A Systematic Approach to Multistage Detectors in Multipath Fading Channels

Laura Cottatellucci, *Student Member, IEEE,* and Ralf R. Müller, *Senior Member, IEEE*

***Abstract*—We consider linear multistage detectors with universal (large system) weighting for synchronous code-division multiple access (CDMA) in multipath fading channels with many users. A convenient choice of the basis of the projection subspace allows a joint projection of all users. Taking advantage of this property, the complexity per bit of multistage detectors with universal weights scales linearly with the number of users on the uplink CDMA channel, while other known multistage detectors with universal weights and different bases of the projection subspace keep the same quadratic complexity order per bit as the linear minimum mean-square error (LMMSE) detector. We focus on the design of two kinds of detectors with linear complexity. The detector of Type I is obtained as an asymptotic approximation of the polynomial expansion detector proposed by Moshavi *et al.* The detector of Type II has the same performance as the multistage Wiener filter (MSWF) in large systems.**

Additionally, general performance expressions for large systems, applicable to any multistage detector with the same basis of the projection subspace (e.g., linear parallel interference canceling detectors), are derived. As a by-product, the performance analysis disproves the widespread belief that the MSWF and the polynomial expansion detector are equivalent. We show that, in general, the MSWF outperforms the latter one and they are equivalent only asymptotically in the case of equal received powers.

***Index Terms*—Asymptotic weight, code-division multiple-access (CDMA) system, multipath fading channel, multistage detector, multistage Wiener filter (MSWF), multiuser detection, polynomial expansion detector, random matrix, random spreading.**

I. INTRODUCTION

SINCE wideband code-division multiple access (CDMA) has been selected for the air interface of third-generation wireless systems, significant efforts have been focused on design of detectors for signals contaminated by structured interference from other users. In fact, multiuser detection can achieve a significant increase in spectral efficiency at the cost of a considerable increase in complexity. The optimum receiver investigated in [2] allows a dramatic improvement in performance in exchange for an increase in complexity, which is exponential in the number of users. Therefore, there is a strong demand for the discovery of algorithms that simplify the signal processing required for theoretically optimum communications. The linear minimum mean-square error (LMMSE) detector has been proposed with the goal of finding an acceptable compromise between performance and complexity. It yields substantial improvements in performance, while maintaining a lower complexity than the optimum detector investigated in [2]. However, in systems with time-varying multiple-access interference (MAI)—due to, for example, long spreading sequences or fading channels—its computation in real time is very expensive. In fact, the LMMSE detector requires the inversion of matrices that are at least of size $\min(K,N) \times \min(K,N)$, where K is the number of active users and N the spreading factor. When the system size is large, its complexity is prohibitive for real-time applications.

Linear multistage detectors consist of a projector onto a subspace and a subsequent filter [1], [3]–[6]. For the filter design, different optimization criteria have been proposed [5], [3], [1] and their asymptotic performance has been analyzed both in the case of equal received powers [5], [4] and in the case of unequal received powers [7]–[9] or flat fading with binary phase-shift keying (BPSK) modulation [10], [11]. All the above cited multistage detectors use the same Krylov subspace [12]. Hereinafter, we refer to such a subspace as *the projection subspace*. It enjoys several useful properties.

- It does not need be tracked.
- The subspace rank required to achieve a fixed level of performance does not scale with the system size [4].
- The multistage detector output signal-to-interference and noise ratio (SINR) converges exponentially in the detector rank toward to the LMMSE detector output SINR [13] so that a low number of stages is sufficient to achieve near-LMMSE performance.
- Multistage detectors are even more attractive in asynchronous CDMA since there exists a convenient implementation that does not suffer from truncation effects due to finite delay-windowing effects [14], [15].
- In symbol-asynchronous but chip-synchronous systems, with sufficiently large delay, optimum multistage detectors can outperform the LMMSE detector when it is constrained to a certain observation window [14], [15].

The use of subspace methods does not allow a significant reduction in complexity by itself. In fact, the filter design, optimum in a mean-square error (MSE) sense, has the same complexity order as the LMMSE detector. A relevant reduction in complexity can be obtained by approximating the optimum filter coefficients (called also weights) by asymptotic

---

Manuscript received February 4, 2003; revised March 23, 2005. This work was supported by the Kplus Programme of the Austrian Government. The material in this paper was presented in part at the IEEE Information Theory Workshop, Bangalore, India, October 2002 and at the Thirty-Sixth Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, November 2002.

L. Cottatellucci is with Forschungszentrum Telekommunikation Wien, Tech Gate Vienna, A-1220 Vienna, Austria (e-mail: cottatellucci@ftw.at).

R. R. Müller was with Forschungszentrum Telekommunikation Wien, Tech Gate Vienna, A-1220 Vienna, Austria. He is now with the Department of Electronics and Telecommunications, Faculty of Information Technology, Mathematics and Electrical Engineering, Norwegian University of Science and Technology, Trondheim, Norway (e-mail: mueller@iet.ntnu.no).

Communicated by D. N. C. Tse, Associate Editor for Communications.

Digital Object Identifier 10.1109/TIT.2005.853298

approximations [16], [5] at the cost of a slight degradation in performance [17]. The asymptotic multistage detectors, proposed first in [16], [5], take advantage of some asymptotic properties of random matrices such as the convergence of the eigenvalue moments to deterministic limits. These are independent of the spreading sequences and the channel realizations. Since these values can be expressed as a linear function of a small set of parameters, the asymptotic multistage weights can also be computed easily and off-line as a function of the eigenvalue moments. The complexity reduction promised by the use of asymptotic filter coefficients in [16], [5] inspired studies to design asymptotic weighting in different scenarios [11], [7], [8], [18], [19]. Multistage detectors for systems with multipath fading channels have been considered only recently in independent works for the downlink [19]–[21] and the uplink [21], [18]. In [18], the multistage approach with asymptotic weights has been applied to both the multiuser channel estimation for multipath fading and symbol detection. The asymptotic performance of multistage detectors with no channel state information at the receiver is also analyzed in [18].

Thanks to the negligible complexity of the asymptotic filter design, the detector complexity order is determined by the complexity of the projection onto the subspaces. Nevertheless, the projection complexity received little attention. From a point of view of the receiver complexity, its is desirable to perform the projection for all users jointly rather than using different projectors for each user if one wants to detect for all users. In such a way, most of the calculations of the projection become identical for all users and complexity drops by a factor of K. This complexity reduction is possible only if the bases of the Krylov subspaces for all users can be chosen in an appropriate way to support the joint projection.

Fortunately, such a set of bases does exists. The low complexity of weight design and the asymptotic performance analysis of multistage detectors using such a set stems from the asymptotic convergence of the diagonal elements of random Gramian matrices and its positive powers. This convergence is established in this paper.

In this work, we design and analyze multistage detectors for CDMA systems in uplink with any kind of phase-shift keying (PSK) symbol alphabet, random spreading, and multipath fading channels. We use subspace bases supporting the joint processing of all users so that all proposed multistage detectors have a linear complexity order per bit. From a conceptual point of view, we focus on two asymptotic multistage detectors differing in the filter coefficients. Detector Type I uses a single set of weights satisfying the MMSE criterion jointly for all users. It corresponds to the polynomial expansion detector proposed in [1]. In detector Type II, the filter weights satisfy the MMSE criterion individually for each user. Detector Type II performs as well as the asymptotic multistage detectors in [11], [21] but its complexity is reduced by almost a factor of K on the uplink CDMA channel. The detector in [11], [21] will be referred to as detector Type III in the following.

Our analysis applies to a wider class of detectors than just Type I and II. It is applicable to any multistage detector using the same projection subspace bases both for finite and asymptotically large system size, e.g., the linear "standard" partial parallel interference cancellation (PIC) detectors (an analysis of a modified version of standard PIC detectors is in [22]). The asymptotic analysis can also be applied to the multistage Wiener filter (MSWF) [3]. In fact, the asymptotic performance of the MSWF, the Type III detector, and the Type II detector are the same. This observation enlightens also the actual relation between the polynomial expansion detector in [1] and the MSWF. In literature, the idea that those two detectors are equivalent is widely spread, explicitly claimed in [12], and implicitly assumed in [11]. In contrast to this position, we show that the two detectors differ and the MSWF outperforms the polynomial expansion detector in [1] for equal number of stages. The latter detector does not maximize the output SINR. This loss of optimality also affects the characteristics of its multiuser efficiency: In contrast to many of the other detectors analyzed in literature, the multiuser efficiency of the Type I detector depends on the received power of the user of interest. The MSWF and the polynomial expansion detector in [1] coincide asymptotically in case of equal received powers for all users and, under this conditions, they are also equivalent to the multistage detector proposed in [5].

Similarly to the asymptotic analysis of the LMMSE detector [23], [24], the performance and the weighting of both the Type I and Type II detectors are independent of the spreading sequences and the fading channel realizations. They depend only on few macroscopic parameters, namely, the number of user per chip, the received power statistics, the noise variance, and the received power of the user of interest. The analysis in this paper provides deep insight into the system behavior and clear guidelines for the design.

This work is structured as follows. Section II introduces the system model and the notation. In Section III, we discuss criteria for the choice of the subspace bases and for filter optimization. We analyze their impact on performance, complexity, and design. The design of Type I and Type II detectors with universal asymptotic weights is illustrated in Section IV. Section V contains the performance analysis in asymptotic conditions. Section VI presents numerical results and simulations assessing the degradation introduced by the asymptotic multistage detectors when used for finite systems and compares detector Type I and detector Type II in terms of performance. Our conclusions are in Section VII.

II. NOTATIONS AND SYSTEM MODEL

Throughout this work, the superscripts $\cdot^T$ and $\cdot^H$ denote the transpose and the conjugate transpose of the matrix argument, respectively. $\boldsymbol{I}_n$ is the identity matrix of size $n \times n$. $\mathbb{C}$ and $\mathbb{Z}^+$ are the fields of complex numbers and positive integers, respectively. $\mathrm{tr}(\,\cdot\,)$, $||\cdot||_2$, and $|\cdot|$ are the trace, the Frobenius norm, and the spectral norm of the argument, respectively, i.e.,

$$||\boldsymbol{A}||_2 = \sqrt{\mathrm{tr}(\boldsymbol{A}\boldsymbol{A}^H)}, |\boldsymbol{A}| = \max_{\boldsymbol{x}^H\boldsymbol{x}\leq 1} ||\boldsymbol{A}\boldsymbol{x}||.$$

$\mathrm{E}\{\cdot\}$ is the expectation operator. δ_{ij} is the Kronecker symbol. $\mathrm{Re}(\,\cdot\,)$ and $\mathrm{Im}(\,\cdot\,)$ are the real and imaginary parts of the argument, respectively. $\mathrm{span}\{\boldsymbol{y}_m\}|_{m=m_0}^{m_1}$, with $\boldsymbol{y}_m \in \mathbb{C}^n$, denotes the subspace in $\mathbb{C}^n$ spanned by the vectors $\{\boldsymbol{y}_{m_0}, \ldots, \boldsymbol{y}_{m_1}\}$.

$1(x)$ is the indicator function equal to 1 for $x \geq 0$ and zero elsewhere.

Let us consider a synchronous CDMA communication system with spreading factor N and K physical users, multipath fading, and additive noise at the receiver. Throughout this work, the delay spread of the channel is small compared to the symbol time such that the intersymbol interference can be neglected. Then, the equivalent baseband signals at the chip matched filter output are given by

$$\boldsymbol{y}(n) = \boldsymbol{H}(n)\boldsymbol{b}(n) + \boldsymbol{n}(n) \tag{1}$$

where $\boldsymbol{y}(n)$ is the N-dimensional received vector and $\boldsymbol{b}(n)$ is the K-dimensional column vector of transmitted symbols (one signal per each physical user) at the instant n. The transmitted symbols belong to a finite alphabet in $\mathbb{C}$, they are zero mean and satisfy the relation $\mathrm{E}\{\boldsymbol{b}(n)\boldsymbol{b}(\ell)^H\} = \boldsymbol{I}\delta_{n,\ell}$. $\boldsymbol{n}(n)$ is the N-dimensional additive noise vector at the instant n. The additive noise is circularly symmetric complex-valued white Gaussian with zero mean and variance σ^2.

The influence of spreading, transmission amplitudes, and fading is described by the $N \times K$ matrix [25]

$$\boldsymbol{H}(n) = \boldsymbol{S}(n)\boldsymbol{A}(n).$$

$\boldsymbol{A}$ is the $KL \times K$ block-diagonal matrix of received channel amplitudes taking into account the fading-channel amplitudes and the transmitted powers. It consists of blocks of size $L \times 1$, assuming that the channels have impulse responses of lengths L with $L < N$. This last condition is implied by the assumption that the intersymbol interference is negligible. $\boldsymbol{a}_k$ is the kth block-diagonal element of $\boldsymbol{A}$. The multipath channels are perfectly known at the receiver. In the asymptotic design and analysis carried out in this work, we assume that the sequence of the joint empirical distributions of $\boldsymbol{a}_k$

$$F_{\boldsymbol{a}}^{(K)}(a_1, a_2, \ldots, a_L) = \frac{1}{K}\sum_{k=1}^{K}\prod_{\ell=1}^{L} 1(a_\ell - (\boldsymbol{a}_k)_\ell)$$

converges almost surely, as $K \to \infty$, to a nonrandom limit distribution function $F_{\boldsymbol{a}}(a_1, a_2, \ldots, a_L)$ with upper-bounded support. The eigenvalues are given by $\lambda_k = \boldsymbol{a}_k^H \boldsymbol{a}_k$. Hereinafter, we denote by $F_{|\boldsymbol{A}|^2}(\lambda)$ their asymptotic distribution. The matrix of random signature sequences $\boldsymbol{S}$ is an $N \times KL$ block random matrix in $\mathbb{C}$ with blocks $\boldsymbol{S}_k = (\boldsymbol{s}_{(k-1)L+1}, \ldots \boldsymbol{s}_{kL})$, $1 \leq k \leq K$, of size $N \times L$. The elements in a column vector $\boldsymbol{s}_{(k-1)L+1}$ are independent and identically distributed (i.i.d.) with zero mean and variance $\mathrm{E}\{|s_{j,(k-1)L+1}|^2\} = \frac{1}{N}$. Additionally, they are also i.i.d. from block to block. Within a block, the vector $\boldsymbol{s}_{(k-1)L+s}$ is a cyclically shifted version of $\boldsymbol{s}_{(k-1)L+1}$ by $s - 1$ positions. These structures of the matrices $\boldsymbol{A}$ and $\boldsymbol{S}$ allow us to take into account the interchip interference due to multipath fading.

In the following, we adopt the notation:

- $\beta = (K/N)$ for the system load;
- $\boldsymbol{h}_k$ denotes the k-th column of $\boldsymbol{H}(s)$;
- $\boldsymbol{T}(n) = \boldsymbol{H}(n)\boldsymbol{H}(n)^H$;
- $\boldsymbol{R}(n) = \boldsymbol{H}(n)^H\boldsymbol{H}(n)$;
- $\boldsymbol{H}_{\sim k}(n)$ is the $N \times (K-1)$ matrix obtained from $\boldsymbol{H}(n)$ by removing the kth column;
- $\boldsymbol{T}_{\sim k}^m(n) = (\boldsymbol{H}_{\sim k}(n)\boldsymbol{H}_{\sim k}(n)^H)^m$;
- $\boldsymbol{R}_{\sim k}^m(n) = (\boldsymbol{H}_{\sim k}(n)^H\boldsymbol{H}_{\sim k}(n))^m$.

By neglecting the intersymbol interference, only quantities at the symbol-time index n appear in the system model. Therefore, the symbol-time index n will be omitted.

III. MULTISTAGE DETECTORS

A. *Definitions*

A linear multistage detector of order M for user k is a multiuser detector performing the following.

1) A projection of the observed signal onto the Krylov subspace

$$\chi_{M,k}(\boldsymbol{H}) = \mathrm{span}\{\boldsymbol{T}_{\sim k}^m \boldsymbol{h}_k\}_{m=0}^{M-1} \tag{2}$$
$$= \mathrm{span}\{\boldsymbol{T}^m \boldsymbol{h}_k\}_{m=0}^{M-1}. \tag{3}$$

Note that, although also other nonorthogonal bases slightly different have been proposed in literature, these two[1] are capable to catch the main features of all proposed nonorthogonal bases. An alternative to the bases (2) and (3) is a basis obtained by a Gram–Schmidt orthogonalization (GSO) of (3) [26]. With such a basis, it is possible to avoid the asymptotic weight design problem at the expense of the GSO, which can cause numerical problems for fixed-point arithmetic. Additionally, it does not support a modular structure based on matched filters in contrast to the nonorthogonal bases. The use of orthogonal bases exceeds the scope of this work and is not considered further.

2) A subsequent processing of the projections by a filter designed according to an optimality criterion.

The choice of the Krylov subspace is motivated by two different observations. First, as shown in [1], the full-rank LMMSE detector lies in $\chi_{K,k}(\boldsymbol{H})$, i.e., it is a linear combination of the basis vectors of $\chi_{K,k}(\boldsymbol{H})$. Second, it was established in [13] that the multistage filter output SINR converges exponentially in the filter rank M toward the full-rank LMMSE filter output SINR. Moroever, as shown in [4], under the MMSE optimality criterion, the dimension M of the subspace needed to obtain a target SINR (e.g., within a small ϵ of the full-rank SINR) does not scale with the system size (i.e., K and N).

Both the projection and the filter design can be performed jointly for all users or individually for each user. This influences both performance and complexity of the resulting multistage detector. The joint projection is obtained using the vectors in (3) as basis of $\chi_{M,k}(\boldsymbol{H})$. In this case, the projector consists of a matched filter $\boldsymbol{H}^H$ and M stages each of them performing respreading—filtering by $\boldsymbol{H}$—and successive matched filtering. The corresponding multistage detector is shown in Fig. 1. Using the vectors in (2), no joint computation of the projections is known for $M > 2$ and K different projectors are required.

For the basis (3), filter design can be performed jointly using the same filter coefficients for all users and choosing them, for example, by enforcing the minimization of the MSE averaged over all users [7]. Alternatively, we can design a different filter for each user minimizing the MSE individually. Table I shows

---

[1] About the identity of the subspaces spanned by the two bases in (2) and (3) see [4].

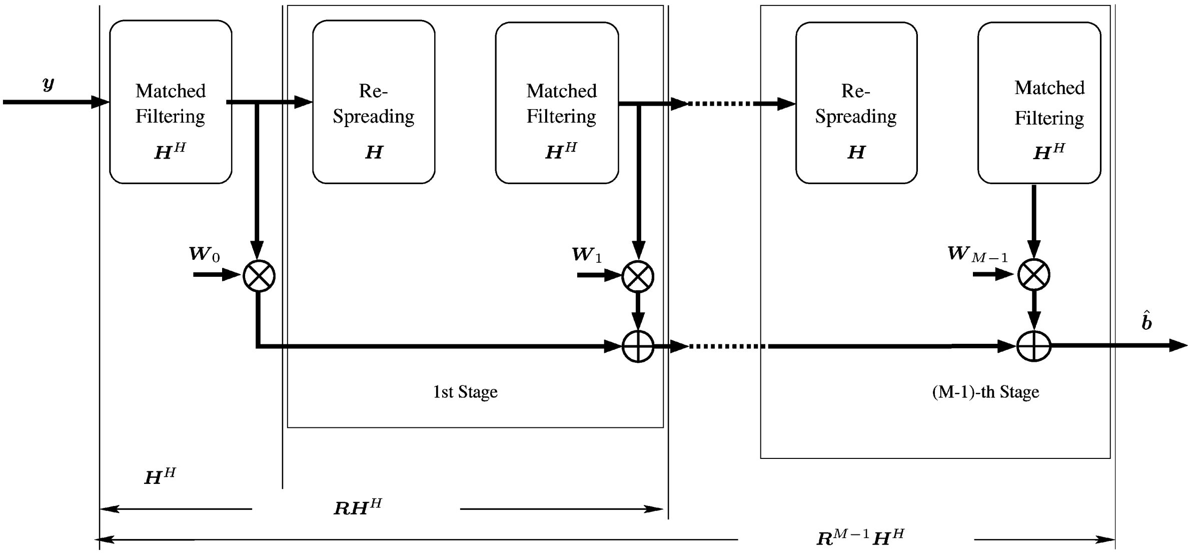

Fig. 1. Type II detector for synchronous systems.

the possible combinations and states the denominations. Detectors Type I are known as polynomial expansion detectors and were proposed in [1]. Detectors Type III are known as MSWFs and were presented first in [3]. Detectors Type II combine the advantages of detectors Type I in terms of complexity and of detectors Type III in terms of performance and are introduced in this work. Detectors Type II and Type III adopt the same optimality criterion in the same subspace and differ only in the choice of the subspace basis. Therefore, they have identical performance. However, they need, in general, different weights.

B. Complexity

To be subspace methods does not imply that the multistage detectors have lower complexity order than the full-rank LMMSE detector. In fact, if we choose the minimization of the MSE as the optimality criterion, the complexity of the filter coefficient design is identical to the complexity order of the LMMSE detector. However, by approximating the optimum filter coefficients with the corresponding asymptotic limits in large systems, i.e., as $K, N \to \infty$ with $(K/N) \to \beta$, as proposed in [16], [5], the complexity of the coefficient design becomes negligible with respect to the projection complexity. This justifies the efforts devoted to determine the asymptotic weighting in this work and, independently, in [20], [11], [21]. According to the taxonomy introduced in Table I, the asymptotic weighting of Type III detectors are designed in [20], [21] for the downlink and in [11], [21] for the uplink.

The complexity order per bit, driven by the projection complexity for detectors with asymptotic filter coefficients, is shown in Table II.[2] Table II distinguishes two cases: a single user is detected, typically in the downlink, and all users are detected at the receiver, typically in the uplink. Considering the advantages of the Type I and Type II detectors in terms of complexity with respect to Type III detectors and LMMSE detectors, we focus on Type I and Type II detectors.

TABLE I
MULTISTAGE DETECTOR CLASSIFICATION

	Joint Projection	**Individual Projection**
Joint Filtering	TYPE I	$\nexists$
Individual Filtering	TYPE II	TYPE III

TABLE II
COMPLEXITY ORDER PER SYMBOL

Detector	**One user's detection**	**All users' detection**
SUMF	$\mathcal{O}(K)$	$\mathcal{O}(K)$
TYPE I	$\mathcal{O}(K^2)$	$\mathcal{O}(K)$
TYPE II	$\mathcal{O}(K^2)$	$\mathcal{O}(K)$
TYPE III^2	$\mathcal{O}(K^2)$	$\mathcal{O}(K^2)$
LMMSE	$\mathcal{O}(K^3)$	$\mathcal{O}(K^2)$

C. Individual Filtering: Type II Detectors

Projecting the received signal onto the subspaces $\chi_{M,k}(\boldsymbol{H})$ with $M < K$, we obtain an M-dimensional nonsufficient statistic of the received signal. We denote this statistic as $\boldsymbol{x}_k$

$$\boldsymbol{x}_k \triangleq \begin{bmatrix} \boldsymbol{h}_k^H \boldsymbol{y} \\ \boldsymbol{h}_k^H \boldsymbol{T} \boldsymbol{y} \\ \vdots \\ \boldsymbol{h}_k^H \boldsymbol{T}^{M-1} \boldsymbol{y} \end{bmatrix}. \quad (4)$$

The Type II detector for user k is defined as the linear operator in $\chi_{M,k}(\boldsymbol{H})$

$$\boldsymbol{m}^H = \sum_{m=0}^{M-1} (\boldsymbol{w}_k)_m \boldsymbol{h}_k^H \boldsymbol{T}^m \quad (5)$$

---

[2] For Type III detectors with one stage ($M = 2$), an implementation with complexity $\mathcal{O}(K)$ is possible as all users are detected (e.g., uplink).

that satisfies the MMSE criterion, i.e., the weight vector $\boldsymbol{w}_k$ is given by

$$\begin{aligned}\boldsymbol{w}_k &= \arg\min_{\boldsymbol{w}_k} \mathrm{E}\{\|\boldsymbol{m}^H\boldsymbol{y} - b_k\|^2\}\\ &= \arg\min_{\boldsymbol{w}_k} \mathrm{E}\left\{\|\boldsymbol{w}_k^H\boldsymbol{x}_k - b_k\|^2\right\}. \end{aligned} \tag{6}$$

From the second expression, the Type II detector reduces to scalar LMMSE estimation on the nonsufficient statistic $\boldsymbol{x}_k$. Thus, the Wiener–Hopf theorem can be applied [27]

$$\boldsymbol{w}_k = \boldsymbol{\Phi}_k^{-1}\boldsymbol{\varphi}_k \tag{7}$$

where $\boldsymbol{\Phi}_k = \mathrm{E}\{\boldsymbol{x}_k\boldsymbol{x}_k^H\}$ and $\boldsymbol{\varphi}_k = \mathrm{E}\{b_k^*\boldsymbol{x}_k\}$. It is straightforward to verify that (8) and (9) at the bottom of the page hold, where $(\boldsymbol{R}^s)_{kk}$ is the kth diagonal element of the matrix $\boldsymbol{R}^s$.

The Type II detector is also the multistage detector in $\chi_{M,k}(\boldsymbol{H})$ that maximizes the signal-to-interference and noise ratio SINR_k of user k at the detector output. This will be shown in Section III-E.

The Type II detector for all users has structure

$$\boldsymbol{M} = \sum_{m=0}^{M-1}\boldsymbol{W}_m\boldsymbol{H}^H\boldsymbol{T}^m = \sum_{m=0}^{M-1}\boldsymbol{W}_m\boldsymbol{R}^m\boldsymbol{H}^H \tag{10}$$

where $\boldsymbol{W}_m$ is the diagonal matrix whose kth diagonal element is the mth component of $\boldsymbol{w}_k$. It minimizes $\mathrm{E}\{\|\boldsymbol{M}\boldsymbol{y} - \boldsymbol{b}\|^2\}$.

D. Joint Filtering: Type I Detectors

The Type I detector is the linear operator in $\chi_M(\boldsymbol{H}) = \mathrm{span}\{\boldsymbol{H}^H\boldsymbol{T}^m\}_{m=0}^{M-1}$

$$\boldsymbol{L} = \sum_{m=0}^{M-1} w_m\boldsymbol{H}^H\boldsymbol{T}^m = \sum_{m=0}^{M-1} w_m\boldsymbol{R}^m\boldsymbol{H}^H \tag{11}$$

such that the scalar weights w_m minimize the MSE $\mathrm{E}\{\|\sum_{m=0}^{M-1} w_m\boldsymbol{R}^m\boldsymbol{H}^H\boldsymbol{y} - \boldsymbol{b}\|^2\}$. Let us compare the joint LMMSE detector in $\chi_M(\boldsymbol{H})$ with the individual LMMSE detector. They differ in the weights: scalar weights characterize $\boldsymbol{L}$ while matrix weights appear in $\boldsymbol{M}$. The Type I detector minimizes also the MSE between the full-rank LMMSE detector output and its own output [1].

The weighting is [1]

$$\boldsymbol{w} = \boldsymbol{\Phi}^{-1}\boldsymbol{\varphi} \tag{12}$$

where the elements of the M-dimensional vector $\boldsymbol{\varphi}$ and the elements of the $M \times M$ matrix $\boldsymbol{\Phi}$ can be expressed in terms of the traces of the powers of $\boldsymbol{R}$ as $(\boldsymbol{\Phi})_{ij} = \mathrm{tr}(\boldsymbol{R}^{i+j}) + \sigma^2\mathrm{tr}(\boldsymbol{R}^{i+j-1})$ and $(\boldsymbol{\varphi})_i = \mathrm{tr}(\boldsymbol{R}^i)$. This implies

$$\boldsymbol{\varphi} = \sum_{k=1}^{K}\boldsymbol{\varphi}_k \quad\text{and}\quad \boldsymbol{\Phi} = \sum_{k=1}^{K}\boldsymbol{\Phi}_k.$$

E. Performance

For the full-rank LMMSE detector, it is well known that the minimization of the MSE per user is equivalent to the minimization of their sum and also to the maximization of the SINR at the output of the filter for each physical user [28]. This is due to the fact that no constraint is enforced on the space where the matrix of the filter lies. This property does not hold if the detector is forced to lie in a specific subspace as in the case of multistage detectors. Here, a difference between the joint minimization of the MSE proposed in [1] and the minimization of the MSE for each user proposed in [4] appears. The maximization of the SINR is achieved only in the latter case.

The MSE of user k is given by

$$\mathrm{MSE}_k = \bar{\boldsymbol{w}}_k^H\mathrm{E}\left\{\boldsymbol{x}_k\boldsymbol{x}_k^H\right\}\bar{\boldsymbol{w}}_k - 2\mathrm{Re}\left(\bar{\boldsymbol{w}}_k^H\mathrm{E}\{\boldsymbol{x}_kb_k\}\right) + 1. \tag{13}$$

for any multistage detector in $\chi_{M,k}(\boldsymbol{H})$ with weight vector $\bar{\boldsymbol{w}}_k$. Recalling that $\mathrm{E}\{\boldsymbol{x}_k\boldsymbol{x}_k^H\} = \boldsymbol{\Phi}_k$ and $\mathrm{E}\{\boldsymbol{x}_kb_k^*\} = \boldsymbol{\varphi}_k$, we obtain

$$\mathrm{MSE}_k = 1 - 2\,\mathrm{Re}\left(\boldsymbol{\varphi}_k^T\bar{\boldsymbol{w}}_k\right) + \bar{\boldsymbol{w}}_k^H\boldsymbol{\Phi}_k\bar{\boldsymbol{w}}_k. \tag{14}$$

The corresponding SINR for user k is given by

$$\mathrm{SINR}_k = \frac{P_k}{P - P_k} \tag{15}$$

where P_k is the useful power of user k at the detector output and P is the total power. For P_k and P, we have

$$P = \mathrm{E}\left\{\bar{\boldsymbol{w}}_k^H\boldsymbol{x}_k\boldsymbol{x}_k^H\bar{\boldsymbol{w}}_k\right\} = \bar{\boldsymbol{w}}_k^H\boldsymbol{\Phi}_k\bar{\boldsymbol{w}}_k \tag{16}$$

and

$$\begin{aligned}P_k &= \mathrm{E}\left\{\left|\sum_{m=0}^{M-1}(\bar{w}_k)_m\boldsymbol{h}_k^H\boldsymbol{T}^m\boldsymbol{H}\boldsymbol{e}_kb_k\right|^2\right\}\\ &= \bar{\boldsymbol{w}}_k^H\boldsymbol{\varphi}_k\boldsymbol{\varphi}_k^T\bar{\boldsymbol{w}}_k \end{aligned} \tag{17}$$

where $\boldsymbol{e}_k$ is a K-dimensional vector with all components equal to zero except the kth that is equal to 1. It yields

$$\mathrm{SINR}_k = \frac{\bar{\boldsymbol{w}}_k^H\boldsymbol{\varphi}_k\boldsymbol{\varphi}_k^T\bar{\boldsymbol{w}}_k}{\bar{\boldsymbol{w}}_k^H\left(\boldsymbol{\Phi}_k - \boldsymbol{\varphi}_k\boldsymbol{\varphi}_k^T\right)\bar{\boldsymbol{w}}_k}. \tag{18}$$

with arbitrary weight vector $\bar{\boldsymbol{w}}_k$. Specializing (14) and (18) with (7) to Type II detectors, we obtain

$$\mathrm{MSE}_{II,k} = 1 - \boldsymbol{\varphi}_k^T\boldsymbol{\Phi}_k^{-1}\boldsymbol{\varphi}_k \tag{19}$$

$$\mathrm{SINR}_{II,k} = \frac{\boldsymbol{\varphi}_k^T\boldsymbol{\Phi}_k^{-1}\boldsymbol{\varphi}_k}{1 - \boldsymbol{\varphi}_k^T\boldsymbol{\Phi}_k^{-1}\boldsymbol{\varphi}_k} \tag{20}$$

$$= \frac{1}{\mathrm{MSE}_{II,k}} - 1. \tag{21}$$

Calculating the gradient of SINR_k in (18) with respect to $\boldsymbol{w}_k$ we verify that Type II detector maximizes each SINR_k as already noticed in Section III-C.

Using (12), the preformance of the Type I detector is given by

$$\mathrm{MSE}_{I,k} = 1 - 2\boldsymbol{\varphi}_k^T\boldsymbol{\Phi}^{-1}\boldsymbol{\varphi} + \boldsymbol{\varphi}^T\boldsymbol{\Phi}^{-1}\boldsymbol{\Phi}_k\boldsymbol{\Phi}^{-1}\boldsymbol{\varphi}^T \tag{22}$$

$$\boldsymbol{\Phi}_k = \begin{pmatrix} (\boldsymbol{R}^2)_{kk} + \sigma^2(\boldsymbol{R})_{kk} & \dots & (\boldsymbol{R}^{M+1})_{kk} + \sigma^2(\boldsymbol{R}^M)_{kk}\\ (\boldsymbol{R}^3)_{kk} + \sigma^2(\boldsymbol{R}^2)_{kk} & \dots & (\boldsymbol{R}^{M+2})_{kk} + \sigma^2(\boldsymbol{R}^{M+1})_{kk}\\ \dots & \dots & \dots\\ (\boldsymbol{R}^{M+1})_{kk} + \sigma^2(\boldsymbol{R}^M)_{kk} & \dots & (\boldsymbol{R}^{2M})_{kk} + \sigma^2(\boldsymbol{R}^{2M-1})_{kk} \end{pmatrix} \tag{8}$$

and

$$\boldsymbol{\varphi}_k = ((\boldsymbol{R})_{kk}, (\boldsymbol{R}^2)_{kk}, \dots, (\boldsymbol{R}^M)_{kk})^T \tag{9}$$

$$\mathrm{SINR}_{I,k} = \frac{1}{\frac{(\boldsymbol{\varphi})^T(\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}_k(\boldsymbol{\Phi})^{-1}\boldsymbol{\varphi}}{(\boldsymbol{\varphi}(\boldsymbol{\Phi})^{-1}\boldsymbol{\varphi})^2} - 1} \tag{23}$$

$$= \frac{\left(\boldsymbol{\varphi}_k^T\boldsymbol{\Phi}^{-1}\boldsymbol{\varphi}\right)^2}{\mathrm{MSE}_{I,k} - \left(\boldsymbol{\varphi}_k^T\boldsymbol{\Phi}^{-1}\boldsymbol{\varphi} - 1\right)^2}. \tag{24}$$

The Type I detector does not null the gradient of each SINR_k or their sum. Therefore, it is not the optimum choice to maximize the SINR. It also follows that the Type II detector outperforms the Type I detector in the same projection subspace.

Note that (21), the relation between SINR_k and MSE_k also holds for the full-rank LMMSE detector, while the corresponding relation (24) is more involved for the Type I detector because of the loss of optimality. For $M \geq K$, both Type I and Type II detectors coincide with the full-rank LMMSE detector [1].

IV. ASYMPTOTIC DETECTOR DESIGN

The asymptotic multistage detectors are based on the idea of approximating the weights of the optimum multistage detectors with the corresponding weights of the detector for large systems. In fact, for finite K and N, both $\mathrm{tr}(\boldsymbol{R}^m)$ and $(\boldsymbol{R}^m)_{kk}$, for $m \in \mathbb{Z}^+$ and $k = 1, \ldots K$, are random variables because of the random assignment of the spreading sequences and of the channel gains. Their computation has complexity $\mathcal{O}(K^3)$. However, it is known that, as $K, N \rightarrow \infty$ with $\frac{K}{N} \rightarrow \beta, \mathrm{tr}(\boldsymbol{R}^m)$ tends to a deterministic value independent of the spreading sequences and depending only on the system load β and the limiting eigenvalue distribution $F_{\boldsymbol{a}}(a_1, a_2, \ldots, a_L)$. These asymptotic values can be computed at complexity $\mathcal{O}(1)$ [29] and need updating only when β and/or $F_{\boldsymbol{a}}(a_1, a_2, \ldots, a_L)$ change. We show that the same property holds also for the diagonal elements of the matrix $\boldsymbol{R}^m$ and we can efficiently use them for the design of Type II asymptotic weighting.

First, for the sake of simplicity, we show the deterministic limit for flat-fading channels and then we extend the results to multipath fading channels.

Theorem 1: Let $\boldsymbol{A}$ be a $K \times K$ diagonal matrix in $\mathbb{C}$ with bounded elements and such that the sequence of the eigenvalue distribution of $\boldsymbol{A}^H\boldsymbol{A}$ converges almost surely, as $K \rightarrow \infty$, to a nonrandom distribution function $F_{|\boldsymbol{A}|^2}(\lambda)$ with upper-bounded support. Let $\boldsymbol{S} \in \mathbb{C}^{N \times K}$ with random i.i.d. zero-mean entries with variance $\mathrm{E}\{|s_{ij}|^2\} = \frac{1}{N}$, and $\lim_{N\rightarrow\infty} \mathrm{E}\{N^3|s_{ij}|^6\} < +\infty$. Let $\boldsymbol{R} = \boldsymbol{A}^H\boldsymbol{S}^H\boldsymbol{S}\boldsymbol{A}$. Conditioned on a_{kk}, the kth diagonal element of $\boldsymbol{A}$, $(\boldsymbol{R}^\ell)_{kk}$ converges almost surely, as $N, K \rightarrow \infty$ with $\frac{K}{N} \rightarrow \beta$, to the conditionally deterministic quantity $R^\ell_{kk,\infty}$

$$R^\ell_{kk,\infty} = |a_{kk}|^2 \sum_{s=0}^{\ell-1} R^s_{kk,\infty}\beta m_{\boldsymbol{R}}^{\ell-s-1}, \qquad \ell > 1 \tag{25}$$

for any $k, \ell \in \mathbb{Z}^+$. $m^s_{\boldsymbol{R}} = \mathrm{E}\{\frac{1}{K}\mathrm{tr}(\boldsymbol{R}^s)\}$. The initial values of the recursion are $R^0_{kk,\infty} = 1$ and $m^0_{\boldsymbol{R}} = \beta^{-1}$. □

Theorem 1 is proven in Appendix I.

A closed-form expression for the moments $m^s_{\boldsymbol{R}}$ can be found in [29]. An alternative recursive expression can be obtained noting that

$$\lim_{K=\beta N\rightarrow\infty} \frac{\mathrm{trace}(\boldsymbol{R}^\ell)}{K} = \lim_{K=\beta N\rightarrow\infty} \frac{1}{K}\sum_{k=1}^{K}(\boldsymbol{R}^\ell)_{kk}$$
$$= \int R^l_\infty(\lambda)\mathrm{d}F_{|\boldsymbol{A}|^2}(\lambda) \tag{26}$$

where

$$R^\ell_\infty(\lambda) = \lambda\sum_{s=0}^{\ell-1} R^s_\infty(\lambda)\beta m_{\boldsymbol{R}}^{\ell-s-1}, \qquad \ell > 1.$$

Substituting (25) in (26) we obtain Corollary 1.

Corollary 1: Let $\boldsymbol{A}$ and $\boldsymbol{S}$ be as in Theorem 1 and let $m^\ell_{|\boldsymbol{A}|^2}$ be the eigenvalue moments of the diagonal matrix $\boldsymbol{A}\boldsymbol{A}^H$. Then, the asymptotic eigenvalue moments of $\boldsymbol{R}$ are given by

$$m^\ell_{\boldsymbol{R}} = \beta\sum_{s=0}^{\ell-1} m_{\boldsymbol{R}}^{\ell-s-1}\mathrm{E}\{|a_{kk}|^2 R^s_{kk,\infty}\}. \tag{27}$$

The initializing moment is $m^0_{\boldsymbol{R}} = \beta^{-1}$. □

Note that (27) can be also used to calculate the eigenvalue moments for random matrices whose elements s_{ij} do not satisfy the constraint on the sixth moment thanks to [30].

Theorem 1 and Corollary 1 suggest a simple algorithm to determine $R^l_{kk,\infty}$ and $m^l_{\boldsymbol{R}}$:

ALGORITHM 1

```
Initialization:  Let $\rho_0(x) = 1$ and $\mu_0 = \beta^{-1}$.
$l$th step:      • Define $\rho_{\ell+1}(x) = \beta x \sum_{s=0}^{\ell} \rho_s(x)\mu_{\ell-s}$ and
                   write it as a polynomial in $x$.
                 • Assign $\rho_{\ell+1}(|a_{kk}|^2)$ to $R^{\ell+1}_{kk,\infty}$.
                   Replace all monomials $x, x^2, \ldots, x^{\ell+1}$ in the
                   polynomial $\rho_{\ell+1}(x)$ by the moments $m^1_{|\boldsymbol{A}|^2}$,
                   $m^2_{|\boldsymbol{A}|^2}, \ldots, m^{\ell+1}_{|\boldsymbol{A}|^2}$, respectively, and assign
                   the result to $m^{\ell+1}_{\boldsymbol{R}}$.
```

A closed-form expression for $R^{k+1}_{kk,\infty}$, $k \in \mathbb{Z}^+$, is provided in Appendix II. However, the formula in Appendix II requires an exhaustive search over the sum indices since they are not explicitly given. An exhaustive search is also required in the closed-form expression for the moments $m^k_{\boldsymbol{R}}$ in [29]. Therefore, the recursive approach is more practical.

The extension of the previous results to multipath fading is supported by the following theorem.

Theorem 2: Let $\boldsymbol{S}$ be an $N \times KL$ block random matrix in $\mathbb{C}$ with blocks

$$\boldsymbol{S}_k = (\boldsymbol{s}_{(k-1)L+1}, \ldots \boldsymbol{s}_{kL}), \qquad 1 \leq k \leq K$$

of size $N \times L$. The elements in a column vector $\boldsymbol{s}_{(k-1)L+1}$ are i.i.d. with zero mean, variance $\frac{1}{N}$, and

$$\lim_{N\rightarrow\infty} \mathrm{E}\{N^3|s_{ij}|^6\} < +\infty.$$

They are also i.i.d. from block to block. Within a block, the vector $\boldsymbol{s}_{(k-1)L+s}$ is a cyclically down-shifted version of $\boldsymbol{s}_{(k-1)L+1}$ by $s-1$ positions. The empirical joint distribution of the received channel amplitudes $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_K$ converges to a limiting joint distribution with upper-bounded support $F_{\boldsymbol{a}}(a_1, a_2, \ldots, a_L)$ and the eigenvalue distribution of $\boldsymbol{R} = \boldsymbol{A}^H \boldsymbol{S}^H \boldsymbol{S} \boldsymbol{A}$ converges almost surely to a limiting distribution.[3]

Additionally, let us assume that the corresponding joint probability density function $f_{\boldsymbol{a}}(a_1, a_2, \ldots, a_L)$, for any k, is an even function[4] of $\mathrm{Re}(a_k)$ and $\mathrm{Im}(a_k)$ for any value of the parameters $(a_1, \ldots, a_{k-1}, a_{k+1}, \ldots, a_L)$. Then, the following equivalences hold.

Equivalence-1: The empirical eigenvalue distribution of $\boldsymbol{R}$ converges to the same limit as the eigenvalue distribution of the covariance matrix of a system with flat-fading channel matrix $\widetilde{\boldsymbol{A}} = (\boldsymbol{A}^H \boldsymbol{A})^{\frac{1}{2}}$ and same system load β. The same property holds for the diagonal elements $R^{\ell}_{kk,\infty}$.

Equivalence-2: The empirical eigenvalue distribution of $\boldsymbol{R}$ converges to the same limit as the eigenvalue distribution of the covariance matrix $\bar{\boldsymbol{R}} = \boldsymbol{A}^H \bar{\boldsymbol{S}}^H \bar{\boldsymbol{S}} \boldsymbol{A}$ where $\bar{\boldsymbol{S}}$ is an $N \times LK$ matrix with all i.i.d. elements. ☐

The proof is in Appendix III.

Thanks to Equivalence 1, we can apply Algorithm 1 substituting $|a_{kk}|^2$ by $\boldsymbol{a}_k^H \boldsymbol{a}_k$ and the eigenvalue moments of $\boldsymbol{A}\boldsymbol{A}^H$ by the eigenvalue moments of $\boldsymbol{A}^H \boldsymbol{A}$ (note that they differ only for a scale factor).

An algorithm to compute the diagonal elements of $\boldsymbol{R}^{\ell}$, as $N, K \to \infty$ with $\frac{K}{N} \to \beta$, in the general case (no assumptions on $f_{\boldsymbol{a}}(a_1, \ldots, a_L)$) is in Appendix III.

The conditions for Theorem 2 are verified when the channel gains are independent and Gaussian distributed. Therefore, Theorem 2 supports the conjecture in [25]. Additionally, the analysis in [25] can be extended to all multipath fading channels whose limit eigenvalue density functions satisfy the conditions of Theorem 2.

In order to derive the asymptotic weighting let us define the $M \times M$ matrices

$$\boldsymbol{\Phi}_k^{\infty} = \lim_{K=\beta N \to \infty} \boldsymbol{\Phi}_k \tag{28}$$

and

$$\boldsymbol{\Phi}^{\infty} = \lim_{K=\beta N \to \infty} \frac{\boldsymbol{\Phi}}{K}. \tag{29}$$

Their generic elements are

$$(\boldsymbol{\Phi}_k^{\infty})_{st} = R^{s+t}_{kk,\infty} + \sigma^2 R^{s+t-1}_{kk,\infty}$$

and

$$(\boldsymbol{\Phi}^{\infty})_{st} = m^{s+t}_{\boldsymbol{R}} + \sigma^2 m^{s+t-1}_{\boldsymbol{R}}$$

respectively. Additionally, let $\boldsymbol{\varphi}_k^{\infty}$ and $\boldsymbol{\varphi}^{\infty}$ be the M-dimensional vectors with respective elements $(\boldsymbol{\varphi}_k^{\infty})_s = R^s_{kk,\infty}$ and $(\boldsymbol{\varphi}^{\infty})_s = m^s_{\boldsymbol{R}}$.

The Type II detector with asymptotic weights is obtained using the weights that minimize (14) or, equivalently, maximize (18) as $K, N \to \infty$ with $\frac{K}{N} \to \infty$

$$\boldsymbol{w}_k^{\infty} = (\boldsymbol{\Phi}_k^{\infty})^{-1} \boldsymbol{\varphi}_k^{\infty}. \tag{30}$$

The asymptotic weights of the Type I detector in (11) minimize the quantity

$$\lim_{K \to \infty} \sum_{k=1}^{K} \frac{\mathrm{MSE}_k}{K} = \boldsymbol{w}^T \boldsymbol{\Phi}^{\infty} \boldsymbol{w} - 2\boldsymbol{w}^T \boldsymbol{\varphi}^{\infty} + 1. \tag{31}$$

This yields

$$\boldsymbol{w}^{\infty} = (\boldsymbol{\Phi}^{\infty})^{-1} \boldsymbol{\varphi}. \tag{32}$$

Let us consider the case when all received signals have the same power, i.e., $\boldsymbol{A}^H \boldsymbol{A} = \mathcal{P}\boldsymbol{I}$. Then, $(\boldsymbol{R}^l)_{kk}$ converges almost surely, as $K, N \to \infty$ with $\frac{K}{N} \to \beta$, to a value that does not depend on the index k due to Theorem 1 and Corollary 1.

Corollary 2: Let $\boldsymbol{S}, \boldsymbol{A}$, and $\boldsymbol{R}$ be as in Theorem 1. Additionally, let the matrix $\boldsymbol{A}$ be such that $\boldsymbol{A}^H \boldsymbol{A} = \mathcal{P}\boldsymbol{I}$. Then, for any $i, k \in \mathbb{Z}^+$, $(\boldsymbol{R}^{\ell})_{kk}$ converges almost surely, as $N, K \to \infty$ with $\frac{K}{N}$ constant, to the deterministic quantity

$$(\boldsymbol{R}^{\ell})_{kk} \xrightarrow{\text{a.s.}} m^{\ell}_{\boldsymbol{R}}. \tag{33}$$

☐

Corollary 2 ensures that $\boldsymbol{\Phi}_k^{\infty} = \boldsymbol{\Phi}^{\infty}$ and $\boldsymbol{\varphi}_k^{\infty} = \boldsymbol{\varphi}^{\infty}$ as $\boldsymbol{A}^H \boldsymbol{A} = \mathcal{P}\boldsymbol{I}$. Thus, Type I and Type II detectors coincide asymptotically in the equal power case. Additionally, in this case, the two asymptotic multistage detectors coincide also with the detector proposed in [5], which maximizes the ratio between the total useful power and the total noise and interference power at the detector output. For $\boldsymbol{A}^H \boldsymbol{A} = \mathcal{P}\boldsymbol{I}$, a closed-form expression of the eigenvalue moments is in [31].

V. ASYMPTOTIC PERFORMANCE ANALYSIS

Let us consider a multistage detector for the kth user using the basis (3) and weighting $\bar{\boldsymbol{w}}_k$. As $K, N \to \infty$ with $\frac{K}{N} \to \beta$, the MSE and the SINR are given by taking the limits of (14) and (18)

$$\mathrm{MSE}_k^{\infty} = 1 - 2\mathrm{Re}\left((\boldsymbol{\varphi}_k^{\infty})^T \bar{\boldsymbol{w}}_k \right) + \bar{\boldsymbol{w}}_k^T \boldsymbol{\Phi}_k^{\infty} \bar{\boldsymbol{w}}_k \tag{34}$$

$$\mathrm{SINR}_k^{\infty} = \frac{1}{\frac{\bar{\boldsymbol{w}}_k^H \boldsymbol{\Phi}_k^{\infty} \bar{\boldsymbol{w}}_k}{\bar{\boldsymbol{w}}_k^H \boldsymbol{\varphi}_k^{\infty} (\boldsymbol{\varphi}_k^{\infty})^T \bar{\boldsymbol{w}}_k} - 1}. \tag{35}$$

These equations can be immediately specialized to Type I and Type II detectors with (32) and (30), respectively

$$\mathrm{SINR}_{I,k}^{\infty} = \frac{1}{\frac{(\boldsymbol{\varphi}^{\infty})^T (\boldsymbol{\Phi}^{\infty})^{-1} \boldsymbol{\Phi}_k^{\infty} (\boldsymbol{\Phi}^{\infty})^{-1} \boldsymbol{\varphi}^{\infty}}{\left((\boldsymbol{\varphi}^{\infty})^T (\boldsymbol{\Phi}^{\infty})^{-1} \boldsymbol{\varphi}_k^{\infty} \right)^2} - 1} \tag{36}$$

$$\mathrm{SINR}_{II,k}^{\infty} = \frac{1}{\frac{1}{\left(\boldsymbol{\varphi}_k^{\infty}\right)^T \left(\boldsymbol{\Phi}_k^{\infty}\right)^{-1} \boldsymbol{\varphi}_k^{\infty}} - 1}. \tag{37}$$

They are the asymptotic limits of (23) and (20), respectively. In the asymptotic case, the performance depends only on the limiting distribution function $F_{\boldsymbol{A}}(a_1, \ldots, a_L), \beta$, and σ^2. If the

---

[3] This assumption is of technical nature. Indeed, we conjecture that it follows from the nature of the support of $F_{\boldsymbol{a}}(a_1, \ldots, a_L)$ and the statistics of the spreading sequences.

[4] This condition is satisfied in the case of uncorrelated Rayleigh fading for example.

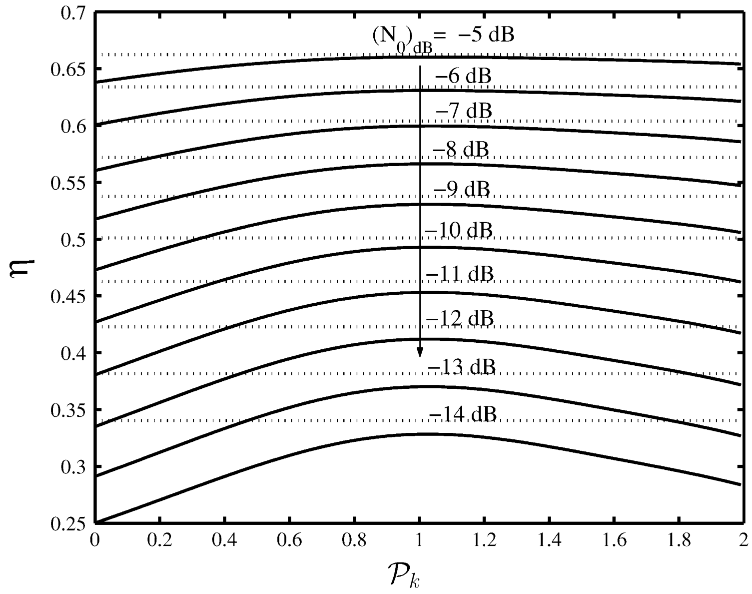

Fig. 2. Multiuser efficiency versus power of the user of interest $\mathcal{P}_k$ for Type I detector (solid line) and Type II detector (dotted line). Frequency-selective fading with exponentially decaying PDP and $L = 15$. Parameter setting: $M = 4$, $\beta = 0.5$.

conditions of Theorem 2 are fulfilled, the performance depends on the eigenvalue moments of $\boldsymbol{A}^H\boldsymbol{A}$, $m_{|\boldsymbol{A}|^2}$, the received power of user $k, \mathcal{P}_k, \beta$, and σ^2.

As shown in [4], the output SINRs of the Type II detector and of the Type III detector are proportional to the received power of user k, $\mathcal{P}_k$. Therefore, the multiuser efficiency

$$\eta_k = \frac{\sigma^2}{\mathcal{P}_k}\mathrm{SINR}_k \tag{38}$$

is independent of $\mathcal{P}_k$. In contrast and differently from many detectors analyzed in the literature, the SINR of the Type I detector depends on $\mathcal{P}_k$ by a nonlinear relation as can easily be verified by inspection. Therefore, the multiuser efficiency of this latter detector does depend on $\mathcal{P}_k$. For any user, the constant multiuser efficiency of the Type II detector is an upper bound for the multiuser efficiency of the Type I detector (see Fig. 2). In Section VI, this behavior is verified numerically.

Equations (20) and (23) allow the performance evaluation of Type I and Type II detectors when they are used in real scenarios with finite system size by setting $\bar{\boldsymbol{w}}_k = \boldsymbol{\Phi}^\infty\boldsymbol{\varphi}^\infty$ and $\bar{\boldsymbol{w}}_k = \boldsymbol{\Phi}_k^\infty\boldsymbol{\varphi}_k^\infty$, respectively. However, the performance depends on the specific realizations of $\boldsymbol{A}$ and $\boldsymbol{S}$.

VI. NUMERICAL RESULTS

Numerical results and simulations presented in this paper were obtained using an exponentially decaying power-delay profile (PDP) with a decrease of 30 dB within the channel length $L = 15$ for all users and block Rayleigh fading. Denoting the L taps of the PDP with $p_0, p_1, \ldots, p_{L-1}$, the jth tap of each channel is complex Gaussian distributed with variance p_j. Then, the characteristic function of the eigenvalues of $\boldsymbol{A}^H\boldsymbol{A}$ is given by

$$\Phi_{|\boldsymbol{A}|^2}(i\omega) = \prod_{j=0}^{L-1}\frac{1}{1-2ip_j\omega}. \tag{39}$$

We calculate the eigenvalue moments from the relation

$$i^n m_{|\boldsymbol{A}|^2}^n = (\mathrm{d}^n\Phi_{|\boldsymbol{A}|^2})/(\mathrm{d}\omega^n)|_{\omega=0}$$

the SINRs of Type I and Type II detectors in asymptotic conditions, and the multiuser efficiency for the two detectors with (38). In Fig. 3, the families of the curves η_{II}^∞ versus $(E_s)/(N_0)$ parameterized with respect to the system load β are plotted for $M = 2$ in dashed lines and for $M = 4$ in solid lines. The expected improvement in η obtained by increasing the number of stages is negligible for low $\frac{E_s}{N_0}$ and becomes more and more relevant for increasing $\frac{E_s}{N_0}$.

The performance degradation of both Type I and Type II detectors with asymptotic weighting compared to the corresponding detectors with exact weights and the full-rank LMMSE were assessed by simulations. The simulations were performed using $\frac{\pi}{4}$-QPSK modulation, in presence of multipath fading, and assuming perfect knowledge of the channel. Fig. 4 shows the bit-error rate (BER) versus $\frac{E_b}{N_0}$ for multistage detectors with $M = 4$ and $\beta = 0.5$. The performance degradation due to the asymptotic approximation of weights is completely negligible, since the curves of Type I and Type II detectors with asymptotic weights almost match the correspondent detectors with exact weighting. Using the same exponential decaying PDP for all users, the performance degradation of the Type I detector with respect to the Type II detector is irrelevant for small $\frac{E_b}{N_0}$ and becomes visible at larger $\frac{E_b}{N_0}$, as expected from the theoretic performance in Fig. 2.

Fig. 5 shows the performance improvements of Type II detector with asymptotic weights for an increasing number of stages.

VII. CONCLUSION

In this work, we identified a general framework able to catch the main features of multistage detectors with asymptotic weights in terms of performance and complexity. Both the projection onto the Krylov subspace and the filtering can be

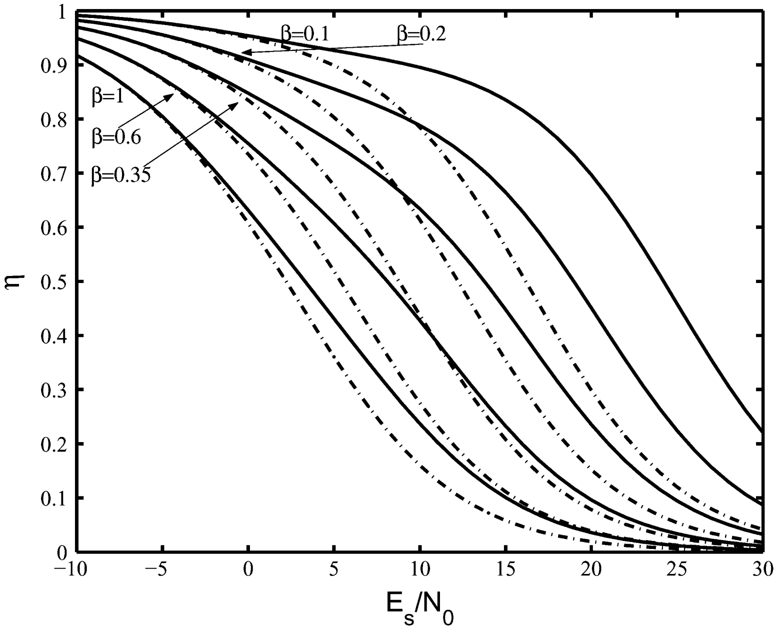

Fig. 3. Multiuser efficiency η versus signal-to-noise ratio (SNR) for Type II detector with $M = 4$ (solid line) and $M = 2$ (dashed line). Frequency-selective fading with exponentially decaying PDP and $L = 15$.

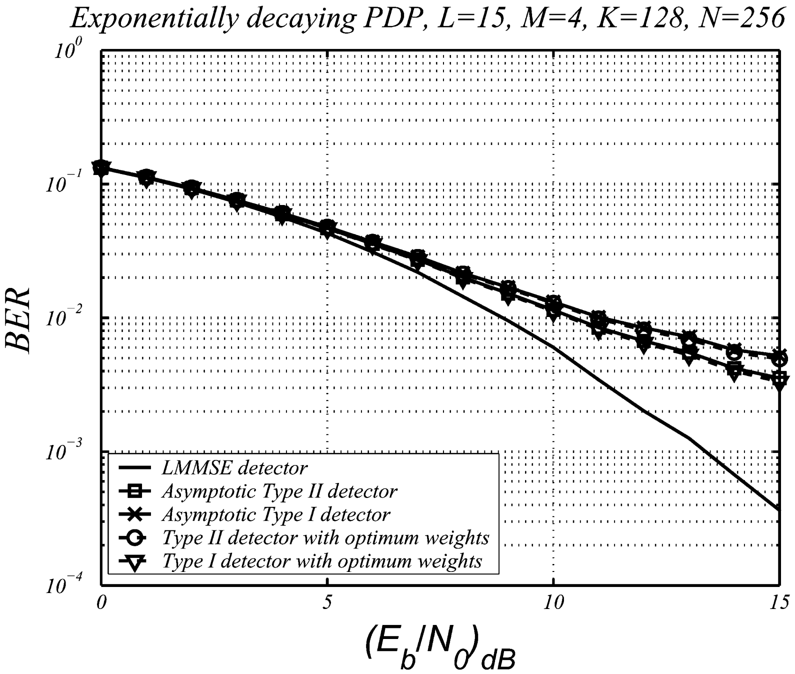

Fig. 4. BER versus $\frac{E_b}{N_0}$ for $\beta = 0.5$.

performed jointly or individually for each single user. The kind of projection affects essentially the complexity while the type of filtering has an impact on the performance.

Considerations on the projection enlightened the fact that only a joint projection can efficiently decrease the complexity order per bit from quadratic to linear.

Considerations on the individual and joint filtering disproved the belief of the equivalence of the MSWF and the polynomial expansion detector in [1].

We proposed two kinds of detectors having linear complexity order per symbol in the uplink. The choice of a basis able to support joint projection required the knowledge of the diagonal elements of $\boldsymbol{R}^\ell$ for design and analysis purposes. We could prove that, similarly to the widely used convergence of the eigenvalue moments of the matrix $\boldsymbol{R}$, also the diagonal elements of $\boldsymbol{R}^\ell$ converge to deterministic limit values depending only on the statistics of the channel and on the system load. This allowed for an efficient design and reduction of complexity.

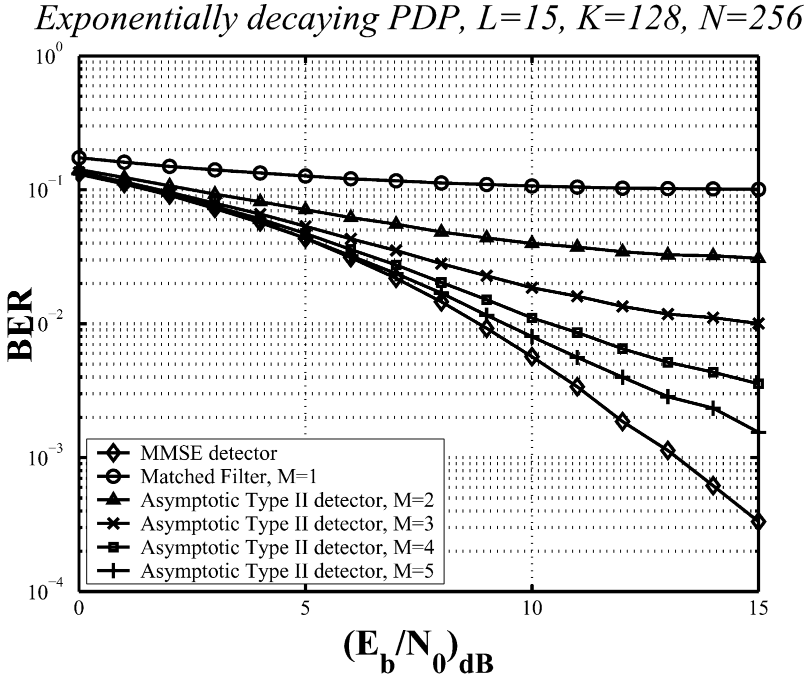

Fig. 5. BER versus $\frac{E_b}{N_0}$ for $\beta = 0.5$ and varying number of stages.

APPENDIX I
PROOF OF THEOREM 1

Proof: Let us consider any realization of the random matrix $\boldsymbol{T}^n_{\sim k}$ of size $N \times N$. Thanks to the almost sure convergence of the empirical eigenvalue distribution of $\boldsymbol{T}$ [32], $\forall \varepsilon, \delta > 0 \, \exists N'$ such that $\forall N > N'$

$$\Pr\left\{\left|\frac{1}{N}\mathrm{tr}\boldsymbol{T}^n_{\sim k} - m^n_{\boldsymbol{T}}\right| < \delta\right\} = 1 - \epsilon \tag{40}$$

where $m^n_{\boldsymbol{T}}$ denotes the limiting eigenvalue moment of order n of the matrix $\boldsymbol{T}$. Since the support of the limiting eigenvalue distribution $F_{|\boldsymbol{A}|^2}$ is upper-bounded, all its eigenvalue moments $m^n_{|\boldsymbol{A}|^2}$, $s \in \mathbb{Z}^+$ are finite. Then, the same property holds for $m^n_{\boldsymbol{T}}$ (see [29]).

By appealing to Lemma 2.7 in [33], we obtain the inequality

$$\begin{aligned} \mathrm{E}\left\{\left|\boldsymbol{s}_k^H \boldsymbol{T}^n_{\sim k}\boldsymbol{s}_k - \frac{\mathrm{tr}\boldsymbol{T}^n_{\sim k}}{N}\right|^3\right\} & \\ \leq CN\left((\mathrm{E}\{|s_{11}|^4\})^{\frac{3}{2}} + \mathrm{E}\{|s_{11}|^6\}\right) & \qquad (41) \\ \leq \frac{C^{'}}{N^2} & \qquad (42) \end{aligned}$$

where C and $C^{'}$ are constants depending on $\max((m^{2n}_{\boldsymbol{T}})^{\frac{3}{2}}, m^{6n}_{\boldsymbol{T}})$ but not on N. $\boldsymbol{s}_k$ is the kth column of $\boldsymbol{S}$. We use the Lyapunov inequality (see, e.g., [34]) to bound $\mathrm{E}\{|s_{11}|^4\}$.

The almost sure convergence as $N \to \infty$

$$\boldsymbol{s}_k^H \boldsymbol{T}^n_{\sim k}\boldsymbol{s}_k \xrightarrow{\text{a.s.}} m^n_{\boldsymbol{T}} \tag{43}$$

follows along the same lines as the proof of Lemma 4.1 in [35].

The strong law of large numbers (see, e.g., [36]) yields the almost sure convergence $\boldsymbol{s}_k^H \boldsymbol{s} \xrightarrow{\text{a.s.}} 1$ as $N \to \infty$. Then

$$\boldsymbol{R}_{kk} = |a_{kk}|^2 \boldsymbol{s}_k^H \boldsymbol{s}_k \xrightarrow{\text{a.s.}} |a_{kk}|^2. \tag{44}$$

For $\ell \geq 2$

$$\begin{aligned} (\boldsymbol{R}^\ell)_{kk} &= |a_{kk}|^2 \boldsymbol{s}_k^H \boldsymbol{T}^{\ell-1}\boldsymbol{s}_k \\ &= |a_{kk}|^2 \boldsymbol{s}_k^H \left(|a_{kk}|^2 \boldsymbol{s}_k \boldsymbol{s}_k^H + \boldsymbol{T}_{\sim k}\right) \\ &\quad \times \left(\boldsymbol{T}_{\sim k} + |a_{kk}|^2 \boldsymbol{s}_k \boldsymbol{s}_k^H\right)^{\ell-2} \boldsymbol{s}_k. \end{aligned} \tag{45}$$

Expanding the product, we can rewrite the first term as

$$\begin{aligned} |a_{kk}|^4 \boldsymbol{s}_k^H \boldsymbol{s}_k \boldsymbol{s}_k^H \left(\boldsymbol{T}_{\sim k} + |a_{kk}|^2 \boldsymbol{s}_k \boldsymbol{s}_k^H\right)^{\ell-2} \boldsymbol{s}_k & \\ = |a_{kk}|^2 \boldsymbol{s}_k^H \boldsymbol{s}_k (\boldsymbol{R}^{\ell-1})_{kk}. & \end{aligned} \tag{46}$$

The second term can be further decomposed as

$$\begin{aligned} |a_{kk}|^2 \boldsymbol{s}_k^H \boldsymbol{T}_{\sim k} \left(\boldsymbol{T}_{\sim k} + |a_{kk}|^2 \boldsymbol{s}_k \boldsymbol{s}_k^H\right)^{\ell-2} \boldsymbol{s}_k & \\ = |a_{kk}|^2 \boldsymbol{s}_k^H \boldsymbol{T}_{\sim k} \boldsymbol{s}_k (\boldsymbol{R}^{\ell-2})_{kk} & \\ + |a_{kk}|^2 \boldsymbol{s}_k^H \boldsymbol{T}^2_{\sim k} \left(\boldsymbol{T}_{\sim k} + |a_{kk}|^2 \boldsymbol{s}_k \boldsymbol{s}_k^H\right)^{\ell-3} \boldsymbol{s}_k. & \end{aligned} \tag{47}$$

Iterating the expansion (47) we get

$$(\boldsymbol{R}^\ell)_{kk} = \sum_{s=0}^{\ell-1} |a_{kk}|^2 \boldsymbol{s}_k^H \boldsymbol{T}^{\ell-1-s}_{\sim k} \boldsymbol{s}_k (\boldsymbol{R}^s)_{kk}. \tag{48}$$

Therefore, (44), (43), and the recursion yield

$$R^\ell_{kk,\infty} = \lim_{\substack{K,N\to\infty \\ \frac{K}{N}\to\beta}} (\boldsymbol{R}^s)_{kk} \xrightarrow{\text{a.s.}} \sum_{n=0}^{\ell-1} |a_{kk}|^2 m^{\ell-1-n}_{\boldsymbol{T}} R^n_{kk,\infty}. \tag{49}$$

Making use of the relation $m^n_{\boldsymbol{T}} = \beta m^n_{\boldsymbol{R}}$ we obtain (25). ■

APPENDIX II
A CLOSED-FORM EXPRESSION FOR $R^{\ell}_{kk,\infty}$

Theorem 3: Let $\boldsymbol{A}, \boldsymbol{S}$, and $\boldsymbol{R}$ be as in Theorem 1. Conditioned on a_{kk}, the kth diagonal element of $\boldsymbol{A}$, $(\boldsymbol{R}^{\ell})_{kk}$ converges almost surely, as $N, K \to \infty$, with $\frac{K}{N} \to \beta$, to the deterministic quantity $R^{\ell}_{kk,\infty}$ depending on $|a_{kk}|^2$

$$R^{\ell}_{kk,\infty} = \sum_{\substack{(i_0, i_1, \ldots i_{\ell-1}): \\ i_0 + \sum_{j=1}^{\ell-1} j i_j = \ell \\ i_0 - \sum_{j=1}^{\ell-1} i_j \geq 0}} \left(i_0 - \sum_{j=1}^{\ell-1} i_j, i_1, \ldots i_{\ell-1} \right)! \times |a_{kk}|^{2i_0} \prod_{s=1}^{\ell-1} (\beta m^s_{\boldsymbol{R}})^{i_s} \tag{50}$$

for any $k, \ell \in \mathbb{Z}^+$. $(i_0, i_1, \ldots i_{\ell-1})$ is an ℓ-tuple of nonnegative integers and $(\cdot, \cdot, \ldots, \cdot)!$ denotes the multinomial coefficient.

Proof: By expanding

$$(\boldsymbol{R}^{\ell})_{kk} = |a_{kk}|^2 \boldsymbol{s}_k^H (\boldsymbol{T}_{\sim k} + |a_{kk}|^2 \boldsymbol{s}_k \boldsymbol{s}_k^H)^{\ell-1} \boldsymbol{s}_k$$

and using the asymptotic convergence in (43) and (44), we obtain the asymptotic convergence

$$(\boldsymbol{R}^{\ell})_{kk} \xrightarrow{\text{a.s.}} \sum_{\substack{(i_0, i_1, \ldots i_{\ell-1}): \\ i_0 + \sum_{j=1}^{\ell-1} j i_j = \ell}} \varphi(i_0, i_1, \ldots, i_{\ell-1}) \times |a_{kk}|^{2i_0} \prod_{s=1}^{\ell-1} (\beta m^s_{\boldsymbol{R}})^{i_s} \tag{51}$$

where the coefficients $\varphi(i_0, i_1, \ldots, i_{\ell-1})$ are obtained expanding the binomial $(\boldsymbol{T}_{\sim k} + |a_{kk}|^2 \boldsymbol{s}_k \boldsymbol{s}_k^H)^{\ell-1}$.

Finding a closed-form expression for $R^{\ell}_{kk,\infty}$ is equivalent to the combinatorial problem of determining the coefficients $\varphi(i_0, i_1, \ldots, i_{\ell-1})$ since $m^s_{\boldsymbol{R}}$ are given in closed form in [29].

Let us consider the set $\mathcal{S}$ of all binary strings of length $\ell - 1$. Two elements in $\mathcal{S}$ are defined to be equivalent if both of them contain the same number of runs of ones with the same length, i.e., both of them contain i_1 runs of length 1, i_2 runs of length 2, i_s runs of length s for $1 \leq s \leq \ell - 1$. This equivalence relation induces a partition of $\mathcal{S}$ in classes of equivalence. The subset of the equivalent strings with i_s runs of ones with length s and, by convention, $i_0 = \ell - \sum_{s=1}^{\ell-1} s i_s$ with $i_0 \geq 1$ is denoted by $\mathcal{S}_{i_0, i_1 \ldots i_{\ell-1}}$. It is straightforward to recognize that the number of terms $c^{i_0} \prod_{s=1}^{\ell-1} (\boldsymbol{x}^H \boldsymbol{Y} \boldsymbol{x})^{i_s}$ obtained from the expansion of $c^{i_0} \boldsymbol{x}^H (\boldsymbol{Y} + c \boldsymbol{x} \boldsymbol{x}^H)^{\ell-1} \boldsymbol{x}$ is equal to the cardinality of $\mathcal{S}_{i_0, i_1 \ldots i_{\ell-1}}$. The latter equals the number of distinct permutations of a multiset with $i'_0 = i_0 - \sum_{k=1}^{\ell-1} i_k \geq 0$ elements equal to zero, i_s elements equal to s, for $1 \leq s \leq \ell - 1$, i.e., the multinomial coefficient $(i'_0, i_1, \ldots i_{\ell-1})!$ [37]. ■

APPENDIX III
PROOF OF THEOREM 2: $(\boldsymbol{R}^{\ell})_{kk,\infty}$ FOR MULTIPATH FADING CHANNELS

For $k = 1 \ldots K$ we define

- the $L(K-1) \times (K-1)$ matrix

$$\boldsymbol{A}_{\sim k} = \mathrm{diag}(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots \boldsymbol{a}_{k-1}, \boldsymbol{a}_{k+1} \ldots \boldsymbol{a}_K);$$

- the $N \times N$ permutation matrix corresponding to a cyclic down-shift by ℓ positions, $\boldsymbol{\Pi}_{\ell}$;
- The $N \times L$ spreading block of user k

$$\begin{aligned} \boldsymbol{S}_k &= (\boldsymbol{s}_{(k-1)L-1}, \ldots, \boldsymbol{s}_{kL}) \\ &= (\boldsymbol{\Pi}_0 \boldsymbol{s}_{(k-1)L-1}, \ldots, \boldsymbol{\Pi}_{L-1} \boldsymbol{s}_{(k-1)L-1}). \end{aligned}$$

Let

$$r_v^{(u)}(\boldsymbol{a}_k) = \boldsymbol{a}_k^H \boldsymbol{S}_k^H \boldsymbol{T}^{u-1} \boldsymbol{\Pi}_v \boldsymbol{S}_k \boldsymbol{a}_k, \qquad \text{for } u \in \mathbb{Z}^+.$$

By substituting $\boldsymbol{T} = \boldsymbol{S}_k \boldsymbol{a}_k \boldsymbol{a}_k^H \boldsymbol{S}_k + \boldsymbol{T}_{\sim k}$ and proceeding as in Theorem 1 we obtain

$$\begin{aligned} r_v^{(u)}(\boldsymbol{a}_k) = \sum_{\ell=0}^{u-2} & \boldsymbol{a}_k^H \boldsymbol{S}_k^H \boldsymbol{T}_{\sim k}^{\ell} \boldsymbol{S}_k \boldsymbol{a}_k r_v^{(u-\ell-1)}(\boldsymbol{a}_k) \\ & + \boldsymbol{a}_k^H \boldsymbol{S}_k^H \boldsymbol{T}_{\sim k}^{u-1} \boldsymbol{\Pi}_v \boldsymbol{S}_k \boldsymbol{a}_k, \qquad \text{for } u \in \mathbb{Z}^+ \end{aligned} \tag{52}$$

with the convention $\sum_{\ell=0}^{-1}(\,\cdot\,) = 0$.

Let us define (53) at the bottom of the page.

Using the same arguments as in Theorem 1 it is straightforward to show the following limits for $N \to \infty$:

$$\begin{aligned} \boldsymbol{s}_k^H \boldsymbol{\Pi}_s \boldsymbol{s}_k &\xrightarrow{\text{a.s.}} \delta_{0,s} \\ \boldsymbol{S}_k^H \boldsymbol{\Pi}_s \boldsymbol{S}_k &\xrightarrow{\text{a.s.}} \boldsymbol{I}_L(s) \\ \boldsymbol{S}_k^H \boldsymbol{T}_{\sim k} \boldsymbol{\Pi}_s \boldsymbol{S}_k &\xrightarrow{\text{a.s.}} \boldsymbol{\mathcal{T}}_s^{\ell} \end{aligned}$$

where $\boldsymbol{I}_L(s)$ denotes an $L \times L$ matrix with $i_{u,u+s} = 1$, for $u = 1, \ldots, L - s$, and zero elsewhere.

Then, $r_v^{(u)}(\boldsymbol{a}_k)$ converges almost surely to the deterministic limit $\rho_v^{(u)}(\boldsymbol{a}_k)$ where $\rho_v^{(u)}(\boldsymbol{a})$ satisfies the recursive expression

$$\rho_v^{(u)}(\boldsymbol{a}) = \sum_{\ell=0}^{u-2} \boldsymbol{a}^H \boldsymbol{\mathcal{T}}_0^{\ell} \boldsymbol{a} \rho_v^{(u-\ell-1)}(\boldsymbol{a}) + \boldsymbol{a}^H \boldsymbol{\mathcal{T}}_v^{u-1} \boldsymbol{a}. \tag{54}$$

$\boldsymbol{\mathcal{T}}_s^0 = \boldsymbol{I}_L(s)$. The matrix $\boldsymbol{\mathcal{T}}_s^{\ell}$ can be obtained as

$$\boldsymbol{\mathcal{T}}_s^{\ell} = \beta \begin{bmatrix} \bar{\rho}_s^{(\ell)} & \bar{\rho}_{s+1}^{(\ell)} & \ddots & \bar{\rho}_{s+L-1}^{(\ell)} \\ \bar{\rho}_{s-1}^{(\ell)} & \bar{\rho}_s^{(\ell)} & \ddots & \bar{\rho}_{s+L-2}^{(\ell)} \\ \ddots & \ddots & \ddots & \ddots \\ \bar{\rho}_{s-L+1}^{(\ell)} & \ddots & \ddots & \bar{\rho}_s^{(\ell)} \end{bmatrix} \tag{55}$$

where $\bar{\rho}_j^{(\ell)} = \mathrm{E}\{\rho_j^{\ell}\}$. Note that $\bar{\rho}_{-v}^{(\ell)} = (\bar{\rho}_v^{(\ell)})^*$.

$$\boldsymbol{\mathcal{T}}_s^{\ell} = \lim_{\substack{K,N \to \infty \\ \frac{K}{N} \to \infty}} \begin{bmatrix} \frac{1}{N}\mathrm{tr}(\boldsymbol{T}^{\ell} \boldsymbol{\Pi}_s) & \frac{1}{N}\mathrm{tr}(\boldsymbol{T}^{\ell} \boldsymbol{\Pi}_{s+1}) & \ddots & \frac{1}{N}\mathrm{tr}(\boldsymbol{T}^{\ell} \boldsymbol{\Pi}_{s+L-1}) \\ \frac{1}{N}\mathrm{tr}(\boldsymbol{T}^{\ell} \boldsymbol{\Pi}_{s-1}) & \frac{1}{N}\mathrm{tr}(\boldsymbol{T}^{\ell} \boldsymbol{\Pi}_s) & \ddots & \ddots \\ \ddots & \ddots & \ddots & \ddots \\ \frac{1}{N}\mathrm{tr}(\boldsymbol{T}^{\ell} \boldsymbol{\Pi}_{s-L+1}) & \ddots & \ddots & \frac{1}{N}\mathrm{tr}(\boldsymbol{T}^{\ell} \boldsymbol{\Pi}_s) \end{bmatrix}. \tag{53}$$

In order to prove Equivalence 1 we show first that $\bar{\rho}_v^{(u)} = 0$ for $v \neq 0$ and $u \in \mathbb{Z}^+$. This property is shown by induction. We have

$$\bar{\rho}_v^{(1)} = \begin{cases} \mathrm{E}\left\{\sum_{i=1}^{L-v} a_i^* a_{i+v}\right\}, & 1 \leq v \leq L-1 \\ 0, & v \geq L. \end{cases} \tag{56}$$

Since the argument of the expectation $\mathrm{E}\{a_i^* a_{i+v}\}$ is an odd function in a_i while the probability density function is even, $\bar{\rho}_v^{(1)} = 0 \; \forall v \geq 1$. Note that $\boldsymbol{T}_0^{(0)}$ is proportional to the identity matrix. For the induction in s, we assume that $\bar{\rho}_v^{(s)} = 0$ for $v \neq 0$ and $s < u$. Then, all diagonal elements of $\boldsymbol{T}_v^{(s)}$, for $v \neq 0$ and $s < u$, are zero and $\boldsymbol{T}_0^u$ is proportional to the identity matrix. This implies that $\rho_v^{(u)} = \boldsymbol{a}^H \boldsymbol{T}_v^{u-1} \boldsymbol{a}$, for $v \in \mathbb{Z}^+$, is an odd function in all variables that appear in it so that $\bar{\rho}_v^{(u)} = 0 \forall v \neq 0$. This completes the induction. Moreover, $\boldsymbol{T}_0^u$ is proportional to the identity matrix.

By using induction again and the fact that $\boldsymbol{T}_0^u = \beta \bar{\rho}_0^{(u)} \boldsymbol{I}_L$, $\forall\, u \in \mathbb{Z}^+$, with the convention that $\bar{\rho}_0^{(0)} = (1)/(\beta)$ we obtain

$$\rho_0^{(u)}(\boldsymbol{a}) = \sum_{\ell=0}^{u-1} \boldsymbol{a}^H \boldsymbol{a} \beta \bar{\rho}_0^{(\ell)} \rho_0^{(u-\ell-1)}(\boldsymbol{a}). \tag{57}$$

Since $\rho_0^{(u)}$ depends on $\boldsymbol{a}$ only through the scalar $\mathcal{P} = \boldsymbol{a}^H \boldsymbol{a}$, the previous recursion is rewritten as

$$\widetilde{\rho}_0^{(u)}(\mathcal{P}) = \sum_{\ell=0}^{u-1} \mathcal{P} \beta \bar{\rho}_0^{(\ell)} \widetilde{\rho}_0^{(u-\ell-1)}(\mathcal{P}). \tag{58}$$

Since $R_{kk,\infty}^u \xrightarrow{\text{a.s.}} \widetilde{\rho}_0^u(\mathcal{P}_k)$, where $\mathcal{P}_k = \boldsymbol{a}_k^H \boldsymbol{a}_k$, and (58) coincides with (25) once done the obvious substitutions, Equivalence 1 is proven.

To prove Equivalence 2 we introduce the $N \times LK$ matrix $\bar{\boldsymbol{S}}$ whose elements are i.i.d., zero mean with variance $\mathrm{E}\{|\bar{s}_{11}|^2\} = \frac{1}{N}$, and sixth moment such that

$$\lim_{N \to \infty} \mathrm{E}\{N^3 |\bar{s}_{ij}|^6\} < \infty.$$

$\bar{\boldsymbol{s}}_k$ is the kth column of $\bar{\boldsymbol{S}}$. $\bar{\boldsymbol{S}}_k = (\bar{\boldsymbol{s}}_{(k-1)L+1}, \ldots, \bar{\boldsymbol{s}}_{kL})$ is the $N \times L$ "spreading" block of user k, and $\bar{\boldsymbol{S}}_{\sim k} = (\bar{\boldsymbol{S}}_1, \ldots, \bar{\boldsymbol{S}}_{k-1}, \bar{\boldsymbol{S}}_{k+1}, \ldots \bar{\boldsymbol{S}}_K)$. $\bar{\boldsymbol{T}}, \bar{\boldsymbol{T}}_{\sim k}$ and $\bar{\boldsymbol{R}}$ are defined similarly to $\boldsymbol{T}$, $\boldsymbol{T}_{\sim k}$, and $\boldsymbol{R}$, respectively, substituting $\boldsymbol{S}$ with $\bar{\boldsymbol{S}}$. An approach similar to Lemma 2.7 in [33] (see also [25]) yields the inequality $\mathrm{E}\{\bar{\boldsymbol{s}}_\ell^H \bar{\boldsymbol{T}}_{\sim k}^u \bar{\boldsymbol{s}}_m\} \leq \frac{k_p}{N^3} \mathrm{tr} \bar{\boldsymbol{T}}_{\sim k}^u$, for $\ell, m = (k-1)L+1, \ldots, kL, \ell \neq m$, and k_p scalar independent of N. Then, for $K, N \to \infty$ with $\frac{K}{N} \to \beta$

$$\bar{\boldsymbol{s}}_\ell^H \bar{\boldsymbol{T}}_{\sim k}^u \bar{\boldsymbol{s}}_m \xrightarrow{\text{a.s.}} 0. \tag{59}$$

An expansion of $(\bar{\boldsymbol{R}}^u)_{kk}$ along the lines of (52) yields

$$(\bar{\boldsymbol{R}}^u)_{kk} = \sum_{\ell=0}^{u-2} \boldsymbol{a}_k^H \bar{\boldsymbol{S}}_k^H \bar{\boldsymbol{T}}_{\sim k}^\ell \bar{\boldsymbol{S}}_k \boldsymbol{a}_k (\bar{\boldsymbol{R}}^{u-\ell-1})_{kk}. \tag{60}$$

Since $\bar{\boldsymbol{S}}_k^H \bar{\boldsymbol{T}}_k^\ell \bar{\boldsymbol{S}}_k \xrightarrow{\text{a.s.}} m_{\bar{\boldsymbol{T}}}^\ell \boldsymbol{I}_L(0)$, it is straightforward to recognize that (60) and (58) coincide asymptotically.

This proof suggests also an algorithm to determine $R_{kk,\infty}^l$ and $m_{\boldsymbol{R}}^l$ for channels where the limiting probability density function is not even.

ALGORITHM 2

INITIALIZATION: Let $\boldsymbol{x} = (x_1, x_2, \ldots x_L)^T, \boldsymbol{T}_0^s = \boldsymbol{I}_L(s)$, and $\rho_v^{(1)}(\boldsymbol{x}) = \sum_{r=1}^{L-v} x_{v+r}^* x_r, \; v = 1, \ldots L-1$.

RECURSION:

- Assume $\rho_v^{(s)} = 0$ for $s = 1, \ldots, \ell - 1$ and

$$(\ell - 1)(L - 1) \leq v \leq \ell(L-1).$$

- Define

$$\rho_v^{(\ell)}(x_1, x_2, \ldots, x_L) = \sum_{s=0}^{\ell-2} \boldsymbol{x}^H \boldsymbol{T}_0^s \boldsymbol{x} \rho_v^{\ell-s-1} + \boldsymbol{x}^H \boldsymbol{T}_v^{\ell-1} \boldsymbol{x}. \qquad 0 \leq v \leq \ell(L-1). \tag{61}$$

  and write them as polynomial in $x_1, x_2, \ldots, x_L$.
- Replace all monomials $\prod_{\ell=1}^L x_\ell^{i_\ell}$ by the mixed moments

$$m_{\boldsymbol{A}}^{(i_1 \ldots i_L)} = \mathrm{E}\{a_1^{i_1} a_2^{i_2} \ldots a_L^{i_L}\}$$

  in $\rho_v^{(\ell)}(x_1, x_2, \ldots, x_L)$, $v = 1, \ldots, \ell(L-1)$ and assign the result to $\bar{\rho}_v^{(\ell)}$.
- Build the matrices

$$\boldsymbol{T}_v^\ell = \beta \begin{bmatrix} \bar{\rho}_v^\ell & \bar{\rho}_{v+1}^\ell & \ddots & \ddots & \bar{\rho}_{v+L-1}^\ell \\ \bar{\rho}_{v-1}^\ell & \bar{\rho}_v^\ell & \ddots & \ddots & \bar{\rho}_{v+L-2}^\ell \\ \ddots & \ddots & \ddots & \ddots & \ddots \\ \bar{\rho}_{v-L+2}^\ell & \ddots & \ddots & \ddots & \bar{\rho}_{v+1}^\ell \\ \bar{\rho}_{v-L+1}^\ell & \bar{\rho}_{v-L+2}^\ell & \ddots & \bar{\rho}_{v-1}^\ell & \bar{\rho}_v^\ell \end{bmatrix} \tag{62}$$

  by using the relation $\bar{\rho}_{-v}^{(\ell)} = (\bar{\rho}_v^{(\ell)})^*$.
- Assign $\rho_0^{(\ell)}(\boldsymbol{a}_k)$ to $R_{kk,\infty}^\ell$ *and* $\bar{\rho}_0^{(\ell)}$ to $m_{\boldsymbol{R}}^\ell$.

REFERENCES

[1] S. Moshavi, E. G. Kanterakis, and D. L. Schilling, "Multistage linear receivers for DS-CDMA systems," *Int. J. Wireless Inf. Netw.*, vol. 3, no. 1, pp. 1–17, Jan. 1996.

[2] S. Verdú, "Minimum probability of error for asynchronous Gaussian multiple-access channels," *IEEE Trans. Inf. Theory*, vol. IT-32, no. 1, pp. 85–96, Jan. 1986.

[3] J. S. Goldstein, I. S. Reed, and L. L. Scharf, "A multistage representation of the Wiener filter based on orthogonal projections," *IEEE Trans. Inf. Theory*, vol. 44, no. 7, Nov. 1998.

[4] M. Honig and W. Xiao, "Performance of reduced-rank linear interference suppression," *IEEE Trans. Inf. Theory*, vol. 47, no. 5, pp. 1928–1946, Jul. 2001.

[5] R. R. Müller and S. Verdú, "Design and analysis of low-complexity interference mitigation on vector channels," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 8, pp. 1429–1441, Aug. 2001.

[6] D. Guo, L. K. Rasmussen, and T. J. Lim, "Linear parallel interference cancellation in long-code CDMA multiuser detection," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 12, pp. 2074–2081, Dec. 1999.

[7] L. Cottatellucci and R. R. Müller, "Asymptotic design and analysis of multistage detectors with unequal powers," in *Proc. Information Theory Workshop*, Bangalore, Karnataka, India, Oct. 2002.

[8] ——, "Multiuser interference mitigation with multistage detectors: Design and analysis for unequal powers," in *Proc. Asilomar Conf. Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 2002, pp. 1948–1952.

[9] L. G. F. Trichard, J. S. Evans, and I. B. Collings, "Optimal linear multistage receiver with unequal power users," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Yokohama, Japan, Jun./Jul. 2003, p. 390.

[10] A. M. Tulino and S. Verdú, "Asymptotic analysis of improved linear receivers for BPSK-CDMA subject to fading," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 8, pp. 1544–1555, Aug. 2001.

[11] L. Li, A. Tulino, and S. Verdú, "Asymptotic eigenvalue moments for linear multiuser detection," *Commun. Inf. Syst.*, vol. 1, no. 3, pp. 273–304, Fall 2001.

[12] W. Chen, U. Mitra, and P. Schniter, "On the equivalence of three reduced rank linear estimators with applications to DS-CDMA," *IEEE Trans. Inf. Theory*, vol. 48, no. 9, pp. 2609–2614, Sep. 2002.

[13] P. Loubaton and W. Hachem, "Asymptotic analysis of reduced rank Wiener filters," in *Proc. IEEE Information Theory Workshop (ITW)*, Paris, France, Apr. 2003, pp. 328–331.

[14] L. Cottatellucci, R. R. Müller, and M. Debbah, "Asymptotic design and analysis of linear detectors for asynchronous CDMA systems," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Chicago, IL, Jun./Jul. 2004, p. 509.

[15] ——, "Efficient implementation of multiuser detectors for asynchronous CDMA," in *Proc. 42nd Allerton Conf. Communication, Control and Computing*, Monticello, IL, Sep./Oct. 2004, pp. 357–366.

[16] R. R. Müller and S. Verdú, "Spectral efficiency of low-complexity multiuser detection," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Sorrento, Italy, Jun. 2000, p. 439.

[17] L. Cottatellucci and R. R. Müller, "Performance assessment of polynomial detector in MIMO channels with measured data," in *Proc. IEE Technical Seminar on MIMO Communication Systems*, London, U.K., Dec. 2001, pp. 5/1–5/6.

[18] ——, "Asymptotic analysis of multistage receivers for multipath fading channels," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Yokohama, Japan, Jun./Jul. 2003, p. 509.

[19] W. Hachem, "Low complexity polynomial receivers for downlink CDMA," in *Proc. Asilomar Conf. Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 2002, pp. 1919–1923.

[20] ——, "Simple polynomial detectors for CDMA downlink transmissions on frequency-selective channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 1, pp. 164–172, Jan. 2004.

[21] L. Li, A. Tulino, and S. Verdú, "Design of reduced-rank MMSE multiuser detectors using random matrix methods," *IEEE Trans. Inf. Theory*, vol. 50, no. 6, pp. 986–1008, Jun. 2004.

[22] L. G. F. Trichard, J. S. Evans, and I. B. Collings, "Large system analysis of linear multistage parallel interference cancellation," *IEEE Trans. Commun.*, vol. 50, no. 11, pp. 1778–1786, Nov. 2002.

[23] D. N. C. Tse and S. Hanly, "Linear multiuser receivers: Effective interference, effective bandwidth and user capacity," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 641–657, Mar. 1999.

[24] S. Verdú and S. Shamai (Shitz), "Spectral efficiency of CDMA with random spreading," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 622–640, Mar. 1999.

[25] J. Evans and D. N. C. Tse, "Large system performance of linear multiuser receivers in multipath fading channels," *IEEE Trans. Inf. Theory*, vol. 46, no. 6, pp. 2059–2078, Sep. 2000.

[26] D. Segovia-Vargas, F. Iñigo, and M. Sierra-Pérez, "Generalized eigenspace beamformer based on CG-lanczos algorithm," *IEEE Trans. Antennas Propag.*, vol. 51, no. 8, pp. 2146–2154, Aug. 2003.

[27] S. Kay, *Fundamentals of Statistical Signal Processing, Estimation Theory*, ser. Prentice Hall Signal Processing. Englewood Cliffs, NJ: Prentice-Hall, 1993, vol. 1.

[28] U. Madhow and M. L. Honig, "MMSE interference suppression for direct-sequence spread-spectrum CDMA," *IEEE Trans. Commun.*, vol. 42, no. 12, pp. 3178–3188, Dec. 1994.

[29] Y. Q. Yin, "Limiting spectral distribution for a class of random matrices," *J. Multivariate Anal.*, vol. 20, pp. 50–68, 1986.

[30] J. W. Silverstein and Z. Bai, "On the empirical distribution of eigenvalues of a class of large dimensional random matrices," *J. Multivariate Anal.*, vol. 54, pp. 175–192, 1995.

[31] D. Jonsson, "Some limit theorems for the eigenvalues of a sample covariance matrix," *J. Multivariate Anal.*, vol. 12, pp. 1–38, 1982.

[32] J. W. Silverstein, "Strong convergence of the empirical distribution of eigenvalues of large dimensional random matrices," *J. Multivariate Anal.*, vol. 55, pp. 331–339, 1995.

[33] Z. D. Bai and J. W. Silverstein, "No eigenvalues outside the support of the limiting spectral distribution of large dimensional sample covariance matrices," *Ann. Probab.*, vol. 26, no. 1, pp. 316–345, 1998.

[34] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3rd ed. New York: McGraw-Hill, 1991.

[35] J. Zhang, E. K. Chong, and D. N. C. Tse, "Output MAI distributions of linear MMSE multiuser receivers in DS-CDMA systems," *IEEE Trans. Inf. Theory*, vol. 47, no. 3, pp. 1128–1144, Mar. 2001.

[36] W. Feller, *An Introduction to Probability Theory and Its Applications*, 3rd ed. New York: Wiley, 1968, vol. 1.

[37] S. Skiena, *Implementing Discrete Mathematics: Combinatorics and Graph Theory With Mathematica*. Reading, MA: Addison-Wesley, 1990.