Design and Analysis of Low-Complexity Interference Mitigation on Vector Channels

Ralf R. Müller and Sergio Verdú, *Fellow, IEEE*

***Abstract*—Linear multiuser detectors for vector channels with crosstalk are approximated by weighted matrix polynomials. The weight optimization problem is overcome using convergence results from random matrix theory. The results are also extended to receivers with subsequent successive decoding.**

In the case of subsequent successive decoding, a novel low-complexity implementation is found for the first-order approximation that is based on matched filter banks only and does not require matrix algebra. Spectral efficiency is obtained analytically and found to be fairly close to optimum.

The paper is focussed on multiuser detection for CDMA, but the results can be easily extended to communication via antenna arrays.

***Index Terms*—Antenna arrays, CDMA, interference cancellation, multiuser detection, random matrices, spectral efficiency.**

I. Introduction

CROSSTALK is a well-known drawback on vector channels in a variety of applications such as code-division multiple-access (CDMA) communication through antenna arrays, e.g., BLAST [1], xDSL, etc. All these systems can be commonly described by an output vector that is the product of a channel matrix multiplied by an input vector. The dimensions of those vectors are the parameters defining the size of the system, i.e., the number of users and the spreading gain in CDMA, the number of transmit and receive antennas in BLAST, the number of twisted pairs per cable for xDSL, etc.

The deleterious effect of the crosstalk can be reduced if the receiver takes into account the structure of the interfering signals. Those receivers do not only depend on the signal of interest but also on the correlation matrix of the channel [2]. Even suboptimum receivers that simply invert the channel or minimize the mean-squared error (MSE) require solving systems of linear equations that scale with the size of the system. For large systems, e.g., the frequency division duplex (FDD) mode in UMTS, some believe that multiuser detection, though it would improve performance significantly, is infeasible with today's technology.

Mathematical results on the convergence of the eigenvalues of large dimensional random matrices yield a completely different view of the complexity required for the mitigation of interference on linear vector channels. As the sizes of the channel matrices increase, the eigenvalues of the channels become more deterministic [3]. Since the quality of a communication link through a linear vector channel is sufficiently determined by the eigenvalues of the channels' covariance matrices, the problem of efficient detection should simplify for system with large size, as they are more structured. Indeed, this paper shows that efficient mitigation of crosstalk is feasible for large scale systems making use of recent results in random matrix theory. The complexity of our new algorithm is of the same order of magnitude as the single-user matched filter (MF).

The paper is organized as follows. Section II briefly reviews known linear methods for interference suppression from the multiuser detection literature. Section III derives our new schemes with the help of results from random matrix theory. Section IV analyzes their spectral efficiency, an information-theoretic performance measure. Section V presents fast implementations of our new algorithms and Section VI points out the conclusions.

II. Existing Low-Complexity Linear Detectors

Let

$$\boldsymbol{r} = \boldsymbol{S}^T(\boldsymbol{S}\boldsymbol{A}\boldsymbol{b} + \boldsymbol{n}) \tag{1}$$

be the notation of a vector-valued additive white Gaussian noise (AWGN) channel with $\boldsymbol{b}$ and $\boldsymbol{r}$ denoting the transmitted and received symbols, respectively. The AWGN is $\boldsymbol{n}$. In context of CDMA, the diagonal matrix $\boldsymbol{A}$ denotes the users' amplitudes and the $N \times K$ matrix $\boldsymbol{S}$ the K real-valued signature sequences of length N. Note that this channel is not restricted to CDMA. It models any linear, synchronous and memoryless AWGN channel with crosstalk, e.g., it directly applies to a single-user communication link with multiple transmitter and receiver antennas. Due to space limitations, however, we will restrict our considerations to CDMA. Those readers who are more interested in the antenna array applications are referred to [4].

It is convenient to define the cross-correlation matrix $\boldsymbol{R} \triangleq \boldsymbol{S}^T\boldsymbol{S}$. We assume that the diagonal entries of $\boldsymbol{R}$ equal unity without loss of generality. In addition, we assume $E\boldsymbol{b}\boldsymbol{b}^{\boldsymbol{H}} = \mathbf{I}, E\{\boldsymbol{n}\boldsymbol{n}^{\boldsymbol{H}}\} = N_0\mathbf{I}$, and

$$\frac{1}{K}\mathrm{tr}(\boldsymbol{A}^2) = 1 \tag{2}$$

Manuscript received December 15, 2000; revised April 20, 2001. The work of R. Müller was performed in part during his visit with the Department of Electrical Engineering, Princeton University, and supported by the German Academic Exchange Service (DAAD) under Grant 332 4 00 510. The work of S. Verdú was supported by the U.S. Army Research Office under Grant ARO DAAH04-96-1-0379. This work was presented in part at the IEEE International Symposium on Information Theory, Sorrento, Italy, June 2000 and the COST 262 Workshop "Multiuser Detection in Spread-Spectrum Communications," Reisensburg, Germany, January 2001.

R. R. Müller is with the Vienna Research Center for Telecommunications, Vienna, Austria (e-mail: Mueller@ftw.at).

S. Verdú is with the Dept. of Electrical Engineering, Princeton University, Princeton, NJ 08544 USA (e-mail: verdu@ee.princeton.edu).

Publisher Item Identifier S 0733-8716(01)07233-X.

with K and N_0 denoting the number of users and the noise power level of the AWGN channel, respectively.

The decorrelating[1] and the minimum mean-square error (MMSE) detector output vectors, for the channel defined above are given by [2]

$$\boldsymbol{d}_{\rm dec} = \boldsymbol{A}^{-1}\boldsymbol{R}^{-1}\boldsymbol{r} \tag{3}$$

$$\boldsymbol{d}_{\rm mmse} = (\boldsymbol{ARA} + N_0\mathbf{I})^{-1}\boldsymbol{Ar} \tag{4}$$

respectively.

In order to avoid matrix inversions, an approximate decorrelator was proposed in [5], [6]. It is based on the first-order Taylor approximation $(1+x)^{-1} = 1 - x + o(x), |x| < 1$. With the definition $\tilde{\boldsymbol{R}} \triangleq \boldsymbol{R} - \mathbf{I}$, it reads in the context of multiuser detection as

$$\boldsymbol{d}_{\rm dec,1} = \boldsymbol{A}^{-1}(\mathbf{I} - \tilde{\boldsymbol{R}})\boldsymbol{r} = \boldsymbol{A}^{-1}(2\,\mathbf{I} - \boldsymbol{R})\boldsymbol{r}. \tag{5}$$

Note that $\boldsymbol{A}^{-1}$ has negligible complexity as $\boldsymbol{A}$ is diagonal. From $(1+x)^{-1} = \sum_{\ell=0}^{\infty}(-x)^{\ell}, |x| < 1$, the first-order approximation (5) can be generalized to an arbitrary Lth order Taylor approximation by

$$\boldsymbol{d}_{{\rm dec},L} = \boldsymbol{A}^{-1}\sum_{\ell=0}^{L}(-\tilde{\boldsymbol{R}})^{\ell}\boldsymbol{r}. \tag{6}$$

This Lth order Taylor approximation is exactly identical to L-stage linear interference cancellation and to Jacobi's algorithm [7] for iterative matrix inversion.

The approximation converges to the exact solution, i.e., $\lim_{L\to\infty}\boldsymbol{d}_{{\rm dec},L} = \boldsymbol{d}_{\rm dec}$, iff all eigenvalues λ_i of the correlation matrix $\boldsymbol{R}$ are bounded by $0 < \lambda_i < 2\ \forall i$ [7]. From this point of view, it seems to be a reasonable and good approach. However, it is well known that finite order approximations that result from tail-cutting of infinite order approximations do not lead to the best fit among all approximations of the same order, in general. Thus, there should exist some weights $\boldsymbol{w}_L \triangleq [w_0, w_1, \ldots, w_L]^T$ such that the linear detector

$$\boldsymbol{d}_{{\rm dec},\boldsymbol{w}_L} = \boldsymbol{A}^{-1}\sum_{\ell=0}^{L}w_\ell\boldsymbol{R}^{\ell}\boldsymbol{r} \tag{7}$$

is a better approximation to the decorrelator than the Lth-order Taylor series. With the help of the Cayley–Hamilton theorem, the weighted polynomial detector can actually be shown to exactly implement the decorrelator and the MMSE detector for any $L \geq K - 1$ [8] if the weights may depend on the eigenvalues of the correlation matrix $\boldsymbol{R}$.

Less comprehensive generalizations of the approximate decorrelator can be found in [9] where also nonlinear detectors are addressed and [10] where methods from numerical mathematics [11] are applied to the problem of linear multiuser detection. The matrix polynomial in (7) can also be expressed as a finite product instead of a finite sum. This approach is followed in [12], [13].

The polynomial approximations to the decorrelator and the MMSE detector are only helpful in practice if the weights can be calculated more easily than applying the exact solution, i.e., performing matrix inversion. As the optimum weights depend on the eigenvalues of $\boldsymbol{R}$ which are not easy to calculate either, [8] suggested to calculate them in advance and store them in tables. This method, however, seems troublesome, as the eigenvalues depend on various changing parameters and it is not clear what advantage is gained over storing the inverse correlation matrix. Adaptive methods to track the weights are discussed in [8].

[1]Note that the operator $\boldsymbol{A}^{-1}$ is not necessary if the signal set has constant amplitude.

III. Weight Optimization

As the MMSE detector contains the decorrelator as a special case for $N_0 \to 0$, we consider the MMSE detector in the following. If needed, results for the decorrelator are obtained by letting $N_0 = 0$. Moreover, for ease of notation, we drop the indices $(\cdot)_{\rm dec}$ and $(\cdot)_{\rm mmse}$.

It seems reasonable to call a vector of weights $\boldsymbol{w}_L$ optimum if it maximizes the signal-to-interference-and-noise ratio (SINR) in the decision vector $\boldsymbol{d}_{\boldsymbol{w}_L}$ for fixed L. These optimum weights obviously depend on the correlation matrix $\boldsymbol{R}$ and the amplitudes $\boldsymbol{A}$, in general. Reference [8] derives a vector of weights that minimizes the MSE between the output signals of the exact MMSE detector and its polynomial approximation for users with equal amplitudes. This optimization criterion is equivalent to minimize the MSE under a constraint on the degree of the matrix polynomial. Its advantage is that it only depends on the eigenvalues of the correlation matrix $\boldsymbol{R}$ and the noise power density N_0, but not on the eigenvectors of $\boldsymbol{R}$. The total SINR, however, will turn out to also depend on the eigenvectors, in general.

For further considerations, it will be helpful to define the linear detector[2]

$$\boldsymbol{M}_{\boldsymbol{w}_L}\boldsymbol{A} \triangleq \sum_{\ell=0}^{L}w_\ell(\boldsymbol{ARA})^{\ell}\boldsymbol{A} \tag{8}$$

and the eigenvector-eigenvalue decomposition of the covariance matrix $\boldsymbol{ARA} \triangleq \boldsymbol{T\Lambda T}^T$. Then, the detector output vector can be written as

$$\boldsymbol{d}_{\boldsymbol{w}_L} = \boldsymbol{M}_{\boldsymbol{w}_L}\boldsymbol{A}r = \boldsymbol{M}_{\boldsymbol{w}_L}\boldsymbol{A}(\boldsymbol{R}\boldsymbol{A}\boldsymbol{b} + \boldsymbol{S}^T\boldsymbol{n}) \tag{9}$$

which, since the data symbols are assumed to be uncorrelated and the matrix $\boldsymbol{R}$ is symmetric, gives the total received power

$$P = E\left\{\boldsymbol{d}_{\boldsymbol{w}_L}^H\boldsymbol{d}_{\boldsymbol{w}_L}\right\} = {\rm tr}\left(\boldsymbol{M}_{\boldsymbol{w}_L}^2(\boldsymbol{ARA})^2\right) + N_0{\rm tr}\left(\boldsymbol{M}_{\boldsymbol{w}_L}^2\boldsymbol{ARA}\right). \tag{10}$$

It depends only on the noise power level and the eigenvalues of $\boldsymbol{ARA}$, but not on the eigenvectors of the covariance matrix $\boldsymbol{ARA}$ as it can be expressed as

$$P = {\rm tr}\left((\boldsymbol{\Lambda}^2 + N_0\boldsymbol{\Lambda})\left(\sum_{\ell=0}^{L}w_\ell\boldsymbol{\Lambda}^{\ell}\right)^2\right). \tag{11}$$

[2]We define the detector as $\boldsymbol{M}_{\boldsymbol{w}_L}\boldsymbol{A}$ in order to involve only symmetric matrices and simplify notation. Note also that the term $N_0\mathbf{I}$ in (4) has been absorbed into the constant term of the power series.

The total received power can be split up into useful signal power and the superposition of multiple-access interference (MAI) and noise power. As the subsequent processing of the users' signals is supposed to be based only on each users' individual signal, the useful signal power of user number k is the squared kth diagonal entry of the matrix $\boldsymbol{M}_{\boldsymbol{w}_L}\boldsymbol{ARA}$. Thus, the total useful signal power is given by

$$S = \operatorname{tr}\left(\operatorname{diag}^2\left(\boldsymbol{M}_{\boldsymbol{w}_L}\boldsymbol{ARA}\right)\right). \tag{12}$$

Let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_K$ denote the diagonal entries of $\boldsymbol{\Lambda}$ and

$$\xi_k \triangleq \sum_{\ell=0}^{L} w_\ell \lambda_k^{\ell+1}. \tag{13}$$

Then, (12) becomes

$$\begin{aligned} S &= \sum_{k=1}^{K}\left(\sum_{\mu=1}^{K}\boldsymbol{T}_{\mu k}^2 \xi_\mu\right)^2 \\ &= \sum_{\mu=1}^{K}\sum_{\nu=1}^{K}\xi_\mu\xi_\nu \sum_{k=1}^{K}\boldsymbol{T}_{\mu k}^2\boldsymbol{T}_{\nu k}^2. \end{aligned} \tag{14}$$

Obviously, not only does S depend on the eigenvalues, but, in general, also on the eigenvectors of the covariance matrix $\boldsymbol{ARA}$.

A. Optimum Weighting

In order to find the optimum weights, we write the SINR as

$$\mathrm{SINR} = \frac{S}{P-S} = \frac{1}{\frac{P}{S}-1}. \tag{15}$$

Maximizing the SINR and maximizing the ratio of useful signal power to total received power S/P are equivalent. As it will turn out to ease notation, we focus on the latter in the following.

We have

$$\frac{S}{P} = \frac{\sum_{\mu=1}^{K}\sum_{\nu=1}^{K}\lambda_\mu\lambda_\nu\sum_{\ell=1}^{L}w_\ell\lambda_\mu^\ell\sum_{l=1}^{L}w_l\lambda_\nu^l\sum_{k=1}^{K}\boldsymbol{T}_{\mu k}^2\boldsymbol{T}_{\nu k}^2}{\sum_{k=1}^{K}(\lambda_k^2+N_0\lambda_k)\sum_{\ell=1}^{L}w_\ell\lambda_k^\ell\sum_{l=1}^{L}w_l\lambda_k^l}. \tag{16}$$

Defining the matrix operator $\cdot^Q$ that squares its argument element-wise and the auxiliary matrix

$$\boldsymbol{C} \triangleq \begin{bmatrix} 1 & \lambda_1 & \lambda_1^2 & \cdots & \lambda_1^L \\ 1 & \lambda_2 & \lambda_2^2 & \cdots & \lambda_2^L \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_K & \lambda_K^2 & \cdots & \lambda_K^L \end{bmatrix} \tag{17}$$

for ease of notation, the signal-to-total-power ratio can be written as a generalized Rayleigh fraction

$$\frac{S}{P} = \frac{\boldsymbol{w}^T\boldsymbol{C}^T\boldsymbol{\Lambda}\boldsymbol{T}^{QT}\boldsymbol{T}^Q\boldsymbol{\Lambda}\boldsymbol{C}\boldsymbol{w}}{\boldsymbol{w}^T\boldsymbol{C}^T(\boldsymbol{\Lambda}^2+N_0\boldsymbol{\Lambda})\boldsymbol{C}\boldsymbol{w}}. \tag{18}$$

Moreover, let an arbitrary factorization of $\boldsymbol{C}^T(\boldsymbol{\Lambda}^2+N_0\boldsymbol{\Lambda})\boldsymbol{C}$ be defined by

$$\boldsymbol{F}^T\boldsymbol{F} \triangleq \boldsymbol{C}^T(\boldsymbol{\Lambda}^2+N_0\boldsymbol{\Lambda})\boldsymbol{C} \tag{19}$$

and

$$\tilde{\boldsymbol{w}} \triangleq \boldsymbol{F}\boldsymbol{w} \tag{20}$$

then the signal-to-total-power ratio may be expressed as an ordinary Rayleigh fraction

$$\frac{S}{P} = \frac{\tilde{\boldsymbol{w}}^T\boldsymbol{F}^{-T}\boldsymbol{C}^T\boldsymbol{\Lambda}\boldsymbol{T}^{QT}\boldsymbol{T}^Q\boldsymbol{\Lambda}\boldsymbol{C}\boldsymbol{F}^{-1}\tilde{\boldsymbol{w}}}{\tilde{\boldsymbol{w}}^T\tilde{\boldsymbol{w}}}. \tag{21}$$

The eigenvector corresponding to the largest eigenvalue of $\boldsymbol{F}^{-T}\boldsymbol{C}^T\boldsymbol{\Lambda}\boldsymbol{T}^{QT}\boldsymbol{T}^Q\boldsymbol{\Lambda}\boldsymbol{C}\boldsymbol{F}^{-1}$ maximizes the Rayleigh fraction (21) and therefore also the SINR. The optimum weight vector can then be found solving (20) for $\boldsymbol{w}$.

With this optimum choice for the weights, the maximum achievable SINR becomes

$$\max_{\boldsymbol{w}_L}\mathrm{SINR} = \frac{\mu_{\max}}{1-\mu_{\max}} \tag{22}$$

where $\mu_{\max}$ denotes the largest eigenvalue of $\boldsymbol{F}^{-T}\boldsymbol{C}^T\boldsymbol{\Lambda}\boldsymbol{T}^{QT}\boldsymbol{T}^Q\boldsymbol{\Lambda}\boldsymbol{C}\boldsymbol{F}^{-1}$.

The solution to the weight optimization problem does not seem to be easier than an inversion of the covariance matrix $\boldsymbol{ARA}$. In particular, it is disadvantageous that it even depends on the eigenvectors of $\boldsymbol{ARA}$ via the matrix $\boldsymbol{T}$. Thus, the solution is not helpful for an efficient approximate implementation of the MMSE detector.

The computational effort for the inversion of the correlation matrix is infeasible only if two conditions are fulfilled. The number of users is large and the spreading sequences are subject to random fluctuations. If the number of users were small, matrix inversion might be easily performed in real time. If the spreading sequences were known in advance to the receiver, there would be no need for real time matrix inversions. Fortunately, for random spreading sequences and a large number of users the weight optimization problem simplifies.

B. Asymptotic Weighting

As optimum weighting has been found to be infeasible, we focus on asymptotic weighting for random sequences in the following. In the asymptotic random sequence model, we assume the spreading gain and the number of users converge to infinity with a fixed finite ratio. In this case, some simplifications are possible which lead to results which are very helpful in practice [12]. Moreover, we assume the spreading sequences to be independent identically distributed (i.i.d.).

1) Equal Power Users: Consider first the case where all users are signaling with unit amplitude, i.e., $\boldsymbol{A} = \mathbf{I}$. Note that in this case, we have

$$\sum_{\mu=1}^{K}\sum_{\nu=1}^{K}\xi_\mu\xi_\nu = \operatorname{tr}^2\left(\boldsymbol{M}_{\boldsymbol{w}_L}\boldsymbol{R}\right) \tag{23}$$

which leads to the following asymptotic equivalence, which is shown in Appendix A:

Lemma 1: For i.i.d. real-valued random sequences with finite moments, the asymptotic equivalence

$$\begin{aligned} &\lim_{K\to\infty}\frac{1}{K}\operatorname{tr}\left(\operatorname{diag}^2\left(\boldsymbol{M}_{\boldsymbol{w}_L}\boldsymbol{R}\right)\right) \\ &\quad= \lim_{K\to\infty}\frac{1}{K^2}\operatorname{tr}^2\left(\boldsymbol{M}_{\boldsymbol{w}_L}\boldsymbol{R}\right) < \infty \end{aligned} \tag{24}$$

holds almost surely for all $\boldsymbol{M}_{\boldsymbol{w}_L}$ given by (8) and $\boldsymbol{A} = \mathbf{I}$.

Lemma 1 allows for an asymptotic expression of the signal-to-total-power ratio which becomes for real-valued spreading sequences[3]

$$\frac{S}{P} \to \lim_{K\to\infty} \frac{\mathrm{tr}^2\left(\boldsymbol{M}_{\boldsymbol{w}_L}\boldsymbol{R}\right)}{K\mathrm{tr}\left(\boldsymbol{M}^2_{\boldsymbol{w}_L}\boldsymbol{R}^2\right) + KN_0\mathrm{tr}\left(\boldsymbol{M}^2_{\boldsymbol{w}_L}\boldsymbol{R}\right)} \tag{25}$$

$$= \lim_{K\to\infty} \frac{\left(\sum_{k=1}^K \sum_{\ell=1}^L w_\ell \lambda_k^{\ell+1}\right)^2}{K\sum_{k=1}^K \left(\lambda_k^2 + N_0\lambda_k\right)\sum_{\ell=1}^L w_\ell\lambda_k^\ell \sum_{l=1}^L w_l\lambda_k^l}. \tag{26}$$

With the definition $\lambda \triangleq [\lambda_1, \lambda_2, \ldots, \lambda_K]^T$, we obtain

$$\frac{S}{P} \to \lim_{K\to\infty} \frac{\boldsymbol{w}^T\boldsymbol{C}^T\boldsymbol{\lambda}\boldsymbol{\lambda}^T\boldsymbol{C}w}{K\boldsymbol{w}^T\boldsymbol{C}^T(\boldsymbol{\Lambda}^2 + N_0\boldsymbol{\Lambda})\boldsymbol{C}\boldsymbol{w}}. \tag{27}$$

Comparing (27) and (18), we observe that $\boldsymbol{C}^T\boldsymbol{\lambda}\boldsymbol{\lambda}^T\boldsymbol{C}$ is an outer product of two identical vectors, i.e., its unique nonzero eigenvalue is $\boldsymbol{\lambda}^T\boldsymbol{C}\boldsymbol{C}^T\boldsymbol{\lambda}$, while $\boldsymbol{C}^T\boldsymbol{\Lambda}\boldsymbol{T}^{QT}\boldsymbol{T}^Q\boldsymbol{\Lambda}\boldsymbol{C}$ is full rank, in general. This fact will turn out to allow for an explicit solution to the asymptotically optimum weight vector.

Taking derivatives, with respect to the weight vector, gives (28), found at the bottom of the page. It can easily be checked that

$$\boldsymbol{w}_{\mathrm{asy},L} = (\boldsymbol{C}^T(\boldsymbol{\Lambda}^2 + N_0\boldsymbol{\Lambda})\boldsymbol{C})^{-1}\boldsymbol{C}^T\lambda \tag{29}$$

is a zero of (28). Moreover, it actually maximizes the SINR globally. Surprisingly, this asymptotically optimum weight vector is identical to that one found in [8] for minimization of the MSE between the output signals of the approximate and the exact MMSE detector in the *non*asymptotic case. With (29), the maximum asymptotic signal-to-total-power ratio becomes

$$\max_{\boldsymbol{w}} \frac{S}{P} \to \lim_{K\to\infty} \frac{1}{K}\lambda^T\boldsymbol{C}(\boldsymbol{C}^T(\boldsymbol{\Lambda}^2 + N_0\boldsymbol{\Lambda})\boldsymbol{C})^{-1}\boldsymbol{C}^T\lambda. \tag{30}$$

[3]The assumption of real-valued spreading sequences is a technical one. The authors believe that the results also extend to complex-valued spreading sequences.

The asymptotically optimum weights in (29) do not depend on the eigenvectors of the correlation matrix $\boldsymbol{R}$. Moreover, they are almost surely independent in total of the correlation matrix due to the following result from random matrix theory.

Theorem 1 [14][4]: Let $\boldsymbol{S}$ be an $N \times K$ matrix whose entries are i.i.d. random variables with zero mean and variance $1/N$, and let λ_k be the eigenvalues of $\boldsymbol{S}^H\boldsymbol{S}$. Moreover, let $K \to \infty$ and $N \to \infty$, but $0 < \beta \triangleq K/N < \infty$. Then, the moments of the eigenvalues

$$\frac{1}{K}\sum_{k=1}^K \lambda_k^m = \frac{1}{K}\mathrm{tr}(\boldsymbol{S}^H\boldsymbol{S})^m \tag{31}$$

converge almost surely to the nonrandom limits

$$\sum_{i=0}^{m-1} \binom{m}{i}\binom{m}{i+1}\frac{\beta^i}{m}. \tag{32}$$

Note that, in context of CDMA, N can be idenified with the number of chips per symbol (spreading factor). This nice convergence property of the eigenvalues of large dimensional random covariance matrices allows for an explicit analytic expression of the asymptotically optimum weight vector which depends only on the noise power density N_0 and the load β.

The moments of the eigenvalue density can be used to express the asymptotically optimum weight as well as the asymptotic signal-to-total-power ratio. For this purpose, it will be helpful to define (33)–(35), found at the bottom of the page. In terms of $\boldsymbol{m}_L(\beta)$ and $\boldsymbol{\Phi}_L(\beta, N_0)$ the asymptotic weights and the signal-to-total-power ratio admit the expressions

$$\boldsymbol{w}_{\mathrm{asy},L} = (\boldsymbol{\Phi}_L(\beta, N_0))^{-1}\boldsymbol{m}_L(\beta) \tag{36}$$

$$\frac{S}{P} \to \boldsymbol{m}_L(\beta)^T\left(\boldsymbol{\Phi}_L(\beta, N_0)\right)^{-1}\boldsymbol{m}_L(\beta). \tag{37}$$

[4]This theorem was shown in [14] for Gaussian random entries only. The present form of Theorem 1 which is not restricted to the distribution of the entries in $\boldsymbol{S}$ follows from more general results in [3], [15].

$$\frac{\partial}{\partial \boldsymbol{w}}\frac{S}{P} \to \lim_{K\to\infty} \frac{\boldsymbol{C}^T\lambda\lambda^T\boldsymbol{C}\boldsymbol{w}\boldsymbol{w}^T\boldsymbol{C}^T(\boldsymbol{\Lambda}^2 + N_0\boldsymbol{\Lambda})\boldsymbol{C}w - \boldsymbol{C}^T(\boldsymbol{\Lambda}^2 + N_0\boldsymbol{\Lambda})\boldsymbol{C}\boldsymbol{w}\boldsymbol{w}^T\boldsymbol{C}^T\lambda\lambda^T\boldsymbol{C}\boldsymbol{w}}{K\left(\boldsymbol{w}^T\boldsymbol{C}^T(\boldsymbol{\Lambda}^2 + N_0\boldsymbol{\Lambda})\boldsymbol{C}\boldsymbol{w}\right)^2} \tag{28}$$

$$\boldsymbol{m}_L(\beta) \triangleq \lim_{K\to\infty}\frac{1}{K}\boldsymbol{C}^T\lambda = \mathop{E}_{\boldsymbol{S}}\{[\lambda, \lambda^2, \ldots, \lambda^{L+1}]^T\} \tag{33}$$

$$\boldsymbol{\Phi}_L(\beta, N_0) \triangleq \lim_{K\to\infty}\frac{1}{K}\boldsymbol{C}^T(\boldsymbol{\Lambda}^2 + N_0\boldsymbol{\Lambda})\boldsymbol{C} \tag{34}$$

$$= \mathop{E}_{\boldsymbol{S}}\left\{\begin{bmatrix} \lambda^2 + N_0\lambda & \lambda^3 + N_0\lambda^2 & \cdots & \lambda^{L+2} + N_0\lambda^{L+1} \\ \lambda^3 + N_0\lambda^2 & \lambda^4 + N_0\lambda^3 & \cdots & \lambda^{L+3} + N_0\lambda^{L+2} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda^{L+2} + N_0\lambda^{L+1} & \lambda^{L+3} + N_0\lambda^{L+2} & \cdots & \lambda^{2L+2} + N_0\lambda^{2L+1} \end{bmatrix}\right\} \tag{35}$$

TABLE I
OPTIMUM WEIGHT VECTORS FOR $L = 1, 2, 3, 4$

$\boldsymbol{w}'_{\text{asy},1} = \begin{bmatrix} w_0 \\ w_1 \end{bmatrix}$	$w_0 = -N_0 w_1 + 2 + 2\beta$ $w_1 = -1$	
$\boldsymbol{w}'_{\text{asy},2} = \begin{bmatrix} w_0 \\ w_1 \\ w_2 \end{bmatrix}$	$w_0 = -N_0 w_1 + 3 + 4\beta + 3\beta^2$ $w_1 = -N_0 w_2 - 3 - 3\beta$ $w_2 = 1$	
$\boldsymbol{w}'_{\text{asy},3} = \begin{bmatrix} w_0 \\ w_1 \\ w_2 \\ w_3 \end{bmatrix}$	$w_0 = -N_0 w_1 + 4 + 6\beta + 6\beta^2 + 4\beta^3$ $w_1 = -N_0 w_2 - 6 - 9\beta - 6\beta^2$ $w_2 = -N_0 w_3 + 4 + 4\beta$ $w_3 = -1$	
$\boldsymbol{w}'_{\text{asy},4} = \begin{bmatrix} w_0 \\ w_1 \\ w_2 \\ w_3 \\ w_4 \end{bmatrix}$	$w_0 = -N_0 w_1 + 5 + 8\beta + 9\beta^2 + 8\beta^3 + 5\beta^4$ $w_1 = -N_0 w_2 - 10 - 18\beta - 18\beta^2 - 10\beta^3$ $w_2 = -N_0 w_3 + 10 + 16\beta + 10\beta^2$ $w_3 = -N_0 w_4 - 5 - 5\beta$ $w_4 = 1$	

Any scalar multiple of the asymptotically optimum weight vector is asymptotically optimum, too. This means it can be rescaled without loss of optimality by the determinant of $\boldsymbol{\Phi}_L(\beta, N_0)$ as well as an arbitrary scalar function of the load and written in terms of the adjoint of $\boldsymbol{\Phi}_L(\beta, N_0)$ as

$$\boldsymbol{w}'_{\text{asy},L} = \frac{\operatorname{adj}(\boldsymbol{\Phi}_L(\beta, N_0))}{\beta^{L(L+1)/2}} \boldsymbol{m}_L(\beta). \tag{38}$$

This scaling yields that its components are Lth order polynomials in β and N_0. If L is not too large, these polynomials can be calculated with commercial programs for symbolic algebra and are surprisingly simply[5] structured. For $L \leq 4$, the results are summarized in Table I. Real time calculation of the asymptotically optimum weight vectors is surprisingly simple, as Table I indicates. Therefore, the implementation problems of polynomial approximations to decorrelating and MMSE detectors have been overcome for large scale systems with random spreading and equal power users. Approximations for the decorrelator can be obtained from approximations to the MMSE detector letting $N_0 = 0$. As Table I indicates, this results only in negligible reduction of complexity.

2) Unequal Power Users: The results for equal power users do not generalize straightforwardly to users with unbalanced powers. Note that Lemma 1 does not hold in that case. Alternatively, the MSE can serve as performance criterion instead of SINR. This approach was followed in [12] independently from this work. It avoids the need for Lemma 1, but leads to complicated expressions for the asymptotically optimum weights which involve the first $2L$ moments of the power distribution. In the following, we analyze a simpler method.

[5]Note that, at first glance, one would expect to obtain polynomials in the load and noise power density of order $L(L+1)$ instead of order L which would imply the need to calculate $L^2(L+1)^2$ products per component of the weight vector.

Let the polynomial detector ignore the unbalanced powers of the interfering users. This means that we address the detector

$$\hat{\boldsymbol{M}}_{\boldsymbol{w}_L} \triangleq \sum_{i=0}^{L} w_i \boldsymbol{R}^i. \tag{39}$$

For its analysis, the following lemma which is proven in Appendix B will be crucial:

Lemma 2: Let $\boldsymbol{R}$ be a $K \times K$ covariance matrix of random sequences with i.i.d. real-valued chips with finite moments and $\boldsymbol{A}$ be a $K \times K$ nonrandom diagonal matrix fulfilling the normalization (2). Then

$$\lim_{K\to\infty} \frac{1}{K} \operatorname{tr}(\boldsymbol{R}^k \boldsymbol{A}^2) = \lim_{K\to\infty} \frac{1}{K} \operatorname{tr}(\boldsymbol{R}^k) \tag{40}$$

holds almost surely for any positive integer k.

The total received power and the useful signal power for the detector defined in (39) are given by

$$\hat{P} = \operatorname{tr}\left(\boldsymbol{R}\hat{\boldsymbol{M}}^2_{\boldsymbol{w}_L}\boldsymbol{R}\boldsymbol{A}^2\right) + N_0 \operatorname{tr}\left(\hat{\boldsymbol{M}}^2_{\boldsymbol{w}_L}\boldsymbol{R}\right) \tag{41}$$

$$\hat{S} = \operatorname{tr}\left(\operatorname{diag}^2\left(\hat{\boldsymbol{M}}_{\boldsymbol{w}_L}\boldsymbol{R}\boldsymbol{A}\right)\right) \tag{42}$$

respectively. Note the discrepancies between (41) and (42), and (10), (12), due to the lack of scaling with $\boldsymbol{A}$ in (39). With the help of Lemma 2, this yields for real-valued spreading sequences to

$$\frac{\hat{S}}{\hat{P}} \to \lim_{K\to\infty} \frac{\operatorname{tr}\left(\operatorname{diag}^2\left(\hat{\boldsymbol{M}}_{\boldsymbol{w}_L}\boldsymbol{R}\boldsymbol{A}\right)\right)}{\operatorname{tr}\left(\boldsymbol{R}\hat{\boldsymbol{M}}^2_{\boldsymbol{w}_L}\boldsymbol{R}\boldsymbol{A}^2\right) + N_0 \operatorname{tr}\left(\hat{\boldsymbol{M}}^2_{\boldsymbol{w}_L}\boldsymbol{R}\right)} \tag{43}$$

$$= \lim_{K\to\infty} \frac{\operatorname{tr}\left(\operatorname{diag}^2\left(\hat{\boldsymbol{M}}_{\boldsymbol{w}_L}\boldsymbol{R}\right)\right)}{\operatorname{tr}\left(\hat{\boldsymbol{M}}^2_{\boldsymbol{w}_L}\boldsymbol{R}^2\right) + N_0 \operatorname{tr}\left(\hat{\boldsymbol{M}}^2_{\boldsymbol{w}_L}\boldsymbol{R}\right)} \tag{44}$$

$$= \lim_{K\to\infty} \frac{S}{P} \tag{45}$$

the same asymptotic signal-to-total-power ratio as found for equal powers in (25). Obviously, this equivalence also yields equivalence of the SINRs in both cases.

The SINR for unequal powers is a quantity averaged over all users. Large SINRs of some users may compensate for low ones of other users. Therefore, it is not obvious whether it is indeed a sensible measure of performance in any case. The distribution of the users' SINRs gives more insight. It is proven in the following that

$$\lim_{K\to\infty} \mathrm{SINR}_k = A_k^2 \lim_{K\to\infty} \mathrm{SINR} \tag{46}$$

where A_k denotes the kth user's amplitude. Note that due to symmetry the total interference and noise power affecting user k is independent of the index k. In contrast, the useful signal power is proportional to the transmitted signal power of user k, as the transmission is linear and both channel and receiver do not depend on the other user's powers. Therefore, the calculated SINR is given by

$$\lim_{K\to\infty} \mathrm{SINR} = \lim_{K\to\infty} \frac{1}{K} \sum_{k=1}^{K} \frac{\mathrm{SINR}_k}{A_k^2}. \tag{47}$$

As the SINR is asymptotically independent of the choice of $\boldsymbol{A}$, this yields (46).

Optimizing the weights for equal powers, but applying them in the unequal power case yields the same average SINR, as applying them in the equal power case which they are designed for. To be more precise, we have Theorem 2.

Theorem 2: Let a polynomial multiuser detector be defined as in (39), i be the index of the user of interest, and the power distribution among the users be normalized to $\sum_k P_k = K$ with $P_i = 1$. Then, the average SINR for random spreading is almost surely independent of the power distribution, as $K \to \infty$.

Note that Theorem 2 holds for any weights. Note also that the polynomial multiuser detector defined in (39) does not involve the users' powers. Therefore, it cannot be the best Lth order approximation to the MMSE detector for any weights, except in the equal power case. This implies that a polynomial multiuser detector as defined in (8) with weights optimized with respect to SINR performs at least as well as that one defined in (39) with equality holding only in the equal power case. This implies that regarding the detector defined in (8) there is a saddle point of the SINR as a function of the weights and the power distribution. To be more precise, we have Theorem 3.

Theorem 3: Let a polynomial multiuser detector be defined as in (8), i be the index of the user of interest, and the power distribution among the users be normalized to $\sum_k P_k = K$ with $P_i = 1$. Moreover, let the spreading sequences be assigned randomly and $K \to \infty$. Then, there is a power distribution and a weight assignment such that for any different weight assignment the average SINR decreases and for any different power distribution the average SINR is unaffected. This power distribution is the equal power case.

The interference caused by unbalanced interferers has turned out to be less harmful than in the equal power case in the asymptotic limit with random spreading. Thus, interferers with equal powers are the worst case scenario. Therefore, our main focus is on the equal power case throughout the rest of this work and all results are also lower bounds on performances for unbalanced interferers.

IV. Analysis of Efficiency

In the previous section, linear multiuser detectors that do not require matrix inversions have been proposed and optimized. Certainly, there is strong interest to find out how well this approximations perform in comparison to the benchmarks set by the decorrelator and the MMSE detector.

The traditional measures of comparison for uncoded performance of multiuser detectors are asymptotic efficiency and near–far resistance [2]. However, they do not help in our particular case. All discussed linear approximations will turn out to exhibit zero near–far resistance. Nevertheless, they could obtain a large performance gain over the conventional MF detector and exhibit excellent performance in moderate near–far scenarios.

If random spreading is considered, a more general means of comparison than near–far resistance, called spectral efficiency, can be applied [16], [17]. It is discussed for decorrelating and/or MMSE detectors with and without successive cancellation in [18]–[22] and [18], [21], respectively. Another framework for analysis of linear multiuser receivers with random spreading is the concept of effective interference [23]. Although it applies to linear multiuser detectors, it is not clear how to generalize it to nonlinear receivers including those involving successive cancellation. Therefore, we will analyze the performance of the approximation to decorrelating and MMSE detectors with and without successive cancellation in terms of spectral efficiency. Hereby, we restrict our considerations to the case of equal powers due to the reasons outlined in Section III–B2.

TABLE II
NORMALIZED DETERMINANTS $\Delta_L(\beta, N_0)$ FOR $L = 0, 1, 2, 3, 4$

Δ_0	$=$	$N_0 + 1 + \beta$
Δ_1	$=$	$N_0^2 + (2+2\beta)N_0 + 1 + \beta + \beta^2$
Δ_2	$=$	$N_0^3 + (3+3\beta)N_0^2 + (3+4\beta+3\beta^2)N_0 + 1 + \beta + \beta^2 + \beta^3$
Δ_3	$=$	$N_0^4 + (4+4\beta)N_0^3 + (6+9\beta+6\beta^2)N_0^2 + (4+6\beta+6\beta^2+4\beta^3)N_0 +$ $+1 + \beta + \beta^2 + \beta^3 + \beta^4$
Δ_4	$=$	$N_0^5 + (5+5\beta)N_0^4 + (10+16\beta+10\beta^2)N_0^3 + (10+18\beta+18\beta^2+10\beta^3)N_0^2 +$ $+(5+8\beta+9\beta^2+8\beta^3+5\beta^4)N_0 + 1 + \beta + \beta^2 + \beta^3 + \beta^4 + \beta^5$

In the nonasymptotic case, an exact analysis is difficult. Results have been reported only for the decorrelator with spherical random sequences [18], [21]. For general sets of sequences, however, [24] indicates that the results should be similar. Due to these reasons and forced by the additional analytical trouble arising in the case of polynomial detectors for finite-length sequences by their dependency on the eigenvectors of the correlation matrix, we restrict to the asymptotic case in the following. This has the additional advantage that no averaging over realizations of the random sequences is required, as Theorem 1 ensures convergence of the eigenvalue distribution in probability.

A basic tool in order to calculate spectral efficiency, is an analytic formula expressing the SINR in terms of the load and noise power density. In the present case, it certainly depends also on the weight vector. Because of (15) it is sufficient to find the signal-to-total power ratio. It becomes

$$\frac{S}{P} \to \frac{\left(\boldsymbol{w}_L^T \boldsymbol{m}_L(\beta)\right)^2}{\boldsymbol{w}_L^T \boldsymbol{\Phi}_L(\beta, N_0) \boldsymbol{w}_L} \tag{48}$$

with (27), (33), and (34) and can be further simplified with (36) to obtain

$$\frac{S}{P} \to \boldsymbol{m}_L^T(\beta) \boldsymbol{w}_{\mathrm{asy},L} = \frac{\boldsymbol{m}_L^T(\beta) \boldsymbol{w}'_{\mathrm{asy},L}}{\Delta_L(\beta, N_0)} \tag{49}$$

where

$$\Delta_L(\beta, N_0) \triangleq \frac{\det \boldsymbol{\Phi}_L(\beta, N_0)}{\beta^{L(L+1)/2}}. \tag{50}$$

The normalized determinants $\Delta_L(\beta, N_0)$ can be calculated in the same way as the asymptotic weight vectors. Some of them are listed in Table II.

With these preliminaries

$$\mathrm{SINR}_L \to \frac{\boldsymbol{m}_L^T(\beta) \boldsymbol{w}'_{\mathrm{asy},L}}{\Delta_L(\beta, N_0) - \boldsymbol{m}_L^T(\beta) \boldsymbol{w}'_{\mathrm{asy},L}} \tag{51}$$

allows to give explicit expressions for the optimum SINRs, e.g., (52)–(56), found at the bottom of the page. Obviously, the zeroth-order approximation is equivalent to the conventional MF. Our first-order approximation is better than the approximate decorrelator (AD) analyzed in ([2], p. 281), where the weights were based on a first-order Taylor series and not optimized with respect to the maximum achievable SINR, see (5).

Note that due to Lemma 1, the asymptotic SINR depends only on the moments of the eigenvalue density which converges in probability. This means we have the following theorem.

Theorem 4: Let $\boldsymbol{A} = \mathbf{I}, K, N \to \infty$, but $0 < \beta \triangleq (K/N) < \infty$ and the random components of $\boldsymbol{S}$ be independent with finite variance. Then, the SINR at the output $\boldsymbol{d} = \boldsymbol{Mr}$ of any linear detector described by a matrix $\boldsymbol{M} = \sum_{\ell=0}^{L} w_\ell(\beta, N_0)\boldsymbol{R}^\ell$ converges almost surely to a nonrandom quantity for arbitrary weight functions $w_\ell(\beta, N_0), 0 \le \ell \le L$, and arbitrary order L.

In the case of vanishing noise, an explicit expression for the SINR is possible for arbitrary order L:

$$\lim_{N_0 \to 0} \max_{\boldsymbol{w}_L} \mathrm{SINR}_L \to \frac{\beta^{-L-1} - 1}{1 - \beta}. \tag{57}$$

It was found in [25] in related but different context and also holds for polynomial approximations to linear multiuser detection. Note that the SINR is finite even when no AWGN is present. However, the asymptotic SINR grows exponentially with the order of the approximation for $\beta < 1$.

A. Linear Receivers

The previous results on the SINR with random spreading can be plugged into the definitions of power and bandwidth efficiency reading

$$\frac{N_0}{E_b} \triangleq N_0 C = N_0 \log_2(1 + \mathrm{SINR}(N_0, \beta)) \tag{58}$$

$$\Gamma \triangleq \beta C = \beta \log_2(1 + \mathrm{SINR}(N_0, \beta)) \tag{59}$$

respectively [18], [19], with C denoting the channel capacity of an individual user's channel.

A parametric description of the functional relationship between power and bandwidth efficiency is given by (58) and (59). While, for the MMSE detector with random spreading, an explicit expression can be found [18], this is not possible for approximations to linear receivers, in general. In the following extreme case, however, we find

$$\lim_{\frac{E_b}{N_0} \to \infty} \Gamma_L = \beta \log_2 \left(\frac{\beta - \beta^{-L-1}}{\beta - 1} \right) \tag{60}$$

which is illustrated in Fig. 1. It can be observed that around $\beta \approx 0.4$ spectral efficiency increases by about half a bit/s/Hz with each additional stage of the approximation. Thus

$$\max \Gamma_L \approx \frac{L}{2} + 1 \tag{61}$$

is an accurate rule of thumb.

Fig. 1 points to a remarkable property of polynomial approximations to MMSE receivers: While for the MMSE receiver, the optimum load converges to $\beta = 1$ as $E_b/N_0 \to \infty$, the optimum load of its polynomial approximation shows a different behavior. In Appendix C, it is shown that

$$\lim_{L \to \infty} \operatorname*{argmax}_{\beta} \lim_{N_0 \to 0} \max_{\boldsymbol{w}_L} \beta \log_2(1 + \mathrm{SINR}_L) = \frac{1}{\mathrm{e}} \tag{62}$$

i.e., the asymptotically optimum load converges to $\exp(-1)$. On the other hand, the approximation becomes exact as the order approaches the number of users, i.e., $L \to K$. As $K \to \infty$, this means

$$\lim_{N_0 \to 0} \operatorname*{argmax}_{\beta} \lim_{L \to \infty} \max_{\boldsymbol{w}_L} \beta \log_2(1 + \mathrm{SINR}_L) = 1. \tag{63}$$

Obviously, the behavior differs if the limits are exchanged. However, it does not matter from a practical point of view because the limits $N_0 \to 0, L \to \infty$ are far away from the operating conditions of real-world communication systems.

For less extreme situations, i.e., SNRs E_b/N_0 that range within intervals that are of practical interest, the polynomial approximation receiver shows promising performance. This is observed from Fig. 2. It shows spectral efficiency as a function of the load for a fixed SNR that represents a typical setting in practice. The approximate MMSE receiver is found to

$$\max_{\boldsymbol{w}_0} \mathrm{SINR}_0 \to \frac{1}{\beta + N_0} = \mathrm{SINR}_{\mathrm{MF}} \tag{52}$$

$$\max_{\boldsymbol{w}_1} \mathrm{SINR}_1 \to \frac{1 + \beta + N_0}{\beta^2 + N_0(1 + 2\beta) + N_0^2} \overset{\beta > 0}{>} \frac{1 - 2\beta + \beta^2}{\beta^2 + \beta^3 + N_0(1 - \beta + \beta^2)} \leftarrow \mathrm{SINR}_{\mathrm{AD}} \tag{53}$$

$$\max_{\boldsymbol{w}_2} \mathrm{SINR}_2 \to \frac{1 + \beta + \beta^2 + N_0(2 + 2\beta) + N_0^2}{\beta^3 + N_0(1 + 2\beta + 3\beta^2) + N_0^2(2 + 3\beta) + N_0^3} \tag{54}$$

$$\max_{\boldsymbol{w}_3} \mathrm{SINR}_3 \to \frac{1 + \beta + \beta^2 + \beta^3 + N_0(3 + 4\beta + 3\beta^2) + N_0^2(3 + 3\beta) + N_0^3}{\beta^4 + N_0(1 + 2\beta + 3\beta^2 + 4\beta^3) + N_0^2(3 + 6\beta + 6\beta^2) + N_0^3(3 + 4\beta) + N_0^4} \tag{55}$$

$$\max_{\boldsymbol{w}_4} \mathrm{SINR}_4 \to \frac{1 + \beta + \beta^2 + \beta^3 + \beta^4 + N_0(4 + 6\beta + 6\beta^2 + 4\beta^3) +}{\beta^5 + N_0(1 + 2\beta + 3\beta^2 + 4\beta^3 + 5\beta^4) + N_0^2(4 + 9\beta + 12\beta^2 + 10\beta^3) +} \cdots \frac{+N_0^2(6 + 9\beta + 6\beta^2) + N_0^3(4 + 4\beta) + N_0^4}{+N_0^3(6 + 12\beta + 10\beta^2) + N_0^4(4 + 5\beta) + N_0^5}. \tag{56}$$

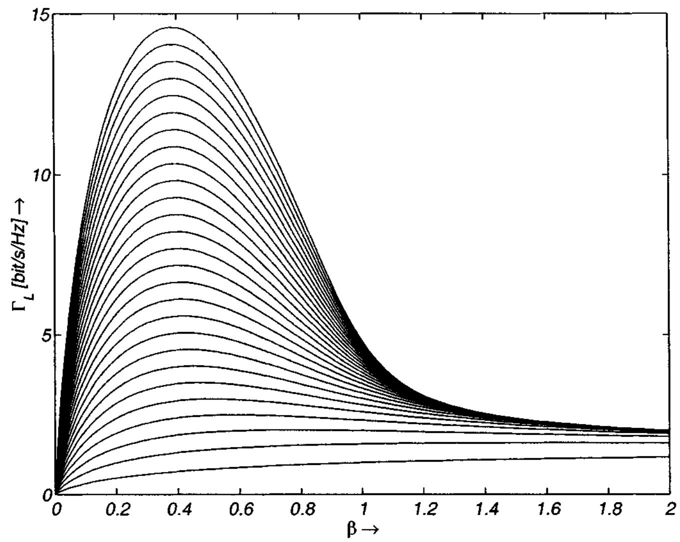

Fig. 1. Spectral efficiency versus load for approximations of zeroth to 26th order as $E_b/N_0 \to \infty$.

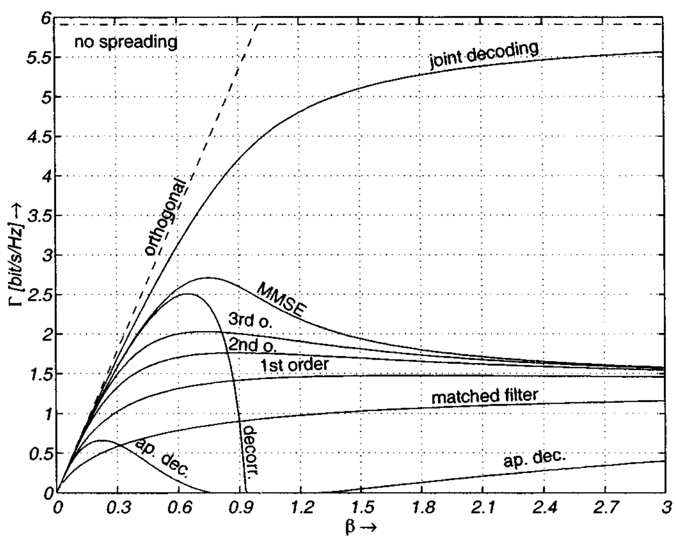

Fig. 2. Spectral efficiency versus system load for several linear multiuser receivers and fixed $10\log_{10}(E_b/N_0) = 10$ dB in comparison to limits for orthogonal systems and joint decoding reported in [19].

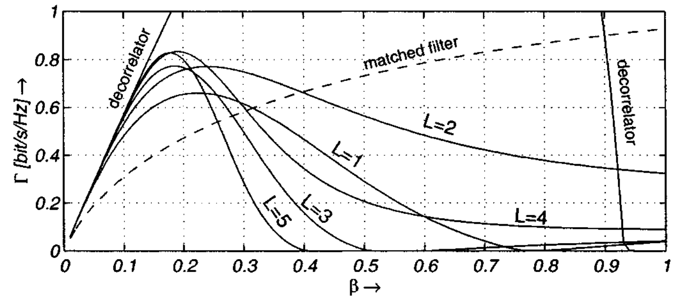

Fig. 3. Spectral efficiency versus system load for the first five Taylor approximations to the decorrelator and fixed $10\log_{10}(E_\mathrm{b}/N_0) = 10$ dB.

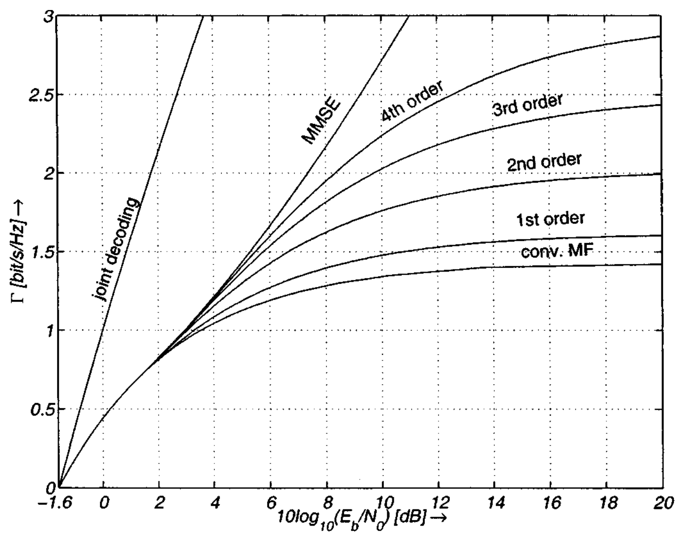

Fig. 4. Spectral efficiency versus power efficiency for several linear multiuser receivers with optimized loads.

approximate its exact counterpart rather fast, while the approximate decorrelator specified in (5) does not show promising performance for $\beta > 0.3$. Additionally, spectral efficiency is hardly affected by fluctuations of the load.

All these properties result from proper weighting of the powers of the correlation matrix. Without sophisticated weighting, spectral efficiency may drop far behind that of the conventional MF even for multistage approximations. This is illustrated in Fig. 3 for the Taylor approximation defined in (6). For all orders shown, the range of β where reasonable performance is achieved is very limited, as the spectral radius of $\boldsymbol{S}$ exceeds the convergence interval of the Taylor approximation if $\beta > (N_0\sqrt{2}-1)^2 \approx 0.17$ [10]. Remarkably, the Taylor approximations with even order are better with respect to a performance-complexity tradeoff.

The tradeoff between power and bandwidth efficiency is well illustrated in Fig. 4. There the load is optimized for each individual receiver. The figure shows that the number of approximation orders L that is reasonable to implement heavily depends on the SNR on the channel. As already observed in Figs. 1, Fig. 4 also confirms the rule of thumb found in (61). Only for $L = 0$, i.e., the conventional MF, it is not accurate.

B. Successive Cancellation Receivers

Postdecoding successive cancellation (SC) receivers with linear MMSE predetection were shown to incur no loss in comparison with joint multiuser decoding [26]. Their general structure is depicted in Fig. 5. First, we assume that all users signal at identical powers. In this case, spectral efficiency can be easily obtained by integration of the respective spectral efficiencies without successive cancellation [18], [19], [21], i.e.,

$$\Gamma_{\mathrm{SC}}(\beta) = \int_0^\beta \Gamma_{\mathrm{linear}}(\beta')\frac{d\beta'}{\beta'}. \tag{64}$$

For Lth-order polynomial approximations to the linear multiuser MMSE predetector as front end to successive decoding, this means

$$\Gamma_L = \int_0^\beta \log_2(1 + \mathrm{SINR}_L(N_0, \beta'))\,d\beta'. \tag{65}$$

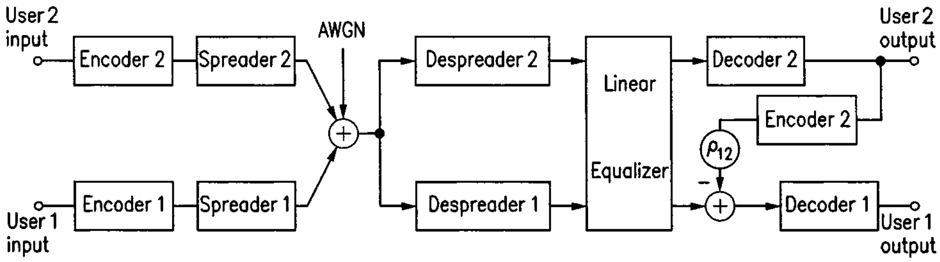

Fig. 5. Postdecoding successive cancellation for two users.

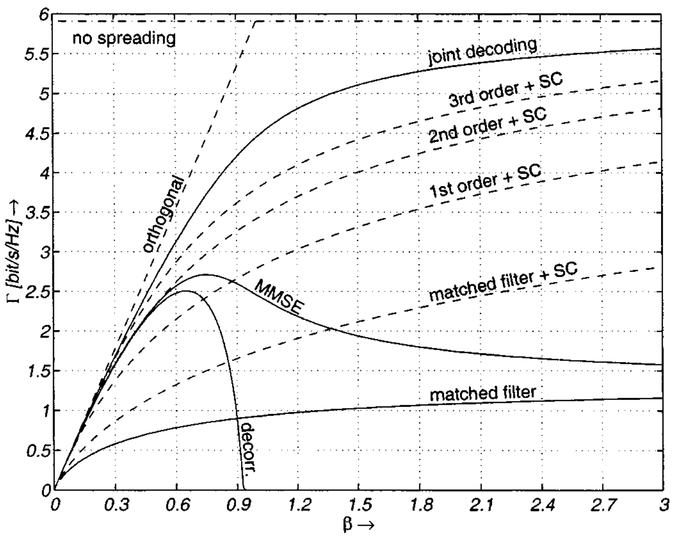

Fig. 6. Spectral efficiency versus system load for several multiuser receivers and fixed $10\log_{10}(E_b/N_0) = 10$ dB.

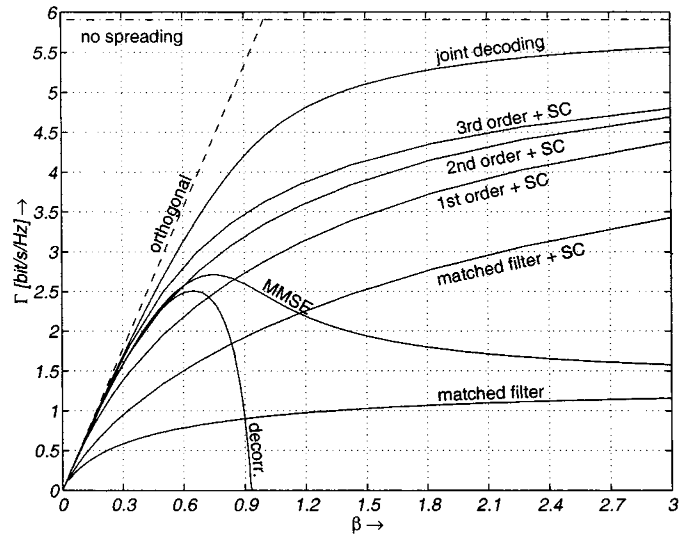

Fig. 7. Spectral efficiency versus system load for equal rate user and fixed $10\log_{10}(E_b/N_0) = 10$ dB.

Spectral efficiency is plotted in Fig. 6 versus the load for a fixed SNR and compared to some standard methods of multiuser detection. The linear MMSE detector can only keep step with the second-order approximation for loads up to about 0.5. For larger loads, it falls behind even the first-order approximation. The interference cancellation without linear predetection also keeps far behind all $L \geq 2$ order approximations. The third-order approximation is already close to the theoretical limit.

The previous considerations are based on successive decoding with equal powers for all users. This implies an imbalanced rate assignment to the users. Although the rate imbalances can be compensated by methods like time-sharing or rate-splitting, such methods involve additional complexity. Alternatively, the powers can be adjusted in such a way that a target rate distribution is achieved. This approach is analyzed in the following at the example of equal rates for all users. Following the approach in Section III–B2 to ignore the power imbalances of the weight design, the common channel capacity of all users becomes

$$C = \log_2\left(1 + \frac{(k-1)A_k^2}{\sum_{i=1}^{k-1} A_i^2} \times \mathrm{SINR}\left((k-1)N_0 \Big/ \sum_{i=1}^{k-1} A_i^2,\ \frac{k}{N}\right)\right) \quad \forall k \quad (66)$$

since the noise variance is scaled by the average interference power. This uniquely defines a power allocation $A_1^2, \ldots, A_K^2$ which can be computed recursively from (66). With knowledge of the power allocation, it is a routine procedure to calculate spectral efficiency, see e.g., [21] for details. The results are depicted in Fig. 7. It is interesting to observe that the lower order approximations improve in comparison to their respective counterparts in the equal power case while the higher order approximations lose performance. This phenomena is not surprising. The considered detector converges from the MF to the MMSE detector, as the order grows from zero to infinity. The results in [21] show that for the MF the equal rate case is advantageous while the MMSE detector performs better with equal powers.

V. Fast Implementations

The overarching goal of all this work is to simplify the implementation effort of multiuser detection. This means the proposed Lth order approximations to the MMSE detector would be meaningless if there were no simple ways to implement them.

A fast structure to implement a linear Lth order detector for fixed weights has been found in [27]. There, the computations are accelerated compared to (7) using a generalization of Horner's scheme [28]:

$$\boldsymbol{d}_{\boldsymbol{w}_L} = (((\cdots((w_L \boldsymbol{S}^T\boldsymbol{S} + w_{L-1}\boldsymbol{I})\boldsymbol{S}^T\boldsymbol{S}\cdots)\cdots + w_2\boldsymbol{I})\boldsymbol{S}^T\boldsymbol{S} + w_1\boldsymbol{I})\boldsymbol{S}^T\boldsymbol{S} + w_0\boldsymbol{I})\boldsymbol{r}. \quad (67)$$

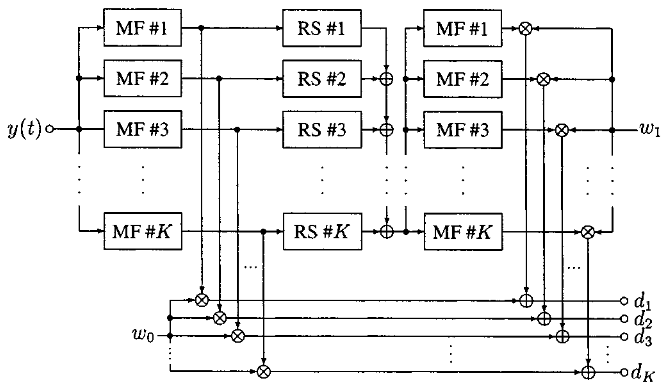

Fig. 8. Block structure for first-order approximate linear MMSE detection [27].

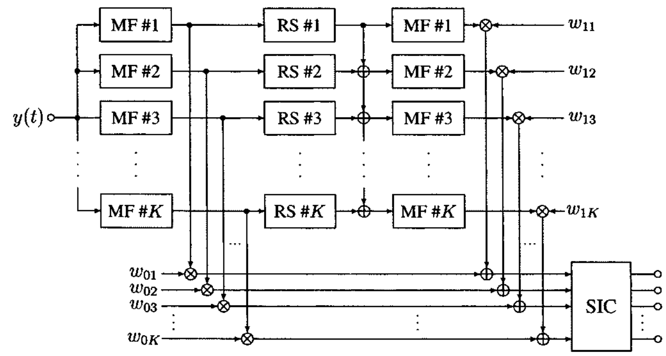

Fig. 9. Proposed structure for first-order approximate MMSE detection with subsequent SIC.

This means that multiplications between matrices are avoided by subsequent respreading and matched filtering. In effect, only multiplications between matrices and vectors are required. For $L = 1$, the receiver is illustrated in Fig. 8 in its equivalent complex baseband representation. Obviously, calculation of the cross-correlation matrix $\boldsymbol{R}$ is avoided by matched-filtering (MF) and respreading (RS) of the receiver input signal $y(t)$. For Lth order approximations with $L > 1$, the unit consisting of RS, summation, and matched-filtering is simply to be repeated L times. Note that in case of multipath channels, the initial MFs have to be replaced by rake receivers [29], and the respreaders and the subsequent MFs should be based on the effective spreading sequences estimated by the rake receivers.

Measuring in units of the complexity of the conventional MF, the complexity for a linear Lth order approximation is given by

$$Z_{\text{lin}} = 3L + 1. \tag{68}$$

Hereby, the factor three takes into account L times RS, summation, and matched-filtering in addition to the initial matched-filtering. Note that the complexity of the exact MMSE detector implemented by Gaussian elimination exhibits complexity

$$Z_{\text{MMSE}} = \text{const.} \cdot K + 1 \tag{69}$$

which can be substantially higher than (68) for large K. Moreover, the polynomial approximation can be computed by K to N processors in parallel while Gaussian elimination is a sequential algorithm.

The problem remaining unsolved in [27], i.e., to find a method for adjusting the weights in real time, has been solved in this work for large number of users K making use of random matrix theory. Note that for a small number of users the benefits of the polynomial approximation are limited, as in that case the complexities Z_{lin} and Z_{MMSE} do not differ significantly.

Although (67) provides an easy implementation for approximate linear multiuser receivers, it does not straightforwardly extend to receivers involving successive cancellation. Moreover, the invention of an algorithm whose complexity measured in multiples of the MF complexity being less than proportional to the number of users seems to be a difficult task for general L. In the special case $L = 1$, an algorithm with complexity

$$Z_{\text{SC},1} = 4 \tag{70}$$

is given next.

Algorithm 1: Let $\boldsymbol{a}_0 = \boldsymbol{0}$ and $(\cdot)_i$ denote the ith component and the ith column of a vector and a matrix, respectively. Let $\boldsymbol{w}_L(K, N)$ denote the (optimized) weight vector for a system with K users, and spreading factor N. Then, the detector's output vector $\boldsymbol{d}$ is calculated recursively by

$$\boldsymbol{a}_k = \boldsymbol{a}_{k-1} + (\boldsymbol{r})_k(\boldsymbol{S})_k \tag{71}$$
$$(\boldsymbol{d})_k = \boldsymbol{a}_k^H(\boldsymbol{S})_k(\boldsymbol{w}_1(k, N))_1 + (\boldsymbol{r})_k(\boldsymbol{w}_1(k, N))_0. \tag{72}$$

Proof: As the users are canceled in the inverse order of their indices and cancellation is assumed to be perfect, we find

$$(\boldsymbol{d})_k = (\boldsymbol{w}_1(k, N))_0(\boldsymbol{r})_k + (\boldsymbol{w}_1(k, N))_1 \times \sum_{\kappa=1}^{k}(\boldsymbol{S})_\kappa^T(\boldsymbol{S})_k(\boldsymbol{r})_\kappa. \tag{73}$$

The users with indices larger than k do not matter as their influence is cancelled. It remains to be shown that

$$\boldsymbol{a}_k = \sum_{\kappa=1}^{k}(\boldsymbol{S})_\kappa(\boldsymbol{r})_\kappa. \tag{74}$$

This can be directly seen from (71). □

This algorithm is illustrated in Fig. 9. The main modification to get from the symmetric structure in the linear case, see Fig. 8, to the asymmetric structure for subsequent successive inference cancellation (SIC), is to replace the overall summation between RS and second matched-filtering by a cumulative summation. In addition, interference cancellation provides virtually different system loads and channel loads among the users. Therefore, the weights significantly vary from user to user. They can be calculated according to the formulas in Section III which hold exactly for large systems and are a very good approximation for small sized systems.

The structure in Fig. 9 is more general than it needs to be. Obviously, either w_{0i} or w_{1i} can be arbitrarily assigned any nonzero value for any i. Note also that the path consisting of the second MF #1 and the weighting with w_{11} is redundant. The following choice of weights seems to yield the least complex structure for CDMA:

$$w_{11} = 0 \tag{75}$$
$$w_{1i} = -1 \quad \forall\, 2 \le i \le K \tag{76}$$
$$w_{01} = \frac{2K}{N} + 1 + N_0 \tag{77}$$
$$w_{0i} = \frac{2(K-i+1)}{N} + 2 + \frac{(i-1)N_0}{\sum_{\mu=1}^{i-1} A_\mu^2} \qquad \forall\, 2 \le i \le K. \tag{78}$$

In the equal power case, (78) can be further simplified as $A_i = 1$ for all users. The corresponding results for antenna arrays are straightforward. Note, however, that both system and channel load have to be adapted to the virtually present interferers.

In order to find a simple structure for higher order approximation, it is not sufficient to repeat the unit consisting of RS, cumulative summation, and matched filtering as it is possible for symmetric receivers: In the linear case, the signal of user k after Lth order weighted filtering can be represented as

$$w_0(\boldsymbol{r})_k + (\boldsymbol{S})_k^T \sum_{\ell=1}^{L} w_\ell \left(\sum_{i=1}^{K} (\boldsymbol{S})_i (\boldsymbol{S})_i^T \right)^{\ell-1} \boldsymbol{S} r. \tag{79}$$

This representation yields the efficient implementation of (67). In case of subsequent successive cancellation the users with larger indices than k are virtually not present as they get canceled (almost) error-free later on. Thus, the signal corresponding to (79) becomes

$$w_0(\boldsymbol{r})_k + (\boldsymbol{S})_k^T \sum_{\ell=1}^{L} w_\ell \left(\sum_{i=1}^{k} (\boldsymbol{S})_i (\boldsymbol{S})_i^T \right)^{\ell-1} \times \sum_{\kappa=1}^{k} (\boldsymbol{S})_\kappa (\boldsymbol{r})_\kappa. \tag{80}$$

In this case, the $(\ell - 1)$st power of the correlation matrix is dependent on the user index k and needs to be calculated for each user, separately. Hereby, the complexity becomes proportional to the number of users. Only for $L = 1$, the term $(\sum_{i=1}^{k} (\boldsymbol{S})_i (\boldsymbol{S})_i^T)^{\ell-1}$ vanishes for all ℓ, which allows for Algorithm 1.

VI. Conclusion

MMSE multiuser detection approximated by weighted polynomial matrix-filtering, in particular in conjunction with subsequent successive cancellation, has been show to offer an excellent tradeoff between performance and complexity. Hereby, the misconception that increasing spectral efficiency by multiuser detection involves significant additional complexity has been debunked.

The underlying principles that led to our results are more general and based on fundamental properties of large covariance matrices. Therefore, the low-complexity detectors found in this paper are also applicable to other vector channels.

Appendix

A. Proof of Lemma 1

Note from (23) and (12) to (14), it is sufficient for proving the lemma to show that

$$\lim_{K\to\infty} \frac{1}{K} \sum_{\mu=1}^{K} \sum_{\nu=1}^{K} \xi_\mu \xi_\nu \sum_{k=1}^{K} \boldsymbol{T}_{\mu k}^2 \boldsymbol{T}_{\nu k}^2 = \lim_{K\to\infty} \frac{1}{K^2} \sum_{\mu=1}^{K} \sum_{\nu=1}^{K} \xi_\mu \xi_\nu < \infty \tag{81}$$

holds almost surely. The double sum on the right hand side is obviously finite, as it sums the eigenvalues of a bounded matrix. Thus, only the equivalence of the triple sum on the left-hand side and the double sum on the right hand side remains to be shown.

Note that the rows (and columns) of $\boldsymbol{T}$ are eigenvectors of $\boldsymbol{R}$. Thus

$$\sum_{k=1}^{K} \boldsymbol{T}_{\mu k}^2 = 1 \quad \forall \mu \tag{82}$$

since the eigenvectors are normalized to unit norm. However, according to ([30], Theorem 2), we have a stronger result in Lemma 3.

Lemma 3: Let all entries of $\boldsymbol{S}$ be real and i.i.d. random variables with all moments finite. Then, for any $t \in [0, 1]$

$$\lim_{K\to\infty} \sum_{k=1}^{[Kt]} \boldsymbol{T}_{\ell k}^2 = t \tag{83}$$

holds almost surely for any ℓ with $[\cdot]$ denoting the greatest integer being smaller than the argument.

Note that Lemma 3 does not only show asymptotic unit norm for the rows (and columns) of $\boldsymbol{T}$, but also that the cumulative sum grows proportional to the number of terms being summed. This implies Lemma 4.

Lemma 4: Let all entries of $\boldsymbol{S}$ be real i.i.d. random variables with all moments finite and the series $\langle \xi\mu/k \rangle$ be absolutely summable. Then

$$\lim_{K\to\infty} \sum_{k=1}^{K} \xi_k \boldsymbol{T}_{\ell k}^2 = \lim_{K\to\infty} \frac{1}{K} \sum_{k=1}^{K} \xi_k \tag{84}$$

holds almost surely for any ℓ.

Proof: Since $\langle \xi\mu/k \rangle$ is summable, any ξ_k can be decomposed into a sum of $K-k+1$ terms $\tilde{\xi}_\ell, 1 \le \ell \le K-k+1$ such that

$$\xi_k = \sum_{\ell=1}^{K-k+1} \tilde{\xi}_\ell \quad \forall k. \tag{85}$$

Let $\boldsymbol{\Xi}$ be a triangular matrix such that

$$\boldsymbol{\Xi}_{\ell k} = \begin{cases} \tilde{\xi}_\ell, & \text{for } k \le \ell \\ 0, & \text{elsewhere} \end{cases}. \tag{86}$$

Then, we get almost surely

$$\begin{aligned}
&\lim_{K\to\infty}\sum_{k=1}^{K}\xi_k \boldsymbol{T}_{\ell k}^2 \\
&= \lim_{K\to\infty}\sum_{k=1}^{K}\sum_{\ell=1}^{K-k+1}\tilde{\xi}_\ell \boldsymbol{T}_{\ell k}^2 = \lim_{K\to\infty}\sum_{k=1}^{K}\sum_{\ell=1}^{K}\boldsymbol{\Xi}_{\ell k}\boldsymbol{T}_{\ell k}^2 \\
&= \lim_{K\to\infty}\sum_{\ell=1}^{K}\sum_{k=1}^{K}\boldsymbol{\Xi}_{\ell k}\boldsymbol{T}_{\ell k}^2 \qquad (87)\\
&= \lim_{K\to\infty}\sum_{\ell=1}^{K}\tilde{\xi}_\ell\sum_{k=1}^{\ell}\boldsymbol{T}_{\ell k}^2 \qquad (88)\\
&= \lim_{K\to\infty}\sum_{\ell=1}^{K}\tilde{\xi}_\ell\sum_{k=1}^{\ell}\frac{1}{K} = \lim_{K\to\infty}\frac{1}{K}\sum_{\ell=1}^{K}\sum_{k=1}^{K}\boldsymbol{\Xi}_{\ell k} \qquad (89)\\
&= \lim_{K\to\infty}\frac{1}{K}\sum_{k=1}^{K}\sum_{\ell=1}^{K-k+1}\tilde{\xi}_\ell \\
&= \lim_{K\to\infty}\frac{1}{K}\sum_{k=1}^{K}\xi_k \qquad (90)
\end{aligned}$$

where Lemma 3 is used to obtain (89) from (88) while all other conclusions make use of (85) and (86). □

With Lemma 4, we get almost surely

$$\begin{aligned}
&\lim_{K\to\infty}\frac{1}{K}\sum_{\mu=1}^{K}\sum_{\nu=1}^{K}\xi_\mu\xi_\nu\sum_{k=1}^{K}\boldsymbol{T}_{\mu k}^2\boldsymbol{T}_{\nu k}^2 \\
&= \lim_{K\to\infty}\frac{1}{K}\sum_{k=1}^{K}\sum_{\mu=1}^{K}\xi_\mu\boldsymbol{T}_{\mu k}^2\sum_{\nu=1}^{K}\xi_\nu\boldsymbol{T}_{\nu k}^2 \qquad (91)\\
&= \lim_{K\to\infty}\frac{1}{K^2}\sum_{k=1}^{K}\sum_{\mu=1}^{K}\xi_\mu\boldsymbol{T}_{\mu k}^2\sum_{\nu=1}^{K}\xi_\nu \qquad (92)\\
&= \lim_{K\to\infty}\frac{1}{K^3}\sum_{k=1}^{K}\sum_{\mu=1}^{K}\xi_\mu\sum_{\nu=1}^{K}\xi_\nu \qquad (93)\\
&= \lim_{K\to\infty}\frac{1}{K^2}\sum_{\mu=1}^{K}\xi_\mu\sum_{\nu=1}^{K}\xi_\nu. \qquad (94)
\end{aligned}$$

since $\xi\mu/k$ is absolutely summable. The latter statement follows from (13) and the bounded spectral radius of asymptotically large random matrices. □

B. Proof of Lemma 2

The eigenvalue decomposition $\boldsymbol{R} = \boldsymbol{T\Lambda T}^T$ gives

$$\begin{aligned}
\frac{1}{K}\mathrm{tr}(\boldsymbol{R}^k\boldsymbol{A}^2) &= \frac{1}{K}\mathrm{tr}(\boldsymbol{T\Lambda}^k\boldsymbol{T}^T\boldsymbol{A}^2) = \frac{1}{K}\mathrm{tr}(\boldsymbol{\Lambda}^k\boldsymbol{T}^T\boldsymbol{A}^2\boldsymbol{T}) \\
&= \frac{1}{K}\sum_{i=1}^{K}\lambda_i\sum_{\mu=1}^{K}A_\mu^2\boldsymbol{T}_{\mu i}^2. \qquad (95)
\end{aligned}$$

Lemma 4 (see Appendix A) yields

$$\sum_{\mu=1}^{K}A_\mu^2\boldsymbol{T}_{\mu i}^2 \to \frac{1}{K}\sum_{\mu=1}^{K}A_\mu^2 \qquad (96)$$

almost surely as $K\to\infty$. Thus, we find the almost sure convergence of

$$\begin{aligned}
\frac{1}{K}\mathrm{tr}(\boldsymbol{R}^k\boldsymbol{A}^2) &\to \left(\frac{1}{K^2}\right)\sum_{i=1}^{K}\lambda_i\sum_{\mu=1}^{K}A_\mu^2 \\
&= \frac{1}{K}\mathrm{tr}(\boldsymbol{R}^k)\cdot\frac{1}{K}\mathrm{tr}(\boldsymbol{A}^2). \qquad (97)
\end{aligned}$$

The normalization (2) gives the result to be proven. □

C. Derivation of (62)

Here, the validity of (62) is verified:

$$\begin{aligned}
&\lim_{L\to\infty}\operatorname*{argmax}_{\beta}\lim_{N_0\to 0}\max_{\boldsymbol{w}_L}\beta\log_2(1+\mathrm{SINR}_L) \\
&= \lim_{L\to\infty}\operatorname*{argmax}_{\beta}\beta\log_2\left(\frac{\beta^{-L-1}-\beta}{1-\beta}\right) \qquad (98)\\
&= \operatorname*{argmax}_{\beta}\lim_{L\to\infty}\frac{\beta}{L}\ln\left(\frac{\beta^{-L-1}-\beta}{1-\beta}\right) \\
&= \operatorname*{argmax}_{\beta}\lim_{L\to\infty}\frac{\beta}{L}\ln(\beta^{-L-1}-\beta) \qquad (99)\\
&= \operatorname*{argmax}_{\beta}\lim_{L\to\infty}\frac{\beta\ln\beta}{\beta^{L+2}-1} = \operatorname*{argmax}_{\beta} -\beta\ln\beta = \frac{1}{e}. \qquad (100)
\end{aligned}$$

ACKNOWLEDGMENT

We would like to thank E. Biglieri, A. Tulino, S. Moshavi, K. Kopsa, and L. Cottatellucci for critical comments and a careful reading of the manuscript.

REFERENCES

[1] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Tech. J.*, vol. 1, no. 2, pp. 41–59, 1996.
[2] S. Verdú, *Multiuser Detection*. New York: Cambridge Univ. Press, 1998.
[3] V. A. Marcenko and L. A. Pastur, "Distribution of eigenvalues for some sets of random matrices," *Math. USSR-Sb*, no. 1, pp. 457–483, 1967.
[4] R. R. Müller, "Polynomial expansion equalizers for communication via large antenna arrays," in European Personal Mobile Communications Conf., Vienna, Austria, Feb. 2001.
[5] R. Kohno, M. Hatori, and H. Imai, "Cancellation techniques of co-channel interference in asynchronous spread spectrum multiple access systems," *Electron. Commun.*, vol. 66-A, no. 5, pp. 20–29, 1983.
[6] S. Verdú, "Multiuser detection," in *Advances in Statistical Signal Processing: Signal Detection*, H. V. Poor and J. B. Thomas, Eds. Greenwich, CT: JAI Press, 1993, pp. 369–410.
[7] R. S. Varga, *Matrix Iterative Analysis*. Englewood Cliffs, NJ: Prentice-Hall, 1962.
[8] S. Moshavi, E. G. Kanterakis, and D. L. Schilling, "Multistage linear receivers for DS-CDMA systems," *Int. J. Wireless Inform. Networks*, vol. 3, no. 1, pp. 1–17, Jan. 1996.
[9] D. Divsalar, M. K. Simon, and D. Raphaeli, "Improved parallel interference cancellation for CDMA," *IEEE Trans. Commun.*, vol. 46, pp. 258–268, Feb. 1998.
[10] A. Grant and C. Schlegel, "Iterative implementations for linear multiuser detectors," IEEE Trans. Commun., submitted for publication.
[11] O. Axelsson, *Iterative Solution Methods*. New York: Cambridge Univ. Press, 1994.
[12] D. Guo, L. K. Rasmussen, and T. J. Lim, "Linear parallel interference cancellation in long-code CDMA multiuser detection," *IEEE J. Select. Areas Commun.*, vol. 17, pp. 2074–2081, Dec. 1999.
[13] D. Guo, L. K. Rasmussen, S. Sun, and T. J. Lim, "A matrix-algebraic approach to linear parallel interference cancellation in CDMA," *IEEE Trans. Commun.*, vol. 48, pp. 152–161, Jan. 2000.

[14] D. Jonsson, “Some limit theorems for the eigenvalues of a sample covariance matrix,” *J. Multivariate Anal.*, vol. 12, pp. 1–38, 1982.
[15] J. W. Silverstein and Z. D. Bai, “On the empirical distribution of eigenvalues of a class of large dimensional random matrices,” *J. Multivariate Anal.*, vol. 54, pp. 175–192, 1995.
[16] A. J. Viterbi, “When not to spread spectrum—A sequel,” *IEEE Commun. Mag.*, vol. 23, pp. 12–17, Apr. 1985.
[17] ——, *CDMA*. Reading, MA: Addison-Wesley, 1995.
[18] R. R. Müller, *Power and Bandwidth Efficiency of Multiuser Systems with Random Spreading*. Aachen, Germany: Shaker-Verlag, 1999.
[19] S. Verdú and S. S. Shitz, “Spectral efficiency of CDMA with random spreading,” *IEEE Trans. Inform. Theory*, vol. 45, pp. 622–640, Mar. 1999.
[20] P. Schramm and R. R. Müller, “Spectral efficiency of CDMA systems with linear MMSE interference suppression,” *IEEE Trans. Commun.*, vol. 47, pp. 722–731, May 1999.
[21] R. R. Müller, “Multiuser receivers for randomly spread signals: Fundamental limits with and without decision-feedback,” *IEEE Trans. Inform. Theory*, vol. 47, pp. 268–283, Jan. 2001.
[22] S. S. Shitz and S. Verdú, “The impact of frequency-flat fading on the spectral efficiency of CDMA,” *IEEE Trans. Inform. Theory*, vol. 47, pp. 1302–1327, May 2001.
[23] D. Tse and S. Hanly, “Linear multiuser receivers: Effective interference, effective bandwidth and user capacity,” *IEEE Trans. Inform. Theory*, vol. 45, pp. 641–657, Mar. 1999.
[24] D. Tse and O. Zeitouni, “Linear multiuser receivers in random environments,” *IEEE Trans. Inform. Theory*, vol. 46, pp. 171–188, Jan. 2000.
[25] M. Honig and W. Xiao, “Performance of reduced-rank linear interference suppression for DS-CDMA,” *IEEE Trans. Inform. Theory*, vol. 47, pp. 1928–1946, July 2001.
[26] M. K. Varanasi and T. Guess, “Optimum decision feedback multiuser equalization with successive decoding achieves the total capacity of the Gaussian multiple-access channel,” in *Proc. Asilomar Conf. Signals, Systems, Computers*, Monterey, CA, Nov. 1997, pp. 1405–1409.
[27] E. Kanterakis and S. Moshavi, “Multistage Linear Receiver for DS-CDMA Systems,”, May 1998.
[28] I. N. Bronstein and K. A. Semendjajew, *Taschenbuch der Mathematik*, 22nd ed., B. G. Teubner, Ed. Frankfurt, Germany, 1985.
[29] W. C. Jakes Jr, Ed., *Microwave Mobile Communications*. New York: Wiley, 1974.
[30] J. W. Silverstein, “Eigenvalues and eigenvectors of large dimensional sample covariance matrices,” in *Random Matrices and Their Applications*, J. E. Cohen, H. Kesten, and C. M. Newman, Eds. Providence, RI: American Mathematical Soc., 1986, pp. 153–159.

Ralf R. Müller was born in Schwabach, Germany, 1970. He received the Dipl.-Ing. and Dr.-Ing. degree with distinction from University of Erlangen-Nuremberg, Germany, in January 1996 and March 1999, respectively.

He was a Visiting Research Fellow with the Department of Electrical Engineering, Princeton University, NJ, from April 1999 to April 2000. Since May 2000, he has been a Senior Researcher with Vienna Research Center for Telecommunications (FTW), Austria. He received a postdoctoral fellowship from German Academic Exchange Service (DAAD) in 1999. He has also served as an expert for both the former German Ministry for Post and Telecommunications as well as the European Patent Court.

Dr. Müller's dissertation, “Power and Bandwidth Efficiency of Multiuser Systems with Random Spreading,” received awards from both the Mannesmann Foundation for Research in Mobile Communications and the Information Technology Society (ITG) of the Association of German Electrical Engineers (VDE) in May 2000 and September 2000, respectively.

Sergio Verdú (S'80–M'84–SM'88–F'93) is Professor with the Department of Electrical Engineering and the Program in Applied and Computational Mathematics, Princeton University, NJ. He is active in the fields of information theory and multiuser communications. In the 1980s, he pioneered the technology of multiuser detection which exploits the structure of multiaccess interference in order to increase the capacity of multiuser communication systems. His textbook, *Multiuser Detection*, was published in 1998 by Cambridge University Press.

Dr. Verdú is a recipient of several paper awards: the IEEE Donald Fink Paper Award, a Golden Jubilee Paper Award, the 1998 Outstanding Paper Award from the IEEE Information Theory Society, and the 2000 Paper Award from the Telecommunications Advancement Foundation of Japan. He also received a Millennium Medal from the IEEE and the 2000 Frederick E. Terman Award from the American Society for Engineering Education. He served as Associate Editor for Shannon Theory of the IEEE TRANSACTIONS ON INFORMATION THEORY. He served on the Board of Governors of the Information Theory Society in 1989–1999, and was President of the Society in 1997. He was Co-Chairman of the Program Commitee of the 1998 IEEE International Symposium on Information Theory, and Co-Chairman of the 2000 IEEE International Symposium on Information Theory. He served as Editor of the Special 1948–1998 Commemorative Issue of the IEEE TRANSACTIONS ON INFORMATION THEORY, reprinted by IEEE Press in 1999 as *Information Theory: Fifty Years of Discovery*.