Design and Analysis of Low–Complexity Interference Mitigation for CDMA

Ralf R. Müller[*]
Forschungszentrum
Telekommunikation Wien

Sergio Verdú[†]
Department of EE
Princeton University

Abstract

Linear MMSE detectors for CDMA with crosstalk are approximated by weighted matrix polynomials. The weight optimization problem is overcome using convergence results from random matrix theory. Imbalanced powers among users are shown to have minor impact using properties of free random variables. The performance in terms of spectral efficiency is analyzed analytically and found near optimum.

1 Introduction

Multiuser interference is a well–known drawback in code–division multiple–access (CDMA) communications. The deleterious effect of interference 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 [1]. Even suboptimum receivers that simply invert the channel or minimize the mean–squared error require solving systems of linear equations that scale with the number of users. For large systems, e.g. the FDD (frequency division duplex) mode in UMTS (universal mobile telecommunication system), 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 of multiuser detection. As the size of the spreading matrix increases, the eigenvalues of the correlation matrix become more and more deterministic [2]. This paper shows that efficient mitigation of multiuser interference 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.

The paper is organized as follows: Section 2 briefly reviews known linear methods for interference suppression from the multiuser detection literature. Section 3 derives our new schemes with the help of results from random matrix theory. Section 4 analyzes their spectral efficiency, an information–theoretic performance measure, and Section 5 points out the conclusions.

[*]The work of R. Müller was performed during his visit at the Dept. of EE at 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.

2 Existing Low–Complexity Linear Detectors

Let

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

be the vector notation of a CDMA channel in presence of additive white Gaussian noise (AWGN) with $\boldsymbol{b}$ and $\boldsymbol{r}$ denoting the transmitted and received symbols, respectively. The AWGN is $\boldsymbol{n}$. The diagonal matrix $\boldsymbol{A}$ denotes the users' amplitudes and the matrix $\boldsymbol{S}$ the signature sequences.

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

$$\frac{1}{K}\,\mathsf{E}\,\mathrm{tr}\left(\boldsymbol{A}^2\right) = 1 \tag{2}$$

with N_0 denoting the noise power level of the AWGN channel.

The decorrelating and the MMSE detector output vectors, for the channel defined above are given by [1] $\boldsymbol{d}_{\mathrm{dec}} = \boldsymbol{A}^{-1}\boldsymbol{R}^{-1}\boldsymbol{r}$ and $\boldsymbol{d}_{\mathrm{mmse}} = (\boldsymbol{A}\boldsymbol{R}\boldsymbol{A} + N_0\mathbf{I})^{-1}\boldsymbol{A}\boldsymbol{r}$, respectively. In order to avoid matrix inversions, an approximate decorrelator was proposed in [3, 4]. It is based on the first–order Taylor approximation $(1+x)^{-1} = 1 - x + \mathrm{o}(x), |x| < 1$. With the definition $\tilde{\boldsymbol{R}} \triangleq \boldsymbol{R} - \mathbf{I}$, it reads in the context of multiuser detection as

$$\boldsymbol{d}_{\mathrm{dec},1} = \boldsymbol{A}^{-1}\left(\mathbf{I} - \tilde{\boldsymbol{R}}\right)\boldsymbol{r} = \boldsymbol{A}^{-1}(2\,\mathbf{I} - \boldsymbol{R})\boldsymbol{r}. \tag{3}$$

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 (3) can be generalized to an arbitrary L^{th} order Taylor approximation by

$$\boldsymbol{d}_{\mathrm{dec},L} = \boldsymbol{A}^{-1}\sum_{\ell=0}^{L}\left(-\tilde{\boldsymbol{R}}\right)^{\ell}\boldsymbol{r}. \tag{4}$$

This L^{th} order Taylor approximation is exactly identical to L–stage linear interference cancellation and to Jacobi's algorithm [5] for iterative matrix inversion.

The approximation converges to the exact solution, i.e. $\lim_{L\to\infty}\boldsymbol{d}_{\mathrm{dec},L} = \boldsymbol{d}_{\mathrm{dec}}$, iff all eigenvalues λ_i of the correlation matrix $\boldsymbol{R}$ are bounded by $0 < \lambda_i < 2\,\forall i$ [5]. 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 $\boldsymbol{w}_L \triangleq [w_0, w_1, \ldots, w_L]^{\mathrm{T}}$ such that the linear detector

$$\boldsymbol{d}_{\mathrm{dec},\boldsymbol{w}_L} = \boldsymbol{A}^{-1}\sum_{\ell=0}^{L} w_{\ell}\boldsymbol{R}^{\ell}\boldsymbol{r} \tag{5}$$

is a better approximation to the decorrelator than the L^{th}–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 $L \geq K - 1$, [6], 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 [7] where also nonlinear detectors are addressed and [8] where methods from numerical mathematics [9] are applied to the problem of linear multiuser detection. The matrix polynomial in (5) can also be expressed as a finite product instead of a finite sum. This approach is followed in [10, 11].

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, [6] suggested to calculate them in advance and store them in tables. This method, however, is hardly feasible, as the eigenvalues depend on various changing parameters and it is not clear what advantage is gained over storing the inverse correlation matrix. For known weights, however, very efficient algorithms are known to implement polynomial multiuser detectors with and without subsequent successive cancellation [12, 13] that do not even need to compute the correlation matrix of the spreading sequences.

3 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 $\cdot_{\text{dec}}$ and $\cdot_{\text{mmse}}$.

We call a vector of weights $\boldsymbol{w}_L$ optimum if it maximizes the *signal–to–interference–and–noise ratio (SIR)* in the decision vector $\boldsymbol{d}_{\boldsymbol{w}_L}$ for fixed L. These optimum weights depend on the correlation matrix $\boldsymbol{R}$ and the amplitudes $\boldsymbol{A}$, in general. Reference [6] derives a vector of weights that minimizes the mean–squared error between the output signals of the exact MMSE detector and its polynomial approximation for users with equal amplitudes. 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 $\boldsymbol{R}$. As we will show later, it does not maximize the SIR, in general, even for equal power users.

For further considerations it is helpful to define the linear detector[1]

$$\boldsymbol{M}_{\boldsymbol{w}_L}\boldsymbol{A} \triangleq \sum_{\ell=0}^{L} w_\ell \left(\boldsymbol{ARA}\right)^\ell \boldsymbol{A} \tag{6}$$

and the eigenvector–eigenvalue decomposition of the covariance matrix $\boldsymbol{ARA} \triangleq \boldsymbol{T\Lambda T}^{\mathrm{H}}$. Then, the detector output vector can be written as $\boldsymbol{d}_{\boldsymbol{w}_L} = \boldsymbol{M}_{\boldsymbol{w}_L}\boldsymbol{A}\left(\boldsymbol{RAb} + \boldsymbol{S}^{\mathrm{H}}\boldsymbol{n}\right)$ which gives immediately the total received power

$$P = \mathsf{E}\, \boldsymbol{d}_{\boldsymbol{w}_L}^{\mathrm{H}} \boldsymbol{d}_{\boldsymbol{w}_L} = \mathrm{tr}\left(\boldsymbol{M}_{\boldsymbol{w}_L}^2 (\boldsymbol{ARA})^2\right) + N_0 \mathrm{tr}\left(\boldsymbol{M}_{\boldsymbol{w}_L}^2 \boldsymbol{ARA}\right). \tag{7}$$

It depends only on the noise power level and the eigenvalues of $\boldsymbol{ARA}$, but not on the eigenvectors of the covariance matrix $\boldsymbol{ARA}$.

The total received power can be split up into useful signal power and the superposition of multi–access interference 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 k^{th} diagonal entry of the matrix $\boldsymbol{M}_{\boldsymbol{w}_L}\boldsymbol{ARA}$. Thus, the total useful signal power is given by

$$S = \mathrm{tr}\left(\mathrm{diag}^2(\boldsymbol{M}_{\boldsymbol{w}_L}\boldsymbol{ARA})\right). \tag{8}$$

[1] We define the detector as $\boldsymbol{M}_{\boldsymbol{w}_L}\boldsymbol{A}$ in order to involve only Hermitian matrices and simplify notation.

With $\xi_k \triangleq \sum_{\ell=0}^{L} w_\ell \lambda_k^{\ell+1}$, (8) becomes

$$S = \sum_{k=1}^{K} \left(\sum_{\mu=1}^{K} |\boldsymbol{T}_{\mu k}|^2 \xi_\mu \right)^2 = \sum_{\mu=1}^{K} \sum_{\nu=1}^{K} \xi_\mu \xi_\nu \sum_{k=1}^{K} |\boldsymbol{T}_{\mu k}|^2 |\boldsymbol{T}_{\nu k}|^2. \tag{9}$$

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

3.1 Optimum Weighting

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

$$SIR = \frac{S}{P-S} = \frac{1}{\frac{P}{S} - 1}. \tag{10}$$

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

We have

$$\frac{S}{P} = \frac{\sum\limits_{\mu=1}^{K} \sum\limits_{\nu=1}^{K} \lambda_\mu \lambda_\nu \sum\limits_{\ell=1}^{L} w_\ell \lambda_\mu^\ell \sum\limits_{l=1}^{L} w_l \lambda_\nu^l \sum\limits_{k=1}^{K} |\boldsymbol{T}_{\mu k}|^2 |\boldsymbol{T}_{\nu k}|^2}{\sum\limits_{k=1}^{K} (\lambda_k^2 + N_0 \lambda_k) \sum\limits_{\ell=1}^{L} w_\ell \lambda_k^\ell \sum\limits_{l=1}^{L} w_l \lambda_k^l}. \tag{11}$$

Defining the auxiliary matrices

$$\boldsymbol{T}_\mathrm{q} \triangleq \begin{bmatrix} |\boldsymbol{T}_{11}|^2 & \dots & |\boldsymbol{T}_{1K}|^2 \\ \vdots & \ddots & \vdots \\ |\boldsymbol{T}_{K1}|^2 & \dots & |\boldsymbol{T}_{KK}|^2 \end{bmatrix} \quad \text{and} \quad \boldsymbol{C} \triangleq \begin{bmatrix} 1 & \lambda_1 & \dots & \lambda_1^L \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_K & \dots & \lambda_K^L \end{bmatrix} \tag{12}$$

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

$$\frac{S}{P} = \frac{\boldsymbol{w}^\mathrm{T} \boldsymbol{C}^\mathrm{T} \boldsymbol{\Lambda} \boldsymbol{T}_\mathrm{q}{}^\mathrm{T} \boldsymbol{T}_\mathrm{q} \boldsymbol{\Lambda} \boldsymbol{C} \boldsymbol{w}}{\boldsymbol{w}^\mathrm{T} \boldsymbol{C}^\mathrm{T} (\boldsymbol{\Lambda}^2 + N_0 \boldsymbol{\Lambda}) \boldsymbol{C} \boldsymbol{w}}. \tag{13}$$

Moreover, let an arbitrary factorization of $\boldsymbol{C}^\mathrm{T}(\boldsymbol{\Lambda}^2 + N_0\boldsymbol{\Lambda})\boldsymbol{C}$ be defined by $\boldsymbol{F}^\mathrm{T}\boldsymbol{F} \triangleq \boldsymbol{C}^\mathrm{T}(\boldsymbol{\Lambda}^2 + N_0\boldsymbol{\Lambda})\boldsymbol{C}$. Let $\tilde{\boldsymbol{w}}$ be the eigenvector corresponding to the largest eigenvalue of $\boldsymbol{F}^{-T}\boldsymbol{C}^\mathrm{T}\boldsymbol{\Lambda}\boldsymbol{T}_\mathrm{q}{}^\mathrm{T}\boldsymbol{T}_\mathrm{q}\boldsymbol{\Lambda}\boldsymbol{C}\boldsymbol{F}^{-1}$. Then $\boldsymbol{w} = \boldsymbol{F}^{-1}\tilde{\boldsymbol{w}}$ maximizes the generalized Rayleigh fraction (13) and therefore also the SIR.

With this optimum choice for the weights the maximum achievable SIR becomes

$$\max_{\boldsymbol{w}_L} SIR = \frac{\mu_\mathrm{max}}{1 - \mu_\mathrm{max}} \tag{14}$$

where μ_max denotes the largest eigenvalue of $\boldsymbol{F}^{-T}\boldsymbol{C}^\mathrm{T}\boldsymbol{\Lambda}\boldsymbol{T}_\mathrm{q}{}^\mathrm{T}\boldsymbol{T}_\mathrm{q}\boldsymbol{\Lambda}\boldsymbol{C}\boldsymbol{F}^{-1}$.

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 $\boldsymbol{T}_\mathrm{q}$. 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.

3.2 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 [10]. Moreover, we assume the spreading sequences to be independent identically distributed.

3.2.1 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

$$\sum_{\mu=1}^{K} \sum_{\nu=1}^{K} \xi_\mu \xi_\nu = \mathrm{tr}^2 \left(\boldsymbol{M}_{\boldsymbol{w}_L} \boldsymbol{R} \right) \tag{15}$$

which leads to the following asymptotic equivalence [13]:

Lemma 1 *For independent identically distributed random sequences the asymptotic equivalence*

$$\lim_{K\to\infty} \frac{1}{K} \,\mathrm{tr} \left(\mathrm{diag}^2 (\boldsymbol{M}_{\boldsymbol{w}_L} \boldsymbol{R}) \right) = \lim_{K\to\infty} \frac{1}{K^2} \,\mathrm{tr}^2 (\boldsymbol{M}_{\boldsymbol{w}_L} \boldsymbol{R}) < \infty$$

holds with probability one for all $\boldsymbol{M}_{\boldsymbol{w}_L}$ *given by (6) and* $\boldsymbol{A} = \mathbf{I}$.

Lemma 1 allows for an asymptotic expression of the signal–to–total–power ratio which becomes

$$\frac{S}{P} \quad \longrightarrow \quad \lim_{K\to\infty} \frac{\frac{1}{K}\mathrm{tr}^2(\boldsymbol{M}_{\boldsymbol{w}_L}\boldsymbol{R})}{\mathrm{tr}(\boldsymbol{M}_{\boldsymbol{w}_L}^2 \boldsymbol{R}^2) + N_0 \mathrm{tr}(\boldsymbol{M}_{\boldsymbol{w}_L}^2 \boldsymbol{R})} \tag{16}$$

$$= \quad \lim_{K\to\infty} \frac{\frac{1}{K} \left(\sum\limits_{k=1}^{K} \sum\limits_{\ell=1}^{L} w_\ell \lambda_k^{\ell+1} \right)^2}{\sum\limits_{k=1}^{K} (\lambda_k^2 + N_0 \lambda_k) \sum\limits_{\ell=1}^{L} w_\ell \lambda_k^\ell \sum\limits_{l=1}^{L} w_l \lambda_k^l}. \tag{17}$$

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

$$\frac{S}{P} \longrightarrow \lim_{K\to\infty} \frac{\boldsymbol{w}^{\mathrm{T}} \boldsymbol{C}^{\mathrm{T}} \boldsymbol{\lambda} \boldsymbol{\lambda}^{\mathrm{T}} \boldsymbol{C} \boldsymbol{w}}{K \boldsymbol{w}^{\mathrm{T}} \boldsymbol{C}^{\mathrm{T}} (\boldsymbol{\Lambda}^2 + N_0 \boldsymbol{\Lambda}) \boldsymbol{C} \boldsymbol{w}}. \tag{18}$$

Comparing (18) and (13), we observe that $\boldsymbol{C}^{\mathrm{T}} \boldsymbol{\lambda} \boldsymbol{\lambda}^{\mathrm{T}} \boldsymbol{C}$ is an outer product of two identical vectors, i.e. its unique nonzero eigenvalue is $\boldsymbol{\lambda}^{\mathrm{T}} \boldsymbol{C} \boldsymbol{C}^{\mathrm{T}} \boldsymbol{\lambda}$, while $\boldsymbol{C}^{\mathrm{T}} \boldsymbol{\Lambda} \boldsymbol{T}_{\mathbf{q}}{}^{\mathrm{T}} \boldsymbol{T}_{\mathbf{q}} \boldsymbol{\Lambda} \boldsymbol{C}$ is fully ranked, 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 on (18) it can easily be checked that

$$\boldsymbol{w}_{\mathrm{asy},L} = \left(\boldsymbol{C}^{\mathrm{T}} (\boldsymbol{\Lambda}^2 + N_0 \boldsymbol{\Lambda}) \boldsymbol{C} \right)^{-1} \boldsymbol{\lambda}^{\mathrm{T}} \boldsymbol{C} \tag{19}$$

maximizes (18). Surprisingly, this asymptotically optimum weight vector is identical to that one found in [6] for minimization of the mean–squared error between the output signals of the approximate and the exact MMSE detector in the *non*–asymptotic case. With (19) the maximum asymptotic signal–to–total–power ratio becomes

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

The asymptotically optimum weights in (19) do not depend on the eigenvectors of the correlation matrix $\boldsymbol{R}$. Moreover, they are even almost surely totally independent of the correlation matrix due to the following result from random matrix theory:

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

$$\frac{1}{K} \sum_{k=1}^{K} \lambda_k^m = \frac{1}{K} \mathrm{tr} \left(\boldsymbol{S}^{\mathrm{H}} \boldsymbol{S} \right)^m \tag{20}$$

converge almost surely to the non–random limits

$$\sum_{i=0}^{m-1} \binom{m}{i} \binom{m}{i+1} \frac{\beta^i}{m}. \tag{21}$$

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

$$\boldsymbol{m}_L(\beta) \triangleq \lim_{K\to\infty} \frac{\boldsymbol{C}^{\mathrm{T}} \boldsymbol{\lambda}}{K} = \mathop{\mathsf{E}}_{\boldsymbol{S}} \left[\lambda, \lambda^2, \ldots, \lambda^{L+1} \right]^{\mathrm{T}}, \tag{22}$$

$$\boldsymbol{\Phi}_L(\beta, N_0) \triangleq \lim_{K\to\infty} \frac{\boldsymbol{C}^{\mathrm{T}} \left(\boldsymbol{\Lambda}^2 + N_0 \boldsymbol{\Lambda} \right) \boldsymbol{C}}{K} = \mathop{\mathsf{E}}_{\boldsymbol{S}} \begin{bmatrix} \lambda^2 + N_0 \lambda & \ldots & \lambda^{L+2} + N_0 \lambda^{L+1} \\ \vdots & \ddots & \vdots \\ \lambda^{L+2} + N_0 \lambda^{L+1} & \ldots & \lambda^{2L+2} + N_0 \lambda^{2L+1} \end{bmatrix}. \tag{23}$$

In terms of $\boldsymbol{m}_L(\beta)$ and $\boldsymbol{\Phi}_L(\beta, N_0)$ the asymptotic weights and the signal–to–total–power ratio simply read

$$\boldsymbol{w}_{\mathrm{asy},L} = \left(\boldsymbol{\Phi}_L(\beta, N_0) \right)^{-1} \boldsymbol{m}_L(\beta) \tag{24}$$

$$\frac{S}{P} \longrightarrow \boldsymbol{m}_L(\beta)^{\mathrm{T}} \left(\boldsymbol{\Phi}_L(\beta, N_0) \right)^{-1} \boldsymbol{m}_L(\beta). \tag{25}$$

Any scalar multiple of the asymptotically optimum weight vector is asymptotically optimum, too. This means it can be rescaled without loss of optimality by any scalar. Proper scaling yields that its components become L^{th} 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 simplystructured. For $L \leq 4$ the results are summarized in Table 1.

$L=1$	$w_0 = -N_0 w_1 + 2 + 2\beta$ $w_1 = -1$
$L=2$	$w_0 = -N_0 w_1 + 3 + 4\beta + 3\beta^2$ $w_1 = -N_0 w_2 - 3 - 3\beta$ $w_2 = 1$
$L=3$	$w_0 = -N_0 w_1 + 4 + 6\beta + 6\beta^2 + 4\beta^3$ $w_1 = -N_0 w_2 - 6 - 9\beta - 6\beta^2$ $w_2 = -N_0 w_3 + 4 + 4\beta$ $w_3 = -1$

Table 1: Optimum weight vectors $\boldsymbol{w}$asy$, L$ for $L = 1, 2, 3, 4$ after proper scaling.

Real time calculation of the asymptotically optimum weight vectors is surprisingly simple, as Table 1 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 1 indicates, this results only in negligible reduction of complexity.

3.2.2 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 mean–squared error can serve as performance criterion instead of signal–to–noise ratio. This approach was followed in [10] 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.

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

$$\hat{\boldsymbol{M}}_{\boldsymbol{w}_L} \triangleq \sum_{i=0}^{L} w_i \boldsymbol{R}^i. \tag{26}$$

For its analysis, the following lemma which follows form the definition of free random variables in [15] will be crucial:

Lemma 2 *Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be two independent[2] $K \times K$ random matrices with entries having finite moments. Then, with probability one, we have*

$$\lim_{K\to\infty} \frac{1}{K} \operatorname{tr}(\boldsymbol{XY}) = \lim_{K\to\infty} \frac{1}{K} \operatorname{tr}(\boldsymbol{X}) \cdot \lim_{K\to\infty} \frac{1}{K} \operatorname{tr}(\boldsymbol{Y}). \tag{27}$$

The total received power and the useful signal power for the detector defined in (26) are obviously given by

$$\hat{P} = \operatorname{tr}\left(\boldsymbol{R}\hat{\boldsymbol{M}}^2_{\boldsymbol{w}_L}\boldsymbol{R}\boldsymbol{A}^2\right) + N_0 \operatorname{tr}\left(\hat{\boldsymbol{M}}^2_{\boldsymbol{w}_L}\boldsymbol{R}\right) \tag{28}$$

$$\hat{S} = \operatorname{tr}\left(\operatorname{diag}^2\left(\hat{\boldsymbol{M}}_{\boldsymbol{w}_L}\boldsymbol{R}\boldsymbol{A}\right)\right), \tag{29}$$

[2]It would be sufficient to assume freeness (see [15] for definition).

respectively. Note the discrepancies to (7) and (8) due to the lack of scaling with $\boldsymbol{A}$ in (26). With the help of Lemma 2 and the normalization of $\boldsymbol{A}$ in (2), we get

$$\frac{\hat{S}}{\hat{P}} \longrightarrow \lim_{K\to\infty} \frac{S}{P} \tag{30}$$

the same asymptotic signal–to–total–power ratio as found for equal powers in (16). Obviously, this equivalence also yields equivalence of the SIRs in both cases.

The signal–to–noise ratio for unequal powers is a quantity averaged over all users. Large signal–to–noise ratios 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' SIRs gives more insight. It is proven in the following that

$$\lim_{K\to\infty} \mathit{SIR}_k = A_k^2 \lim_{K\to\infty} \mathit{SIR} \tag{31}$$

where A_k denotes the k^{th} 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 signal–to–noise ratio is given by

$$\lim_{K\to\infty} \mathit{SIR} = \lim_{K\to\infty} \frac{1}{K} \sum_{k=1}^{K} \frac{\mathit{SIR}_k}{A_k^2}. \tag{32}$$

As SIR is asymptotically independent of the choice of $\boldsymbol{A}$, this yields (31).

Optimizing the weights for equal powers, but applying them in the unequal power case yields the same average SIR, as applying them in the equal power case which they are designed for. To be more precise:

Theorem 2 *Let a polynomial multiuser detector be defined as in (26), 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 SIR 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 (26) does not involve the users' powers. Therefore, it cannot be the best L^{th} 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 (6) with weights optimized with respect to SIR performs at least as well as that one defined in (26) with equality holding only in the equal power case. This implies that regarding the detector defined in (6) there is a saddle point of the SIR as a function of the weights and the power distribution. To be more precise:

Theorem 3 *Let a polynomial multiuser detector be defined as in (6), 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 $P_i = 1$. 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 SIR decreases and for any different power distribution the average SIR 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.

4 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 assumption of random spreading is an accurate model for long–code CDMA. It is discussed for decorrelating and/or MMSE detectors in [16, 17, 18, 19, 20]. We will analyze the performance of the approximation to decorrelating and MMSE detectors in terms of spectral efficiency. Hereby, we restrict our considerations to the case of equal powers due to the reasons outlined in Section 3.2.2.

Due to 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 SIR in terms of the load and noise power density. Because of (10) it is sufficient to find the signal–to–total power ratio. With (24) and (25), it becomes

$$\frac{S}{P} \longrightarrow \boldsymbol{m}_L^{\mathrm{T}}(\beta)\boldsymbol{w}_{\mathrm{asy},L}. \tag{33}$$

With these preliminaries (10) allows to give explicit expressions for the optimum SIRs

$$\begin{aligned}
\max_{\boldsymbol{w}_0} SIR_0 &\longrightarrow \frac{1}{\beta+N_0} = SIR_{\mathrm{MF}}\\
\max_{\boldsymbol{w}_1} SIR_1 &\longrightarrow \frac{1+\beta+N_0}{\beta^2+N_0(1+2\beta)+N_0^2} \overset{\beta>0}{>} \frac{1-2\beta+\beta^2}{\beta^2+\beta^3+N_0(1-\beta+\beta^2)} \longleftarrow SIR_{\mathrm{AD}}\\
\max_{\boldsymbol{w}_2} SIR_2 &\longrightarrow \frac{1+\beta+\beta^2+N_0(2+2\beta)+N_0^2}{\beta^3+N_0(1+2\beta+3\beta^2)+N_0^2(2+3\beta)+N_0^3}\\
\max_{\boldsymbol{w}_3} SIR_3 &\longrightarrow \frac{1+\beta+\beta^2+\beta^3+N_0(3+4\beta+3\beta^2)+N_0^2(3+3\beta)+N_0^3}{\beta^4+N_0(1+2\beta+3\beta^2+4\beta^3)+N_0^2(3+6\beta+6\beta^2)+N_0^3(3+4\beta)+N_0^4}
\end{aligned}$$

Obviously, the 0^{th} order approximation is equivalent to the conventional matched filter (MF). Our first order approximation is better than the approximate decorrelator (AD) analyzed in [1, p. 281], where the weights were based on a first order Taylor series and not optimized with respect to the maximum achievable SIR, cf. (3).

Note that due to Lemma 1, the asymptotic SIR depends only on the moments of the eigenvalue density which converge in probability. This means:

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

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

$$\lim_{N_0\to 0}\max_{\boldsymbol{w}_L} SIR_L \longrightarrow \frac{\beta^{-L-1}-1}{1-\beta}. \tag{34}$$

It was found in [21, 22] in related, but different, context and also holds for polynomial approximations to linear multiuser detection. Note that the SIR is finite even when no AWGN is present. However, the asymptotic SIR grows exponentially with the order of the approximation for $\beta < 1$.

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

$$\frac{N_0}{E_\mathrm{b}} \triangleq N_0 C = N_0 \log_2\left(1 + SIR(N_0, \beta)\right), \tag{35}$$

$$\Gamma \triangleq \beta C = \beta \log_2\left(1 + SIR(N_0, \beta)\right), \tag{36}$$

respectively [16, 17], 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 (35) and (36). While for the MMSE detector with random spreading an explicit expression can be found, see [16], this is not possible for approximations to linear receivers, in general. In the following extreme case, however, we find

$$\lim_{\frac{E_\mathrm{b}}{N_0} \to \infty} \Gamma_L = \beta \log_2\left(\frac{\beta - \beta^{-L-1}}{\beta - 1}\right) \tag{37}$$

which is illustrated in Fig. 1. It can be observed that around $\beta \approx 0.4$ spectral efficiency

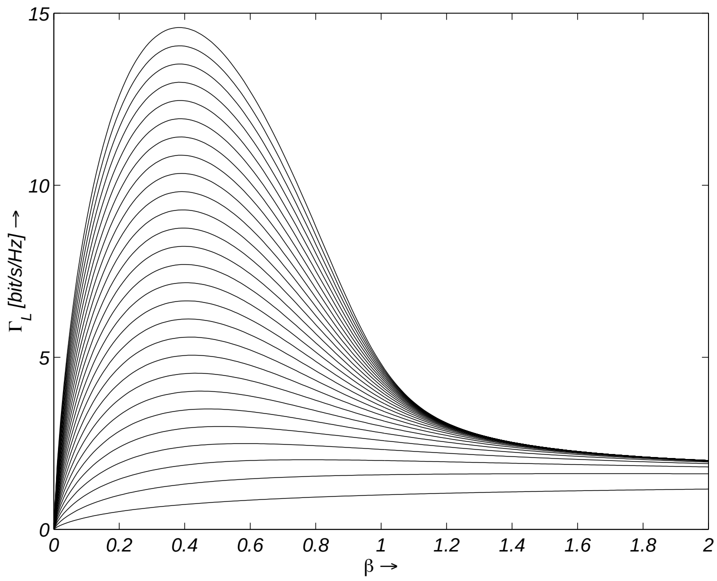

Figure 1: Spectral efficiency vs. load for 0^{th} to 26^{th} order approximation as $E_\mathrm{b}/N_0 \to \infty$.

increases by about half a bit/s/Hz with each additional stage of the approximation. Thus, $\max \Gamma_L \approx L/2 + 1$ is an accurate rule of thumb.

Fig. 2 shows spectral efficiency as a function of the load for a fixed signal–to–noise ratio that represents a typical setting in practice. The approximate MMSE receiver is found to approximate its exact counterpart rather fast, while the approximate decorrelator specified in (3) 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

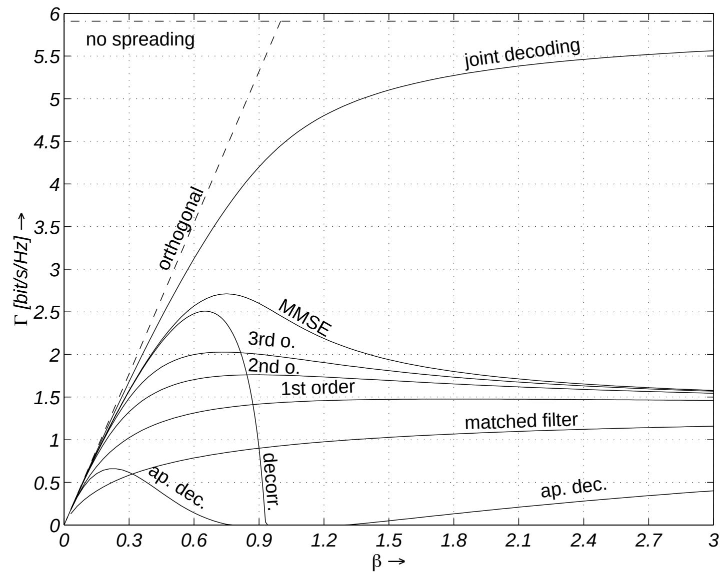

Figure 2: Spectral efficiency vs. system load for several linear multiuser receivers and fixed $10\log_{10}(E_\mathrm{b}/N_0) = 10$ dB.

the conventional matched filter even for multistage approximations. This is illustrated in Fig. 3 for the Taylor approximation defined in (4). For all orders shown, the range of

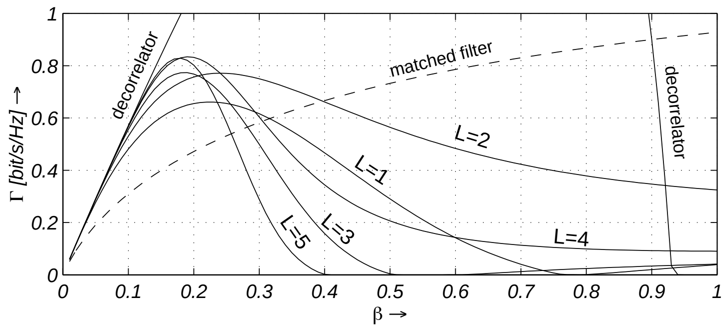

Figure 3: Spectral efficiency vs. system load for the first five Taylor approximations to the decorrelator and fixed $10\log_{10}(E_\mathrm{b}/N_0) = 10$ dB.

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 $\beta > (\sqrt{2}-1)^2 \approx 0.17$ [8]. Remarkably, the Taylor approximations with even order are better with respect to a performance–complexity tradeoff.

5 Conclusions

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. With these results, complexity increase is hardly a justification for denying CDMA receivers based on multiuser detection even for long–code CDMA.

References

[1] S. Verdú. *Multiuser Detection*. Cambridge University Press, New York, 1998.

[2] V.A. Marcenko and L.A. Pastur. Distribution of eigenvalues for some sets of random matrices. *Math. USSR-Sb*, (1):457–483, 1967.

[3] R. Kohno, M. Hatori, H. Imai. Cancellation techniques of co–channel interference in asynchronous spread spectrum multiple access systems. *Electr. & Commun.*, 66–A(5):20–29, 1983.

[4] S. Verdú. Multiuser detection. In H.V. Poor and J.B. Thomas, editors, *Advances in Statistical Signal Processing: Signal Detection*, pp. 369–410. JAI Press, Greenwich, CT, 1993.

[5] R.S. Varga. *Matrix Iterative Analysis*. Prentice–Hall, Englewood Cliffs, NJ, 1962.

[6] S. Moshavi, E.G. Kanterakis, D.L. Schilling. Multistage linear receivers for DS–CDMA systems. *Int'l J. Wireless Inform. Networks*, 3(1):1–17, Jan. 1996.

[7] D. Divsalar, M.K. Simon, D. Raphaeli. Improved parallel interference cancellation for CDMA. *IEEE Trans. Commun.*, 46(2):258–268, Feb. 1998.

[8] A. Grant and C. Schlegel. Iterative implementations for linear multiuser detectors. Submitted to *IEEE Trans. Commun.*, Feb. 1999.

[9] O. Axelsson. *Iterative Solution Methods*. Cambridge University Press, New York, 1994.

[10] D. Guo, L.K. Rasmussen, T.J. Lim. Linear parallel interference cancellation in long–code CDMA multiuser detection. *IEEE J. Sel. Ar. Commun.*, 17(12):2074–2081, Dec. 1999.

[11] D. Guo, L.K. Rasmussen, S. Sun, T.J. Lim. A matrix–algebraic approach to linear parallel interference cancellation in CDMA. *IEEE Trans. Commun.*, 48(1):152–161, Jan. 2000.

[12] E. Kanterakis and S. Moshavi. Multistage linear receiver for DS–CDMA systems, May 1998. US–patent US5757791.

[13] R.R. Müller and S. Verdú. Design and analysis of low–complexity interference mitigation on vector channels. Submitted to *IEEE J. Sel. Ar. Commun.*, Dec. 2000.

[14] D. Jonsson. Some limit theorems for the eigenvalues of a sample covariance matrix. *J. Multivar. Analysis*, 12:1–38, 1982.

[15] D. V. Voiculescu, K. J. Dykema, A. Nica. *Free Random Variables*. American Mathematical Society, Providence, RI, 1992.

[16] R.R. Müller. *Power and Bandwidth Efficiency of Multiuser Systems with Random Spreading*. Shaker–Verlag, Aachen, Germany, 1999.

[17] S. Verdú and S. Shamai. Spectral efficiency of CDMA with random spreading. *IEEE Trans. Inform. Th.*, 45(2):622–640, Mar. 1999.

[18] P. Schramm and R.R. Müller. Spectral efficiency of CDMA systems with linear MMSE interference suppression. *IEEE Trans. Commun.*, 47(5):722–731, May 1999.

[19] R.R. Müller. Multiuser receivers for randomly spread signals: Fundamental limits with and without decision–feedback. *IEEE Trans. Inform. Th.*, 47(1), Jan. 2001.

[20] S. Shamai and S. Verdú. The effect of frequency–flat fading on the spectral efficiency of CDMA. Submitted to *IEEE Trans. Inform. Th.*, 1999.

[21] M.L. Honig and W. Xiao. Large system performance of reduced–rank linear filters. In *Proc. of 37th Ann. Allerton Conf. Commun., Contr. & Comp.*, Monticello, IL, Sep. 1999.

[22] M. Honig and W. Xiao. Performance of reduced–rank linear interference suppression for DS–CDMA. Submitted to *Trans. Inform. Th.*, Dec. 1999.