)$ by a factor $\text{SINR}^{(\text{I})}(\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}})/K$, which can be valuable for low values of the sample size K and/or for high values of $\text{SINR}^{(\text{I})}(\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}})$. For example, referring to the scenario considered in Example 1 of Section 5, it turns out that, for SNR $= 25$ dB and $K = 250$ symbols, $\text{SINR}^{(\text{I})}(\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}) = 10.0952$ (expressed in natural unit) and $\Omega(\mu_0) = 3.6736$, which, accounting for (56), lead to a relative SINR variation of about 15%. According to Theorem 1 and accounting for the discussion reported in Subsection 3.3, this SINR enhancement is expected to improve the performance of the CM algorithm in the second stage, by lowering, with respect to (35), the minimum sample-size K_{min} required to avoid the interference capture.

5 Simulation results

To confirm the results of the analysis previously carried out and to give more insight into the achievable performance of the two-stage receiver proposed in [13] (referred to as TS in the plots) as well as that of its robust implementation (referred to as robust TS in the plots), we present in this section the results of Monte Carlo computer simulations and compare them with the analytical results.

In all the experiments, the following common simulation setting is adopted. The quasi-synchronous MC-CDMA network employs $N = 32$ subcarriers, with a cyclic prefix of length $L_{\text{cp}} = 8$, and QPSK symbol modulation, which implies that the dispersion coefficient to be used in the CM cost function (11) is $\gamma = 1$; the frequency-domain spreading codes are length-32 Walsh-Hadamard sequences. The multipath channel of the jth user is $g_{c,j}(t) = \sum_{m=1}^{4} \beta_{m,j}\, \varphi_c(t - \tau_{m,j})$, where $\varphi_c(t)$ is a Nyquist-shaped pulse with 35% roll-off, the first path ($m = 1$) is assumed to be deterministic with amplitude $\beta_{1,j} = 1$ and propagation time $\tau_{1,j} = 0$, the remaining path gains $\beta_{m,j}$, for $m = 2, 3, 4$, are modeled as mutually independent complex circular Gaussian zero-mean random variables, with standard deviation 0.3, whereas the corresponding propagation times $\tau_{m,j}$ are modeled as mutually independent random variables, uniformly distributed over $L_j + 1 = 5$ sampling periods, for $j = 1, 2, \ldots, J$. The (integer) transmission delays d_j are modeled as discrete random variables, assuming equiprobable values in $\{0, 1, 2\}$, for $j = 1, 2, \ldots, J_{\text{in}}$, and in $\{0, 1, \ldots, P-1\}$, for $j = J_{\text{in}} + 1, J_{\text{in}} + 2, \ldots, J$. The additive noise samples $\widetilde{v}^{(\ell)}(k)$ in (3) are modeled as mutually independent complex circular zero-mean white Gaussian processes, with variance σ_v^2, and the signal-to-noise ratio (SNR) of the desired user at the

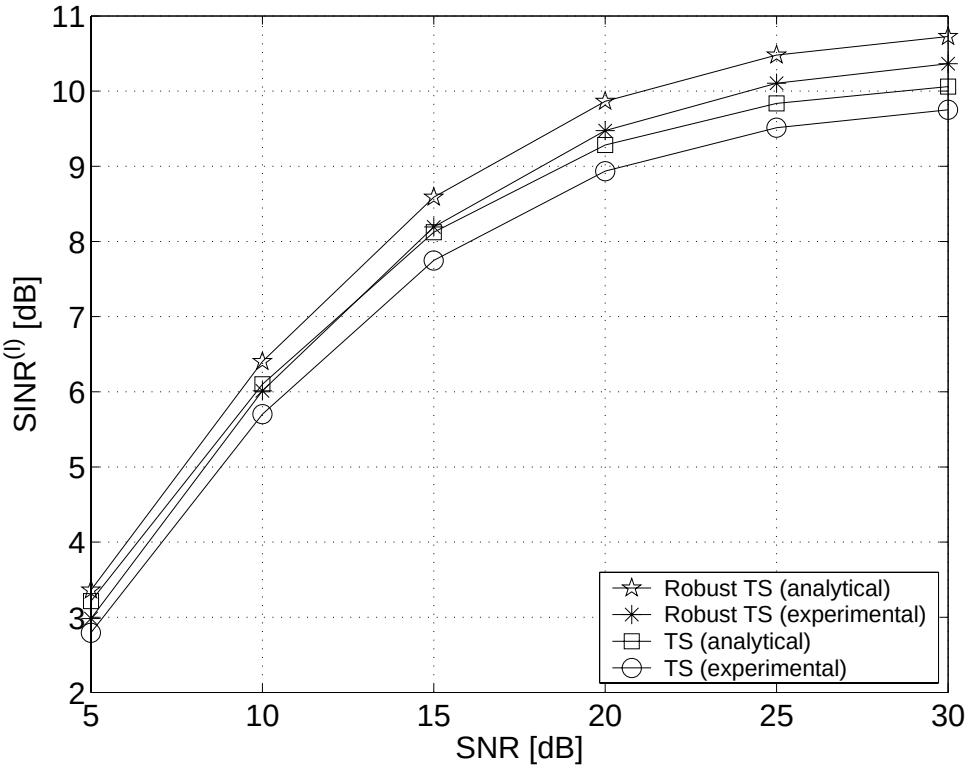

Figure 3: SINR at the output of the first stage versus SNR (first example, $K = 250$).

detector input is defined, according to (7), as

$$\text{SNR} \stackrel{\triangle}{=} \sigma_b^2 \, \frac{\| \boldsymbol{G}_1(0) \, \boldsymbol{W}_{\text{IDFT}} \, \boldsymbol{c}_1 \|^2}{\text{E}[\, \| \boldsymbol{v}(k) \|^2 \,]} \, . \tag{57}$$

We considered a severe near-far scenario: in all the experiments, the path gains of each user channel are adjusted so that each interfering in-cell user is 10 dB stronger than the user of interest ($j = 1$), whereas each out-of-cell user is received with the same power of the desired user (worst case). Unless otherwise specified, the number of the out-of-cell users is fixed to $J_{\text{out}} = 4$. All the results are obtained by carrying out 100 independent trials, with each run using a different set of noise samples and, for each user, a different set of transmission delays, channel parameters (path gains and propagation delays) and data sequences.

Example 1 – SINR performance of the first stage: In this example, we resort to Monte Carlo simulations to evaluate the SINR performance of the first stage of the robust TS, and compare it with that of the first stage of the TS receiver; moreover, the obtained results are compared with the analytical formulas (33) and (51). The number of active users is fixed to $J = 16$ and, after estimating the adaptive matrices $\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}^{(a)}$ and $\widehat{\boldsymbol{\mathcal{F}}}_{\text{rob}}^{(a)}$ on the basis of the given data record of length K, the output SINR is evaluated using (18). As to the robust receiver, in order to validate the first-order analysis of Subsection 4.2, the Lagrange multiplier μ_0 was chosen so as to satisfy the relation $\mu_0 \, \| \, (\boldsymbol{B}_1^H \boldsymbol{R}_{\boldsymbol{dd}} \, \boldsymbol{B}_1)^{-1} \boldsymbol{B}_1^H \, \boldsymbol{\mathcal{Q}}_{\text{in}} \, \boldsymbol{\mathcal{Q}}_{\text{in}}^H \, \boldsymbol{B}_1 \| = 0.9$.

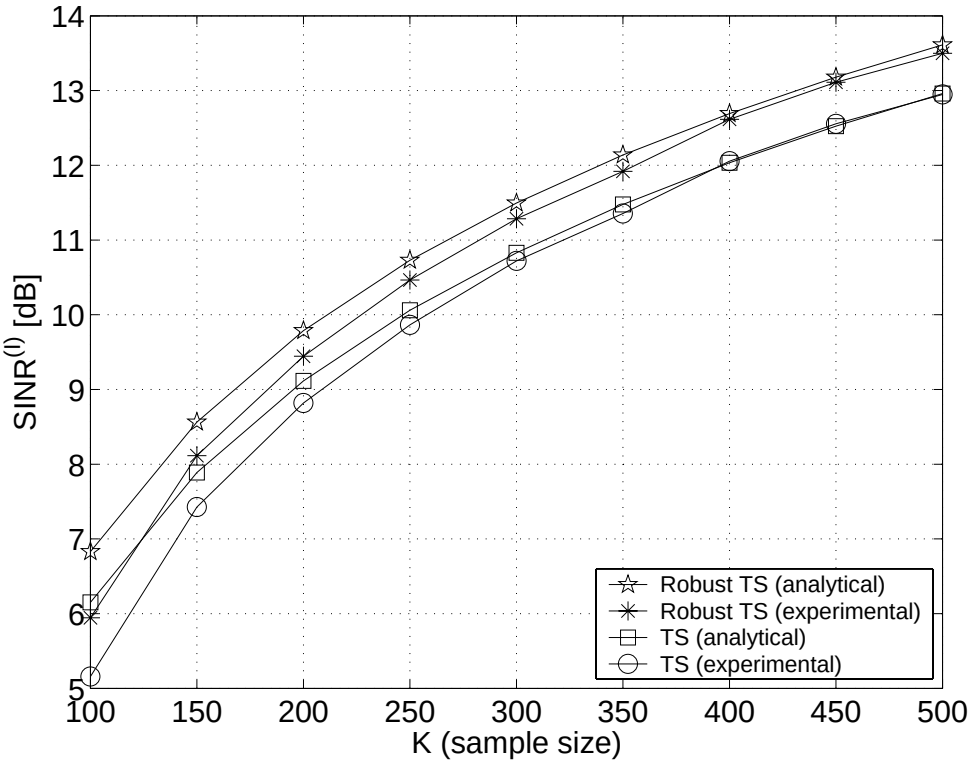

Figure 4: SINR at the output of the first stage versus sample size K (first example, SNR $= 30$ dB).

Figure 3 reports the values of SINR as a function of SNR ranging from 0 to 30 dB, with a sample size $K = 250$ symbols. In this case the order of magnitude of μ_0 varies from 10^{-4} (low values of SNR) to 10^{-6} (high values of SNR). It can be seen that, even though vanishingly small values of μ_0 are employed, the robust TS assures a valuable enhancement of the SINR at the output of the first stage with respect to its TS counterpart; in particular, both the first stages exhibit practically the same performance for low values of SNR, whereas the SINR increase provided by the incorporation of the quadratic constraint becomes more evident for moderate to high values of SNR. Observe that, taking into account the small value used for the sample size K, the absolute and relative behaviours of the two first stages are well predicted by the analytical results. To further corroborate the analysis, we evaluated the performance of the two considered first stages as a function of the sample size K (in symbols) ranging from 100 to 500, in the high SNR region, i.e., for SNR$= 30$ dB; in this region, the order of magnitude of the Lagrange multiplier is 10^{-6}. Results of Fig. 4 evidence a good agreement between experimental and analytical results and, in particular, show that the first stage of the robust TS appreciably outperforms its TS counterpart for all the considered values of the sample size. In this experiment, according to (35), we also evaluated the (average) minimum sample-size $K_{\min}$ required to avoid the interference capture in the second stage, when the CM is initialized with the vector $\boldsymbol{u}_0$ given by (36), with $d_{1,\max} = 2$ and $L_{1,\max} + 1 = 6$. Results show that, for the TS receiver, the interference capture in the second

stage is surely avoided if a minimum sample size of 219 symbols is used, whereas for the robust TS receiver $K_{\min}$ turns out to be equal to 187 symbols.

Example 2 – SER performance of the overall receiver: In this example, we present the Monte Carlo performance analysis of the overall two-stage receivers, together with a comparison with both non-blind (i.e., it is assumed the exact knowledge of the channel impulse response and transmission delay of the desired user) and blind versions of the subspace-based MMSE detector recently proposed in [12] (referred to as MMSE and blind MMSE in the plots, respectively). As (overall) performance measure, we resorted to the symbol error rate (SER) at the output of the considered receivers. After estimating the receiver weights (i.e., the correlation matrix $\boldsymbol{R_{rr}}$) in batch-mode on the basis of the given data record of length K, an independent record of $K_{\text{ser}} = 10^5$ symbols is considered to evaluate the SER at the output of the considered receivers. For the blind receivers, the equalized symbols are first rotated and scaled before evaluating the SER. The Lagrange multiplier μ_0 was chosen according to the algorithm described in Subsection 4.1, with $\Delta\,\mu_0 = 10^{-6}$, whereas the estimate of the optimal weight vector $\boldsymbol{u}_{\text{opt}}$ in (11) is obtained by resorting to the GD method, initialized by using a properly scaled (in accordance with condition C2) version of the vector $\boldsymbol{u}_0$ given by (36), with $d_{1,\text{max}} = 2$ and $L_{1,\text{max}} + 1 = 6$, where the complex gradient vector [26] (with respect to $\boldsymbol{u}^*$) of the CM cost function $\text{E}[(\gamma - |\boldsymbol{u}^H\,\boldsymbol{x}(k)|^2)^2]$ is estimated from the received data in batch-mode (see [27] for details).

In the first part of this example, the SER of the robust TS detector is firstly evaluated as a function of the quadratic constraint value $\epsilon_0 = \delta\,\text{trace}[(\boldsymbol{\mathcal{F}}_{\text{opt}}^{(0)})^H\,\boldsymbol{\mathcal{Q}}_{\text{in}}\,\boldsymbol{\mathcal{Q}}_{\text{in}}^H\,\boldsymbol{\mathcal{F}}_{\text{opt}}^{(0)}]$, with δ ranging from 2 to 22. Fig. 5 reports the SER of the robust TS receiver for different values of SNR, where the number of active users is $J = 16$ and the sample size is fixed to $K = 250$ symbols. It is apparent that, for low values of SNR, the best performance is achieved for $\delta_{\text{opt}} = 4$, whereas, for moderate values of SNR, the optimal choice of δ turns out to be $\delta_{\text{opt}} = 6$; moreover, observe that, except for $\delta = 2$, the SER gracefully degrades as δ deviates from its optimal value, for all the considered values of SNR. Similar considerations apply to Fig. 6, where the SER of the robust TS detector is depicted for different values of the sample size K (in symbols), for a number of users $J = 16$ and SNR $= 20$ dB. It is shown here that, in the considered scenario, the optimal value of δ is practically independent of the sample size. Finally, in Fig. 7 we reported the SER of the robust TS receiver for different values of the number J_{out} of the out-of-cell users; in this experiment, the number J_{in} of in-cell users is fixed to $J_{\text{in}} = 12$, the sample size and the SNR are set to $K = 250$ symbols and SNR $= 20$ dB, respectively. Results show that, for a fixed number of in-cell users, the SER is not considerably affected by increasing or decreasing the number of out-cell-users, provided that the total number of MAI signatures is obviously less than or equal to the number of degrees of freedom for disturbance suppression.

The second part of this example is devoted to the comparison between the two-stage receivers and both

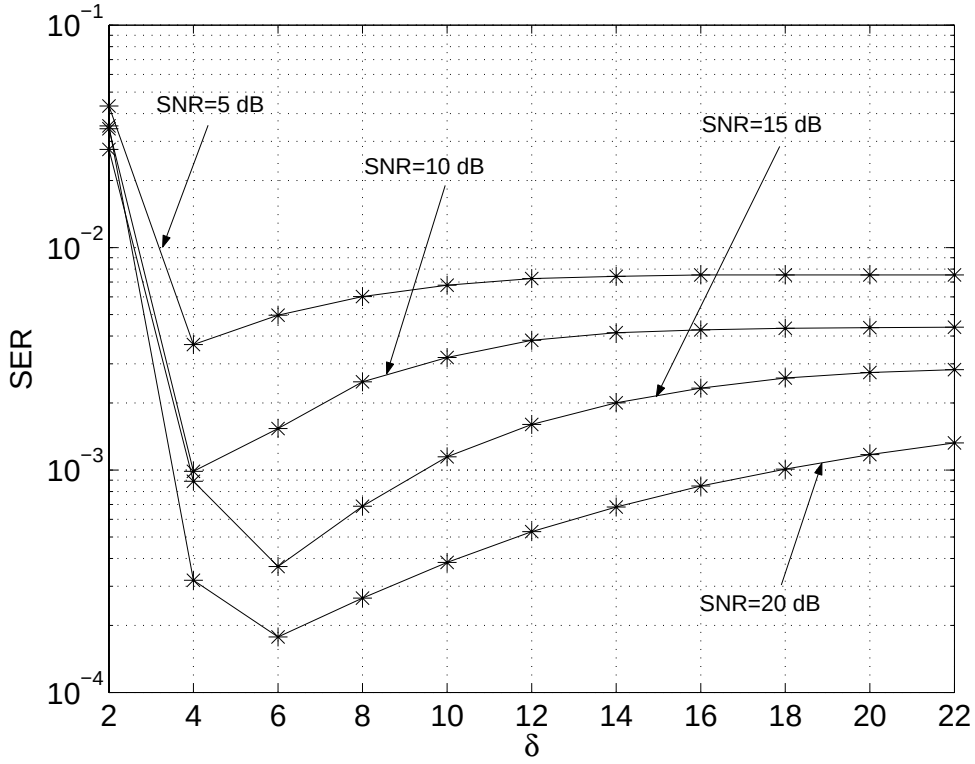

Figure 5: SER versus δ for different values of SNR (second example, $J = 16$, $K = 250$).

non-blind and blind versions of the subspace-based MMSE detector proposed in [12]. In the first experiment, we evaluated the SER of the considered receivers as a function of SNR ranging from 5 to 30 dB. The number of active users is $J = 16$ and the sample size is fixed to $K = 250$ symbols. The quadratic constraint value is chosen equal to $\epsilon_0 = 12\,\mathrm{trace}[(\boldsymbol{\mathcal{F}}_{\mathrm{opt}}^{(0)})^H \boldsymbol{\mathcal{Q}}_{\mathrm{in}}\, \boldsymbol{\mathcal{Q}}_{\mathrm{in}}^H\, \boldsymbol{\mathcal{F}}_{\mathrm{opt}}^{(0)}]$. From Fig. 8, it can be observed that, for high values of SNR (i.e., SNR ≥ 25 dB), the robust TS receiver exhibits performances that are better than or equal to those of the MMSE receivers, assuring a SER significantly inferior to 10^{-3} for SNR $= 20$ dB, whereas the performance of the TS receiver is quite unsatisfactory, showing a SER floor of about $3 \cdot 10^{-3}$ for high values of SNR. It should be observed that, although the blind MMSE receiver outperforms the robust TS for values of SNR ≤ 20 dB, its implementation is much more computational expensive (two eigendecompositions are involved) and, in the considered scenario, requires also the additional knowledge of the number J_{out} of the out-of-cell users. The second experiment investigates the convergence behaviour of the detectors under comparison. We have considered the same simulation setting described in the previous experiment (with $J = 16$ active users and $\epsilon_0 = 12\,\mathrm{trace}[(\boldsymbol{\mathcal{F}}_{\mathrm{opt}}^{(0)})^H \boldsymbol{\mathcal{Q}}_{\mathrm{in}}\, \boldsymbol{\mathcal{Q}}_{\mathrm{in}}^H\, \boldsymbol{\mathcal{F}}_{\mathrm{opt}}^{(0)}]$) and the SNR is fixed to 20 dB. Fig. 9 reports the SER as a function of the sample size K (in symbols) ranging from 100 to 400. It can be observed that the TS robust detector is competitive with the MMSE receivers, especially for small values of the sample size, while significantly outperforming the TS receiver. Moreover, the results of Fig. 9 show that, to obtain the same value of SER,

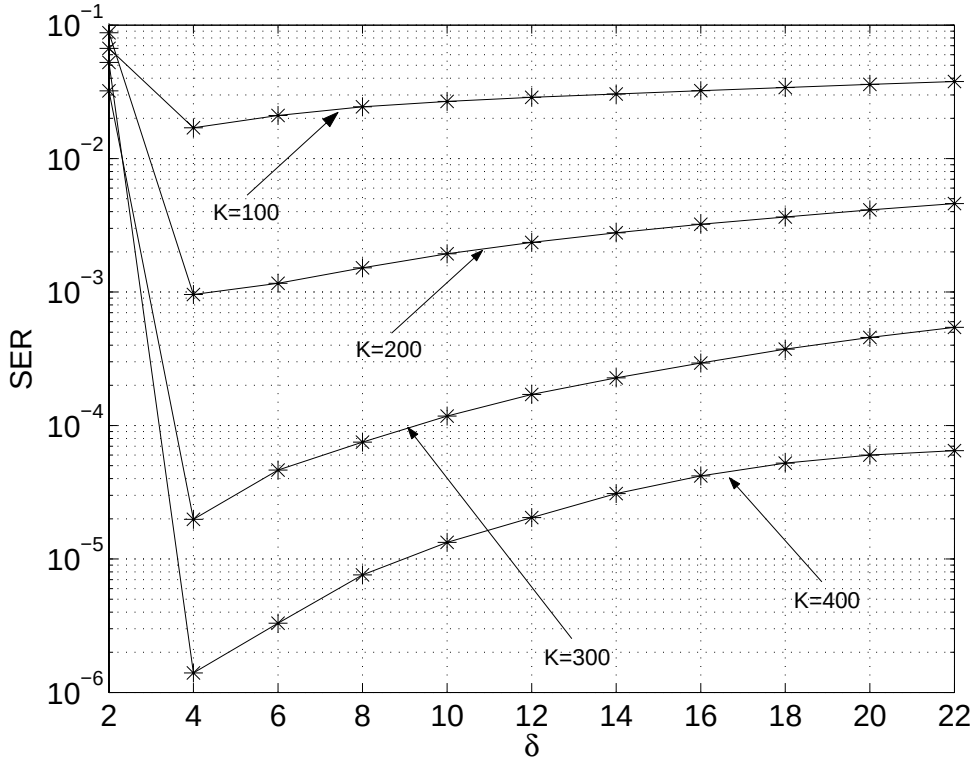

Figure 6: SER versus δ for different values of K (second example, $J = 16$, SNR $= 20$ dB).

the TS detector requires approximately 50 symbols more than the robust TS for $K < 200$, whereas, for $K \geq 200$, it requires 100 symbols more than its robust counterpart. Finally, we have reported in Fig. 10 the values of SER as a function of the number J of active users ranging from 14 to 20, where the SNR is set to 20 dB and the sample size is fixed to $K = 350$ symbols. The quadratic constraint value is chosen[7] equal to $\epsilon_0 = \delta\,\text{trace}[(\boldsymbol{\mathcal{F}}_{\text{opt}}^{(0)})^H \boldsymbol{\mathcal{Q}}_{\text{in}}\, \boldsymbol{\mathcal{Q}}_{\text{in}}^H\, \boldsymbol{\mathcal{F}}_{\text{opt}}^{(0)}]$, with δ being equal to the number of in-cell users, except for $J = 19$ and $J = 20$, where $\delta = 16$ and $\delta = 23$, respectively. Results of Fig. 10 confirm the above observations, showing that, in comparison with the TS detector, the robust TS receiver assures a substantial performance gain for small to moderate values of the number of users, i.e., for $J \leq 20$. Finally, observe that, as the number of users exceeds 16, the robust TS receiver performs comparably to or better than the blind MMSE detector, exhibiting performances that are close to those of the non-blind MMSE receiver.

[7]Results of computer simulations, not reported here, show that, for a fixed number J_{out} of strong out-of-cell users, choosing $\delta = J_{\text{in}}$ assures satisfactory performance, provided that the overall number of users J is less than the number $N - L_{\text{cp}}$ of degrees of freedom. When, however, J becomes comparable with $N - L_{\text{cp}}$, the parameter δ must be chosen much greater than J_{in}.

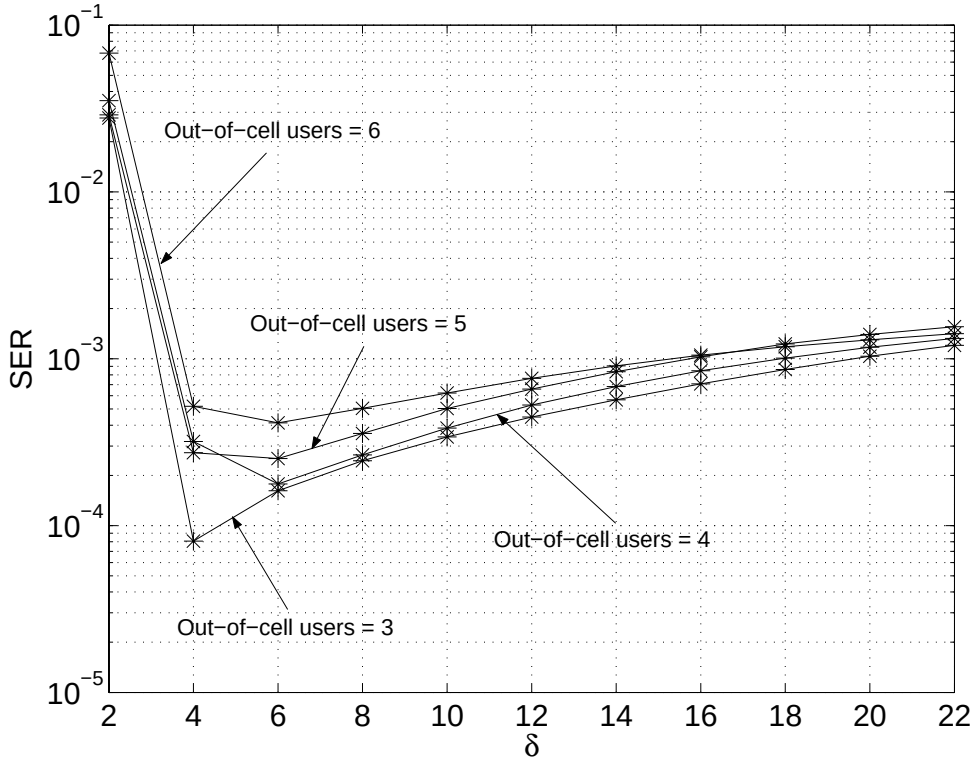

Figure 7: SER versus δ for different values of J_{out} (second example, $J_{\text{in}} = 12$, $K = 250$, SNR $= 20$ dB).

6 Conclusions

In this paper, we have theoretically analyzed the performance of the two-stage receiver recently proposed in [13], when the receiver's parameters are estimated by using a finite sample-size. Results of this analysis have suggested the formulation of a robust version of the two-stage receiver, which is based on the introduction of a suitable quadratic constraint in the synthesis of the first stage. This constraint is constructed by exploiting in the uplink the knowledge of the spreading codes of the in-cell users. The theoretical analysis has evidenced that the incorporation of the quadratic constraint has the effect of slightly reducing the degrees of freedom for disturbance suppression of the first stage, gaining robustness against errors in the estimated statistics of the received data. Moreover, results of computer simulations have shown that, even when small sample size are considered, the proposed receiver performs comparably to the non-blind MMSE receiver, outperforming the two-stage detector proposed in [13] in moderately loaded cells with strong out-of-cell MAI. Finally, our current research is aimed at investigating the feasibility of implementing the first stage of the robust two-stage receiver with recursive least squares updating, where the optimal value of the Lagrange multiplier μ_0 is adaptively adjusted at each step.

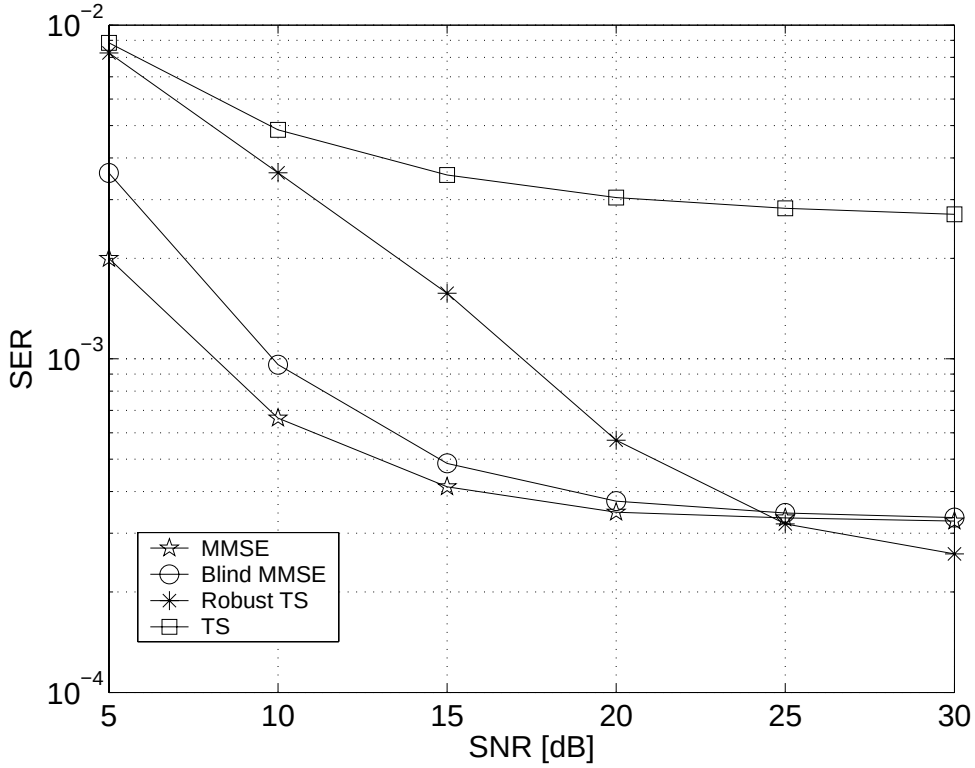

Figure 8: SER versus SNR (second example, $J = 16$, $K = 250$).

A Derivation of SINR for the optimal two-stage receiver

To evaluate the expectation in the denominator of (29), we resort to the conditional expectation rule by writing $\mathrm{E}_{\boldsymbol{d}(k),\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}}[\cdot] = \mathrm{E}_{\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}}\left\{\mathrm{E}_{\boldsymbol{d}(k)|\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}}[\cdot]\right\}$; moreover, we observe that $\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}$, being estimated from $\{\boldsymbol{r}(k)\}_{k=0}^{K-1}$, turns out to be statistically independent from $\boldsymbol{d}(k)$, provided that $k \geq K+2$ [see the signal model (7)]. Thus, one obtains

$$\text{SINR}^{(\text{I})}(\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}) = \frac{\sigma_b^2 \, \|\overline{\boldsymbol{g}}_1\|^2}{\mathrm{E}_{\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}}\left\{\text{trace}[\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}^H \boldsymbol{R}_{\boldsymbol{dd}} \widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}]\right\}} \, . \tag{58}$$

By substituting (32) in the denominator of (58) and invoking assumptions A1 and A2, we obtain, after rearrangement,

$$\mathrm{E}_{\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}}\left\{\text{trace}[\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}^H \boldsymbol{R}_{\boldsymbol{dd}} \widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}]\right\} = \text{trace}(\boldsymbol{\mathcal{F}}_{\text{opt}}^H \boldsymbol{R}_{\boldsymbol{dd}} \boldsymbol{\mathcal{F}}_{\text{opt}}) + \text{trace}\{\boldsymbol{B}_1 \, (\boldsymbol{B}_1^H \boldsymbol{R}_{\boldsymbol{dd}} \boldsymbol{B}_1)^{-1} \boldsymbol{B}_1^H \, \mathrm{E}_{\widehat{\boldsymbol{R}}}[\widehat{\boldsymbol{R}} \, \widehat{\boldsymbol{R}}^H]\} \, , \tag{59}$$

where

$$\mathrm{E}_{\widehat{\boldsymbol{R}}}[\widehat{\boldsymbol{R}} \, \widehat{\boldsymbol{R}}^H] = \frac{\|\overline{\boldsymbol{g}}_1\|^2}{K^2} \sum_{k=0}^{K-1} \sum_{h=0}^{K-1} \mathrm{E}[\boldsymbol{d}(k) \, b_1^*(k) \, b_1(h) \, \boldsymbol{d}^H(h)] \, . \tag{60}$$

From (60), accounting again for A1 and A2, one has

$$\mathrm{E}_{\widehat{\boldsymbol{R}}}[\widehat{\boldsymbol{R}} \, \widehat{\boldsymbol{R}}^H] = \frac{\sigma_b^2 \, \|\overline{\boldsymbol{g}}_1\|^2}{K} \, \boldsymbol{R}_{\boldsymbol{dd}} \, . \tag{61}$$

Finally, by substituting (61) in (59) and the result in (58), we finally get (33).

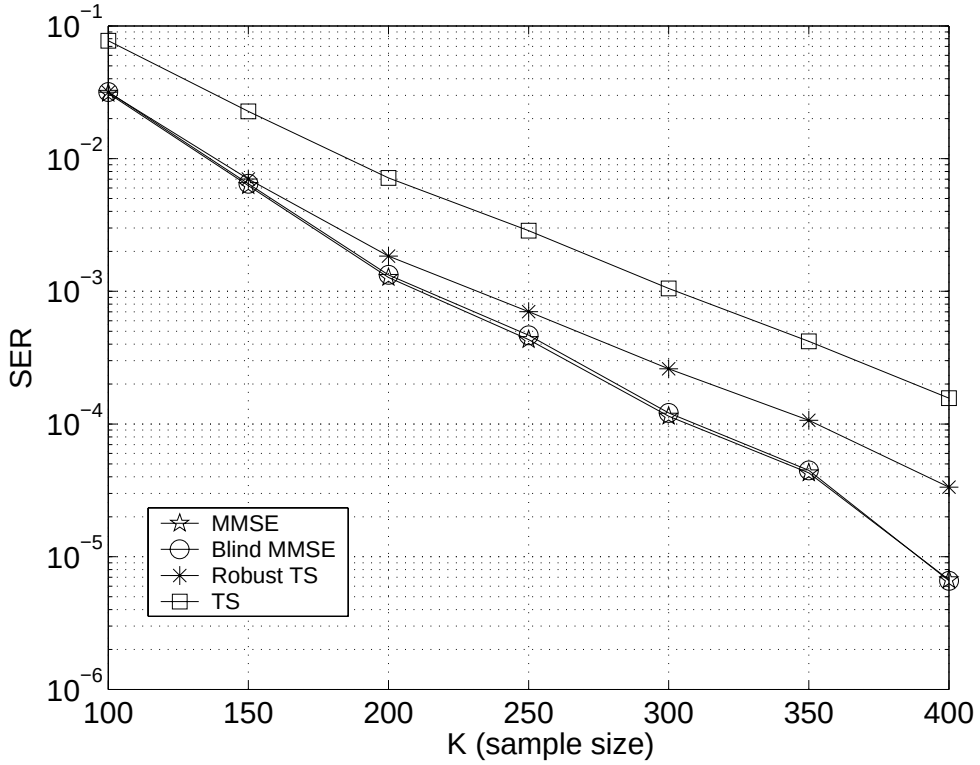

Figure 9: SER versus sample size K (second example, $J = 16$, SNR $= 20$ dB).

B Solution of the quadratically-constrained minimization problem

The problem consists of minimizing the real-valued scalar function

$$\begin{aligned} f(\boldsymbol{\mathcal{F}}^{(a)}) &\stackrel{\triangle}{=} \text{trace}\left\{[\boldsymbol{\mathcal{F}}_{\text{opt}}^{(0)} - \boldsymbol{B}_1\ \boldsymbol{\mathcal{F}}^{(a)}]^H \widehat{\boldsymbol{R}}_{\boldsymbol{rr}}\,[\boldsymbol{\mathcal{F}}_{\text{opt}}^{(0)} - \boldsymbol{B}_1\ \boldsymbol{\mathcal{F}}^{(a)}]\right\} \\ &= \text{trace}\left[(\boldsymbol{\mathcal{F}}_{\text{opt}}^{(0)})^H \widehat{\boldsymbol{R}}_{\boldsymbol{rr}}\,\boldsymbol{\mathcal{F}}_{\text{opt}}^{(0)}\right] - \text{trace}\left[(\boldsymbol{\mathcal{F}}^{(a)})^H \boldsymbol{B}_1^H \widehat{\boldsymbol{R}}_{\boldsymbol{rr}}\,\boldsymbol{\mathcal{F}}_{\text{opt}}^{(0)}\right] \\ &\quad - \text{trace}\left[(\boldsymbol{\mathcal{F}}_{\text{opt}}^{(0)})^H \widehat{\boldsymbol{R}}_{\boldsymbol{rr}}\,\boldsymbol{B}_1\,\boldsymbol{\mathcal{F}}^{(a)}\right] + \text{trace}\left[(\boldsymbol{\mathcal{F}}^{(a)})^H \boldsymbol{B}_1^H \widehat{\boldsymbol{R}}_{\boldsymbol{rr}}\,\boldsymbol{B}_1\,\boldsymbol{\mathcal{F}}^{(a)}\right] \end{aligned} \tag{62}$$

of the matrix $\boldsymbol{\mathcal{F}}^{(a)} \in \mathbb{C}^{(N-L_{\text{cp}})\times L_{\text{cp}}}$, subject to the constraint $g(\boldsymbol{\mathcal{F}}^{(a)}) \leq \epsilon_0$, where

$$\begin{aligned} g(\boldsymbol{\mathcal{F}}^{(a)}) &\stackrel{\triangle}{=} \|[\boldsymbol{\mathcal{F}}_{\text{opt}}^{(0)} - \boldsymbol{B}_1\ \boldsymbol{\mathcal{F}}^{(a)}]^H \boldsymbol{\mathcal{Q}}_{\text{in}}\|^2 = \text{trace}[(\boldsymbol{\mathcal{F}}_{\text{opt}}^{(0)})^H \boldsymbol{\mathcal{Q}}_{\text{in}}\,\boldsymbol{\mathcal{Q}}_{\text{in}}^H\,\boldsymbol{\mathcal{F}}_{\text{opt}}^{(0)}] - \text{trace}[(\boldsymbol{\mathcal{F}}^{(a)})^H \boldsymbol{B}_1^H \boldsymbol{\mathcal{Q}}_{\text{in}}\,\boldsymbol{\mathcal{Q}}_{\text{in}}^H\,\boldsymbol{\mathcal{F}}_{\text{opt}}^{(0)}] \\ &\quad - \text{trace}[(\boldsymbol{\mathcal{F}}_{\text{opt}}^{(0)})^H \boldsymbol{\mathcal{Q}}_{\text{in}}\,\boldsymbol{\mathcal{Q}}_{\text{in}}^H\,\boldsymbol{B}_1\,\boldsymbol{\mathcal{F}}^{(a)}] + \text{trace}[(\boldsymbol{\mathcal{F}}^{(a)})^H \boldsymbol{B}_1^H \boldsymbol{\mathcal{Q}}_{\text{in}}\,\boldsymbol{\mathcal{Q}}_{\text{in}}^H \boldsymbol{B}_1\,\boldsymbol{\mathcal{F}}^{(a)}]\,. \end{aligned} \tag{63}$$

By using the properties of the Kronecker product [28], the optimization problem (62)–(63) is equivalent to the minimization of the real-valued scalar function

$$\widetilde{f}(\boldsymbol{f}^{\,(a)}) \stackrel{\triangle}{=} (\boldsymbol{f}^{\,(a)})^H \widetilde{\boldsymbol{R}}_{\boldsymbol{rr}}\,\boldsymbol{f}^{\,(a)} - (\boldsymbol{f}^{\,(a)})^H \widetilde{\boldsymbol{p}}_{\boldsymbol{rr}} - \widetilde{\boldsymbol{p}}_{\boldsymbol{rr}}^H\,\boldsymbol{f}^{\,(a)} \tag{64}$$

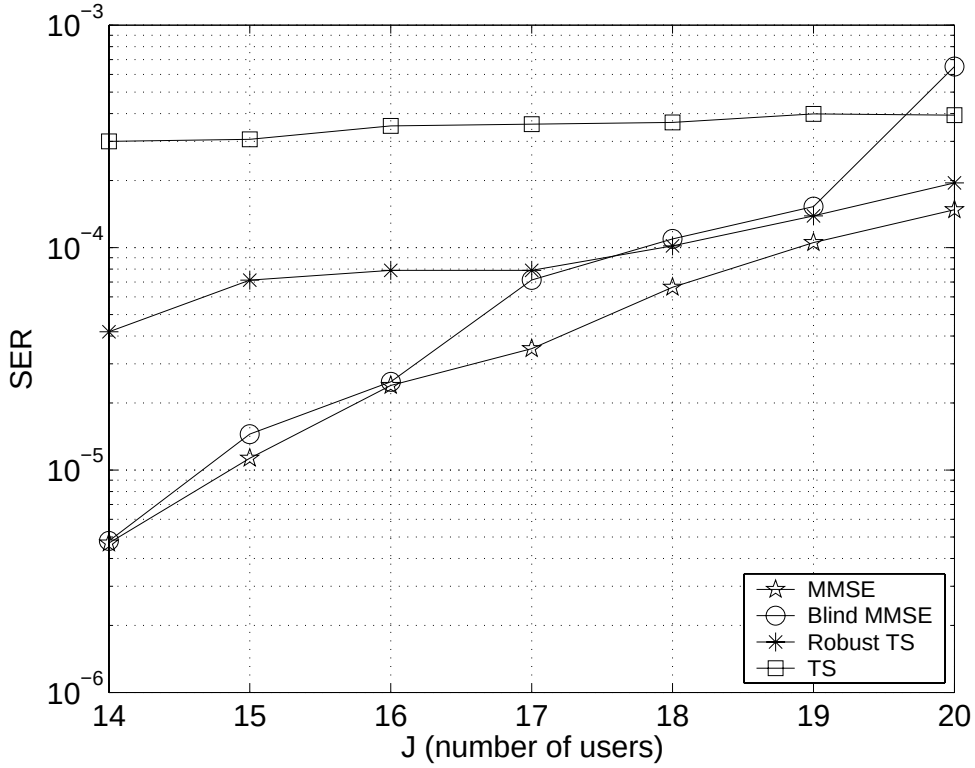

Figure 10: SER versus number J of active users (second example, $K = 350$, SNR $= 20$ dB).

of the vector $\boldsymbol{f}^{(a)} \stackrel{\triangle}{=} \mathrm{vec}[\boldsymbol{\mathcal{F}}^{(a)}] \in \mathbb{C}^{(N-L_{\mathrm{cp}})L_{\mathrm{cp}}}$, with $\widetilde{\boldsymbol{R}}_{\boldsymbol{rr}} \stackrel{\triangle}{=} \boldsymbol{I}_{L_{\mathrm{cp}}} \otimes (\boldsymbol{B}_1^H \widehat{\boldsymbol{R}}_{\boldsymbol{rr}} \boldsymbol{B}_1) \in \mathbb{C}^{(N-L_{\mathrm{cp}})L_{\mathrm{cp}} \times (N-L_{\mathrm{cp}})L_{\mathrm{cp}}}$ and $\widetilde{\boldsymbol{p}}_{\boldsymbol{rr}} \stackrel{\triangle}{=} \mathrm{vec}(\boldsymbol{B}_1^H \widehat{\boldsymbol{R}}_{\boldsymbol{rr}} \boldsymbol{\mathcal{F}}_{\mathrm{opt}}^{(0)}) \in \mathbb{C}^{(N-L_{\mathrm{cp}})L_{\mathrm{cp}}}$, subject to the constraint $\widetilde{g}(\boldsymbol{f}^{(a)}) \leq \beta_0$, where

$$\widetilde{g}(\boldsymbol{f}^{(a)}) \stackrel{\triangle}{=} (\boldsymbol{f}^{(a)})^H \widetilde{\boldsymbol{A}}\, \boldsymbol{f}^{(a)} - (\boldsymbol{f}^{(a)})^H \widetilde{\boldsymbol{q}} - \widetilde{\boldsymbol{q}}^H \boldsymbol{f}^{(a)}\,, \tag{65}$$

with $\widetilde{\boldsymbol{A}} \stackrel{\triangle}{=} \boldsymbol{I}_{L_{\mathrm{cp}}} \otimes (\boldsymbol{B}_1^H \boldsymbol{\mathcal{Q}}_{\mathrm{in}} \boldsymbol{\mathcal{Q}}_{\mathrm{in}}^H \boldsymbol{B}_1) \in \mathbb{C}^{(N-L_{\mathrm{cp}})L_{\mathrm{cp}} \times (N-L_{\mathrm{cp}})L_{\mathrm{cp}}}$, $\widetilde{\boldsymbol{q}} \stackrel{\triangle}{=} \mathrm{vec}(\boldsymbol{B}_1^H \boldsymbol{\mathcal{Q}}_{\mathrm{in}} \boldsymbol{\mathcal{Q}}_{\mathrm{in}}^H \boldsymbol{\mathcal{F}}_{\mathrm{opt}}^{(0)}) \in \mathbb{C}^{(N-L_{\mathrm{cp}})L_{\mathrm{cp}}}$ and $\beta_0 \stackrel{\triangle}{=} \epsilon_0 - \mathrm{trace}[(\boldsymbol{\mathcal{F}}_{\mathrm{opt}}^{(0)})^H \boldsymbol{\mathcal{Q}}_{\mathrm{in}} \boldsymbol{\mathcal{Q}}_{\mathrm{in}}^H \boldsymbol{\mathcal{F}}_{\mathrm{opt}}^{(0)}]$. Clearly, the constrained optimization problem (64)–(65) admits a solution if and only if

$$\min_{\boldsymbol{f}^{(a)}} \widetilde{g}(\boldsymbol{f}^{(a)}) \leq \beta_0\,. \tag{66}$$

In the following, we assume that the constraint value β_0 is set so as to satisfy condition (66). For example, under the assumption that the matrix $\boldsymbol{\mathcal{Q}}_{\mathrm{in}}$ is full-row rank, the matrix $\widetilde{\boldsymbol{A}}$ turns out to be positive definite and, thus, the function $\widetilde{g}(\boldsymbol{f}^{(a)})$ is strictly convex; in this case, it is easily seen that $\widetilde{g}(\boldsymbol{f}^{(a)})$ assumes its minimum value for $\boldsymbol{f}^{(a)} = \widetilde{\boldsymbol{A}}^{-1}\boldsymbol{q}$, which implies that $\min_{\boldsymbol{f}^{(a)}} \widetilde{g}(\boldsymbol{f}^{(a)}) = -\boldsymbol{q}^H \widetilde{\boldsymbol{A}}^{-1} \boldsymbol{q} < 0$ and, therefore, an acceptable choice for the constraint value is $\beta_0 \geq 0$ or, equivalently, $\epsilon_0 \geq \mathrm{trace}[(\boldsymbol{\mathcal{F}}_{\mathrm{opt}}^{(0)})^H \boldsymbol{\mathcal{Q}}_{\mathrm{in}} \boldsymbol{\mathcal{Q}}_{\mathrm{in}}^H \boldsymbol{\mathcal{F}}_{\mathrm{opt}}^{(0)}]$.

In order to solve the optimization problem (64)–(65), we resort to the method of Lagrange multipliers [25].

The Lagrangian for the problem at hand is defined as

$$\mathcal{L}(\boldsymbol{f}^{\,(a)};\mu_0) \stackrel{\triangle}{=} \widetilde{f}(\boldsymbol{f}^{\,(a)}) + \mu_0\,[\widetilde{g}(\boldsymbol{f}^{\,(a)}) - \beta_0] = (\boldsymbol{f}^{\,(a)})^H(\widetilde{\boldsymbol{R}}_{\boldsymbol{rr}} + \mu_0\widetilde{\boldsymbol{A}})\,\boldsymbol{f}^{\,(a)} - (\boldsymbol{f}^{\,(a)})^H(\widetilde{\boldsymbol{p}}_{\boldsymbol{rr}} + \mu_0\,\widetilde{\boldsymbol{q}}) - (\widetilde{\boldsymbol{p}}_{\boldsymbol{rr}} + \mu_0\,\widetilde{\boldsymbol{q}})^H\,\boldsymbol{f}^{\,(a)} - \mu_0\,\beta_0\,, \tag{67}$$

where $\mu_0 \geq 0$ is the Lagrange multiplier. The potential solutions of the constrained optimization problem (64)–(65) are the stationary points of $\mathcal{L}(\boldsymbol{f}^{\,(a)};\mu_0)$, that is, they satisfy the equation

$$\nabla_{(\boldsymbol{f}^{\,(a)})^*}[\mathcal{L}(\boldsymbol{f}^{\,(a)};\mu_0)] = (\widetilde{\boldsymbol{R}}_{\boldsymbol{rr}} + \mu_0\widetilde{\boldsymbol{A}})\,\boldsymbol{f}^{\,(a)} - (\widetilde{\boldsymbol{p}}_{\boldsymbol{rr}} + \mu_0\,\widetilde{\boldsymbol{q}}) = \boldsymbol{0}_{(N-L_{\rm cp})L_{\rm cp}}\,, \tag{68}$$

where $\nabla_{(\boldsymbol{f}^{\,(a)})^*}(\cdot)$ represents the complex gradient operator [26] with respect to $(\boldsymbol{f}^{\,(a)})^*$, and either $\mu_0 = 0$ or the inequality constraint is satisfied with equality [25]. Since our aim is to estimate the detector's parameters from the received data by using small to moderate values of the sample-size, we reasonably assume that the optimal solution $\widehat{\boldsymbol{f}}_{\rm opt}^{\,(a)} = {\rm vec}(\widehat{\boldsymbol{\mathcal{F}}}_{\rm opt}^{\,(a)})$, corresponding to $\mu_0 = 0$, does not allow the inequality constraint to be satisfied, that is $\widetilde{g}(\widehat{\boldsymbol{f}}_{\rm opt}^{\,(a)}) > \beta_0$; in this case, a solution $\boldsymbol{f}_{\rm rob}^{\,(a)}$ of the problem (64)–(65) necessarily occurs on the boundary of the constraint region, that is, $\widetilde{g}(\boldsymbol{f}_{\rm rob}^{\,(a)}) = \beta_0$. Since in general the Hermitian matrix $\widetilde{\boldsymbol{A}}$ is positive semidefinite[8] and $\mu_0 \geq 0$, the matrix $\widetilde{\boldsymbol{R}}_{\boldsymbol{rr}} + \mu_0\widetilde{\boldsymbol{A}}$ is positive definite; in this case, the Lagrangian is strictly convex and, thus, it is minimized, for any μ_0, by

$$\widehat{\boldsymbol{f}}_{\rm rob}^{\,(a)} = (\widetilde{\boldsymbol{R}}_{\boldsymbol{rr}} + \mu_0\widetilde{\boldsymbol{A}})^{-1}(\widetilde{\boldsymbol{p}}_{\boldsymbol{rr}} + \mu_0\,\widetilde{\boldsymbol{q}})\,. \tag{69}$$

By taking into account the above results and using the properties of the Kronecker product, it is seen that the robust solution $\widehat{\boldsymbol{f}}_{\rm rob}^{\,(a)} = {\rm vec}[\widehat{\boldsymbol{\mathcal{F}}}_{\rm rob}^{(a)}]$ can be equivalently written in matrix form as follows

$$\widehat{\boldsymbol{\mathcal{F}}}_{\rm rob}^{(a)} = [\boldsymbol{B}_1^H(\widehat{\boldsymbol{R}}_{\boldsymbol{rr}} + \mu_0\,\boldsymbol{\mathcal{Q}}_{\rm in}\,\boldsymbol{\mathcal{Q}}_{\rm in}^H)\,\boldsymbol{B}_1]^{-1}\boldsymbol{B}_1^H(\widehat{\boldsymbol{R}}_{\boldsymbol{rr}} + \mu_0\,\boldsymbol{\mathcal{Q}}_{\rm in}\,\boldsymbol{\mathcal{Q}}_{\rm in}^H)\,\boldsymbol{\mathcal{F}}_{\rm opt}^{(0)}\,, \tag{70}$$

where the optimum value of the Lagrange multiplier μ_0 is the root of the equation $g(\widehat{\boldsymbol{\mathcal{F}}}_{\rm rob}^{(a)}) = \epsilon_0$.

C Derivation of SINR for the robust two-stage receiver

To evaluate (49), we first need a suitable expression for $\widehat{\boldsymbol{\mathcal{F}}}_{\rm rob}$; to this aim, let us substitute (30) in (45), obtaining thus

$$\begin{aligned}\widehat{\boldsymbol{\mathcal{F}}}_{\rm rob}^{(a)} &= (\boldsymbol{B}_1^H\,\widehat{\boldsymbol{R}}_{\boldsymbol{rr}}\,\boldsymbol{B}_1 + \mu_0\,\boldsymbol{B}_1^H\,\boldsymbol{\mathcal{Q}}_{\rm in}\,\boldsymbol{\mathcal{Q}}_{\rm in}^H\,\boldsymbol{B}_1)^{-1}\,[\boldsymbol{B}_1^H\,\widehat{\boldsymbol{R}}_{\boldsymbol{rr}}\,\boldsymbol{\mathcal{F}}_{\rm opt}^{(0)} + \mu_0\,\boldsymbol{B}_1^H\,\boldsymbol{\mathcal{Q}}_{\rm in}\,\boldsymbol{\mathcal{Q}}_{\rm in}^H\,\boldsymbol{\mathcal{F}}_{\rm opt}^{(0)}] \\ &= (\boldsymbol{B}_1^H\,\widehat{\boldsymbol{R}}_{\boldsymbol{dd}}\,\boldsymbol{B}_1 + \mu_0\,\boldsymbol{B}_1^H\,\boldsymbol{\mathcal{Q}}_{\rm in}\,\boldsymbol{\mathcal{Q}}_{\rm in}^H\,\boldsymbol{B}_1)^{-1}[\boldsymbol{B}_1^H\,\widehat{\boldsymbol{R}}_{\boldsymbol{dd}}\,\boldsymbol{\mathcal{F}}_{\rm opt}^{(0)} + \boldsymbol{B}_1^H\,\widehat{\boldsymbol{R}} + \mu_0\,\boldsymbol{B}_1^H\,\boldsymbol{\mathcal{Q}}_{\rm in}\,\boldsymbol{\mathcal{Q}}_{\rm in}^H\,\boldsymbol{\mathcal{F}}_{\rm opt}^{(0)}]\,.\end{aligned} \tag{71}$$

[8] It turns out to be positive definite if the matrix $\boldsymbol{\mathcal{Q}}_{\rm in}$ is full-row rank.

As in Subsection 3.3, to simplify the analysis, we replace in (71) the sample correlation matrix $\widehat{\boldsymbol{R}}_{\boldsymbol{dd}}$ with the exact one $\boldsymbol{R}_{\boldsymbol{dd}}$, obtaining consequently the approximation

$$\begin{aligned}\widehat{\boldsymbol{\mathcal{F}}}_{\text{rob}}^{(a)} &\approx (\boldsymbol{B}_1^H \boldsymbol{R}_{\boldsymbol{dd}} \boldsymbol{B}_1 + \mu_0 \boldsymbol{B}_1^H \boldsymbol{\mathcal{Q}}_{\text{in}} \boldsymbol{\mathcal{Q}}_{\text{in}}^H \boldsymbol{B}_1)^{-1} [\boldsymbol{B}_1^H \boldsymbol{R}_{\boldsymbol{dd}} \boldsymbol{\mathcal{F}}_{\text{opt}}^{(0)} + \boldsymbol{B}_1^H \widehat{\boldsymbol{R}} + \mu_0 \boldsymbol{B}_1^H \boldsymbol{\mathcal{Q}}_{\text{in}} \boldsymbol{\mathcal{Q}}_{\text{in}}^H \boldsymbol{\mathcal{F}}_{\text{opt}}^{(0)}] \\ &= (\boldsymbol{I}_{N-L_{\text{cp}}} + \mu_0 \boldsymbol{\Phi} \boldsymbol{B}_1)^{-1} [\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}^{(a)} + \mu_0 \boldsymbol{\Phi} \boldsymbol{\mathcal{F}}_{\text{opt}}^{(0)}] \,, \end{aligned} \tag{72}$$

where we have defined the matrix $\boldsymbol{\Phi} \stackrel{\triangle}{=} (\boldsymbol{B}_1^H \boldsymbol{R}_{\boldsymbol{dd}} \boldsymbol{B}_1)^{-1} \boldsymbol{B}_1^H \boldsymbol{\mathcal{Q}}_{\text{in}} \boldsymbol{\mathcal{Q}}_{\text{in}}^H \in \mathbb{C}^{(N-L_{\text{cp}}) \times N}$ and, according to (32), we approximate $\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}^{(a)} \approx \boldsymbol{\mathcal{F}}_{\text{opt}}^{(a)} + (\boldsymbol{B}_1^H \boldsymbol{R}_{\boldsymbol{dd}} \boldsymbol{B}_1)^{-1} \boldsymbol{B}_1^H \widehat{\boldsymbol{R}}$. Under the assumption that the Lagrange multiplier μ_0 satisfies the relation $\mu_0 \|\boldsymbol{\Phi} \boldsymbol{B}_1\| < 1$, the following expansion holds (see [18])

$$(\boldsymbol{I}_{N-L_{\text{cp}}} + \mu_0 \boldsymbol{\Phi} \boldsymbol{B}_1)^{-1} = \sum_{\ell=0}^{+\infty} (-\mu_0)^{\ell} (\boldsymbol{\Phi} \boldsymbol{B}_1)^{\ell} \,. \tag{73}$$

In order to simplify the analysis, our aim is to obtain a first-order approximation of the $\widehat{\boldsymbol{\mathcal{F}}}_{\text{rob}}^{(a)}$ and, thus, we restrict our attention to the case where

$$\mu_0 \|\boldsymbol{\Phi} \boldsymbol{B}_1\| \ll 1 \,. \tag{74}$$

In this case, the matrix $(\boldsymbol{I}_{N-L_{\text{cp}}} + \mu_0 \boldsymbol{\Phi} \boldsymbol{B}_1)^{-1}$ is well approximated by the first two terms of expansion (73), that is, by neglecting the summands of order $o(\mu_0 \|\boldsymbol{\Phi} \boldsymbol{B}_1\|)$, one has

$$(\boldsymbol{I}_{N-L_{\text{cp}}} + \mu_0 \boldsymbol{\Phi} \boldsymbol{B}_1)^{-1} \approx \boldsymbol{I}_{N-L_{\text{cp}}} - \mu_0 \boldsymbol{\Phi} \boldsymbol{B}_1 \,. \tag{75}$$

By substituting (75) in (72) and neglecting the summand of order $o(\mu_0 \|\boldsymbol{\Phi} \boldsymbol{B}_1\|)$, one obtains, after some manipulations, the following first-order approximation

$$\widehat{\boldsymbol{\mathcal{F}}}_{\text{rob}}^{(a)} \approx \widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}^{(a)} + \mu_0 \boldsymbol{\Phi} \widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}} \,, \tag{76}$$

where, according to (32), we approximate $\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}} \approx \boldsymbol{\mathcal{F}}_{\text{opt}} - \boldsymbol{B}_1 (\boldsymbol{B}_1^H \boldsymbol{R}_{\boldsymbol{dd}} \boldsymbol{B}_1)^{-1} \boldsymbol{B}_1^H \widehat{\boldsymbol{R}}$. Equation (76), in its turn, can be substituted in (43), obtaining thus the simple expression

$$\widehat{\boldsymbol{\mathcal{F}}}_{\text{rob}} \approx (\boldsymbol{I}_N - \mu_0 \boldsymbol{\Psi} \boldsymbol{\mathcal{Q}}_{\text{in}} \boldsymbol{\mathcal{Q}}_{\text{in}}^H) \widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}} \,, \tag{77}$$

where we have defined the positive definite Hermitian matrix $\boldsymbol{\Psi} \stackrel{\triangle}{=} \boldsymbol{B}_1 (\boldsymbol{B}_1^H \boldsymbol{R}_{\boldsymbol{dd}} \boldsymbol{B}_1)^{-1} \boldsymbol{B}_1^H \in \mathbb{C}^{(N-L_{\text{cp}}) \times (N-L_{\text{cp}})}$. Equation (77) is particularly useful since it allows one to approximately and directly relate the robust filtering matrix $\widehat{\boldsymbol{\mathcal{F}}}_{\text{rob}}$ with $\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}$. Accounting for (77) and invoking assumptions A1 and A2, after rearrangement, we

obtain

$$\begin{aligned}\mathrm{E}_{\widehat{\boldsymbol{\mathcal{F}}}_{\text{rob}}}\left\{\text{trace}[\widehat{\boldsymbol{\mathcal{F}}}_{\text{rob}}^{H}\,\boldsymbol{R}_{\boldsymbol{dd}}\,\widehat{\boldsymbol{\mathcal{F}}}_{\text{rob}}]\right\} &= \mathrm{E}_{\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}}\left\{\text{trace}(\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}^{H}\,\boldsymbol{R}_{\boldsymbol{dd}}\,\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}})\right\} \\ &\quad - \mu_0\,\text{trace}\left\{\boldsymbol{\mathcal{Q}}_{\text{in}}\,\boldsymbol{\mathcal{Q}}_{\text{in}}^{H}\,\boldsymbol{\Psi}\,\boldsymbol{R}_{\boldsymbol{dd}}\,\mathrm{E}_{\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}}[\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}^{H}\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}]\right\} \\ &\quad - \mu_0\,\text{trace}\left\{\boldsymbol{R}_{\boldsymbol{dd}}\,\boldsymbol{\Psi}\,\boldsymbol{\mathcal{Q}}_{\text{in}}\,\boldsymbol{\mathcal{Q}}_{\text{in}}^{H}\,\mathrm{E}_{\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}}[\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}^{H}\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}]\right\} \\ &\quad + \mu_0^2\,\text{trace}\left\{\boldsymbol{\mathcal{Q}}_{\text{in}}\,\boldsymbol{\mathcal{Q}}_{\text{in}}^{H}\,\boldsymbol{\Psi}\,\boldsymbol{R}_{\boldsymbol{dd}}\,\boldsymbol{\Psi}\,\boldsymbol{\mathcal{Q}}_{\text{in}}\,\boldsymbol{\mathcal{Q}}_{\text{in}}^{H}\,\mathrm{E}_{\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}}[\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}^{H}\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}]\right\}\,. \end{aligned} \tag{78}$$

Observe that, under assumptions A1 and A2, the first summand in (78) has been already evaluated in Appendix A [see equations (59) and (61)], whereas, by invoking again assumptions A1 and A2, it results that

$$\mathrm{E}_{\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}}[\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}^{H}\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}] = \boldsymbol{\mathcal{F}}_{\text{opt}}^{H}\boldsymbol{\mathcal{F}}_{\text{opt}} + \boldsymbol{\Psi}\,\mathrm{E}_{\widehat{\boldsymbol{R}}}[\widehat{\boldsymbol{R}}\,\widehat{\boldsymbol{R}}^{H}]\,\boldsymbol{\Psi}\,. \tag{79}$$

By substituting (61) in (79), one has

$$\mathrm{E}_{\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}}[\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}^{H}\widehat{\boldsymbol{\mathcal{F}}}_{\text{opt}}] = \boldsymbol{\mathcal{F}}_{\text{opt}}^{H}\boldsymbol{\mathcal{F}}_{\text{opt}} + \frac{\sigma_b^2\,\|\overline{\boldsymbol{g}}_1\|^2}{K}\,\boldsymbol{\Psi}\,. \tag{80}$$

Taking into account (59), (61) and (80), equation (78) can be expressed as follows

$$\mathrm{E}_{\widehat{\boldsymbol{\mathcal{F}}}_{\text{rob}}}\left\{\text{trace}[\widehat{\boldsymbol{\mathcal{F}}}_{\text{rob}}^{H}\,\boldsymbol{R}_{\boldsymbol{dd}}\,\widehat{\boldsymbol{\mathcal{F}}}_{\text{rob}}]\right\} = \text{trace}(\boldsymbol{\mathcal{F}}_{\text{opt}}^{H}\,\boldsymbol{R}_{\boldsymbol{dd}}\,\boldsymbol{\mathcal{F}}_{\text{opt}}) + \frac{\sigma_b^2\,\|\overline{\boldsymbol{g}}_1\|^2}{K}\,[N - L_{\text{cp}} - \Omega(\mu_0)]\,, \tag{81}$$

where

$$\Omega(\mu_0) \stackrel{\triangle}{=} 2\,\mu_0\,\text{trace}(\boldsymbol{\mathcal{Q}}_{\text{in}}^{H}\,\boldsymbol{\Psi}\,\boldsymbol{\mathcal{Q}}_{\text{in}}) - \mu_0^2\left\{\text{trace}[(\boldsymbol{\mathcal{Q}}_{\text{in}}^{H}\,\boldsymbol{\Psi}\,\boldsymbol{\mathcal{Q}}_{\text{in}})^2] + \frac{K}{\sigma_b^2\,\|\overline{\boldsymbol{g}}_1\|^2}\,\text{trace}(\boldsymbol{\mathcal{F}}_{\text{opt}}^{H}\boldsymbol{\mathcal{Q}}_{\text{in}}\,\boldsymbol{\mathcal{Q}}_{\text{in}}^{H}\boldsymbol{\Psi}\,\boldsymbol{\mathcal{Q}}_{\text{in}}\,\boldsymbol{\mathcal{Q}}_{\text{in}}^{H}\,\boldsymbol{\mathcal{F}}_{\text{opt}})\right\}, \tag{82}$$

and we have used the two identities $\boldsymbol{\Psi}\,\boldsymbol{R}_{\boldsymbol{dd}}\,\boldsymbol{\Psi} = \boldsymbol{\Psi}$ and $\text{trace}(\boldsymbol{\mathcal{F}}_{\text{opt}}^{H}\boldsymbol{\mathcal{Q}}_{\text{in}}\,\boldsymbol{\mathcal{Q}}_{\text{in}}^{H}\,\boldsymbol{\Psi}\,\boldsymbol{R}_{\boldsymbol{dd}}\,\boldsymbol{\mathcal{F}}_{\text{opt}}) = 0$. Finally, by substituting (81) in (49), we obtain (51).

References

[1] N. Morinaga, M. Nakagawa and R. Kohno, "New concepts and technologies for achieving highly reliable and high capacity multimedia wireless communications systems," *IEEE Commun. Mag.*, pp. 34–40, Jan. 1997.

[2] S. Hara and R. Prasad, "Overview of multicarrier CDMA," *IEEE Commun. Mag.*, pp. 126–133, Dec. 1997.

[3] 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)*, Yokohama, Japan, September 1993, pp. 468-472.

[4] V. M. Da Silva and E. R. Sousa, “Multicarrier orthogonal CDMA signals for quasi-synchronous communication systems,” *IEEE J. Select. Areas Commun.*, pp. 842–852, June 1994.

[5] S. Tsumura and S. Hara, “Design and performance of quasi-synchronous multi-carrier CDMA system,” in *Proc. of IEEE 54th Vehicular Technology Conference, 2001 (VTC’2001 Fall)* Atlantic City, NJ, pp. 843–847, Oct. 2001.

[6] A. Chouly, A. Brajal, and S. Jourdan, “Orthogonal multicarrier technique applied to direct sequence spread spectrum CDMA systems,” in *Proc. of IEEE Global Telecommunications Conference (GLOBECOM’93)*, Houston, TE, Nov. 1993, pp. 1723-1728.

[7] J.G. Proakis, *Digital Communications*. New York: McGraw-Hill, 2001.

[8] S. Hara and R. Prasad, “Design and performance of multicarrier CDMA systems in frequency-selective Rayleigh fading channels,” *IEEE Trans. Vehicular Technology*, pp. 1584–1595, Sept. 1999.

[9] X. Gui and T.S. Ng, “Performance of asynchronous orthogonal multicarrier CDMA system in frequency selective fading channel,” *IEEE Trans. Commun.*, vol. 47, pp. 1084–1091, July 1999.

[10] N. Yee and J.-P. Linnartz, “Wiener filtering of multi-carrier CDMA in Rayleigh fading channel,” in *Proc. of 5th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, (PIMRC’94)*, Den Haag, Netherlands, September 1994. pp. 1344-1347.

[11] P. Zong, K. Wang, and Y. Bar-Ness, “Partial sampling MMSE interference suppression in asynchronous multicarrier CDMA system,” *IEEE J. Select. Areas Commun.*, pp. 1605–1613, Aug. 2001.

[12] F. Verde, “Subspace-based blind multiuser detection for quasi-synchronous MC-CDMA systems”, to appear on *IEEE Signal Processing Letters*, July 2004.

[13] D. Darsena, G. Gelli, L. Paura e F. Verde, “Blind multiuser detection for MC-CDMA systems,” in *Proc. 36nd Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, pp. 1419–1423, 2002.

[14] G. Gelli, L. Paura and F. Verde, “A two-stage CMA-based receivers for blind joint equalization and multiuser detection in high data-rate DS-CDMA systems,” to appear on *IEEE Trans. Wireless Commun.*, July 2004.

[15] A. Høst-Madsen and X. Wang, “Group-blind multiuser detection for uplink CDMA,” *IEEE J. Select. Areas Commun.*, vol. 17, pp. 1971–1984, Nov. 1999.

[16] A. Høst-Madsen and X. Wang, “Performance of blind and group-blind multiuser detectors,” *IEEE Trans. Inform. Theory*, vol. 48, pp. 1849–1872, July 2002.

[17] 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.

[18] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge: Cambridge University Press, 1990.

[19] R. A. Iltis and L. Mailaender, “Multiuser detection of quasisynchronous CDMA signals using linear decorrelators,” *IEEE Trans. Commun.*, pp. 1561–1571, Nov. 1996.

[20] 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*, pp. 1927–1950, Oct. 1998.

[21] P. Schniter and C.R. Johnson Jr., “Sufficient conditions for the local convergence of constant modulus algorithms,” *IEEE Trans. Signal Processing*, pp. 2785–2796, Oct. 2000.

[22] M. Wax and Y. Anu, “Performance analysis of the minimum variance beamformer,” *IEEE Trans. Signal Processing*, pp. 928–937, April 1996.

[23] Z. Tian, K. L. Bell and H. L. Van Trees, “Robust constrained linear receivers for CDMA wireless systems,” *IEEE Trans. Signal Processing*, pp. 1510–1522, July 2001.

[24] H. L. Van Trees, *Optimum Array Processing: Detection, Estimation, and Modulation Theory*, New York: John Wiley & Sons, Inc., 2002.

[25] D.G. Luenberger, *Optimization by Vector Space Methods*, New York: John Wiley & Sons, 1969.

[26] D. H. Brandwood. “A complex gradient operator and its application in adaptive array theory”, *IEE Proc.*, pp. 11–16, Feb. 1993.

[27] J.K. Tugnait, “Blind spatio-temporal equalization and impulse response estimation for MIMO channels using a Godard cost function,” *IEEE Trans. Signal Processing*, vol. 45, pp. 268–271, Jan. 1997.

[28] J. W. Brewer, “Kronecker products and matrix calculus in system theory,” *IEEE Trans. Circuits and Systems*, pp. 772–781, Sept. 1978.