Asymptotic Analysis of Reduced-Rank Chip-Level MMSE Equalizers in the Downlink of CDMA Systems.

Belkacem Mouhouche[1], Philippe Loubaton[2], Walid Hachem[3]

Abstract

In this paper, we address the performance of downlink CDMA receivers that consist in a reduced rank Wiener chip-rate equalizer followed by a despreading operation. In particular, we tackle the question of whether a performance close to the optimum can be achieved for small values of the rank. To answer this question, it is standard to consider the output Signal to Interference plus Noise Ratio (SINR), and to study its convergence speed versus the rank of the receiver. Unfortunately, this task is difficult due to the fact that the SINR expressions depend on the spreading codes allocated to the various users in a rather complicated way. To be able to obtain positive results, we assume that the spreading factor and the number of users tend to infinity while their ratio remains finite. As in the 3rd generation UMTS systems, the spreading codes we consider coincide with orthogonal Walsh Hadamard codes scrambled by an independent identically distributed sequence. In this context, we show that the SINR of each reduced-rank receiver converges towards a deterministic limit which depends only of the rank of the receiver, and not of the spreading codes given to the various users. Using some previous results, we prove that the asymptotic SINRs converge exponentially to the SINR of the plain Wiener receiver when the rank of the receiver increases. We obtain the corresponding convergence rate, and exhibit the parameters that influence the convergence speed. We finally compare our asymptotic performance expressions with the results of numerical simulations. We observe a good agreement for spreading factors as low as 16.

Index Terms

Asymptotic analysis, Reduced-rank filtering, chip-rate MMSE equalizer, downlink CDMA.

1- Freescale Seminconductor, Avenue du Général Eisenhower,31000 Toulouse, France (e-mail: mouhouche@ieee.org).
2- Université de Marne La Vallée, 5 Bld Descartes,77454 Marne La Vallée, France (email: loubaton@univ-mlv.fr).
3- Supélec, Plateau de Moulon, 91192 Gif-sur-Yvette Cedex, France (e-mail: walid.hachem@supelec.fr).

I. INTRODUCTION

Reduced-rank filtering has been considered in many areas of signal processing such as space-time coding, array processing, radar, and channel equalization [18], [22], [10]. Recently, reduced-rank filtering has been applied to interference cancellation for Direct Sequence Code Division Multiple Access (DS-CDMA) communication systems. In the case of short-code CDMA, conventional adaptive filtering algorithms can be applied. However, the Minimum Mean Squared Error (MMSE) filter calculation requires the inversion of the (estimated) received signal covariance matrix after each update. This represents a huge computational burden, especially when the spreading factor is large. Furthermore, for fast varying channels, one needs to estimate the MMSE filter from a small number of training data over which the channel can be reasonably considered as stationary. Trying to adapt a large number of coefficients slows down the speed of convergence of the filter and its tracking ability. Reduced-rank filtering allows to reduce the number of coefficients to be updated, provides a satisfying trade off between performance and complexity, and offers a better tracking ability when a small number of samples are used.

Performance analysis of existing CDMA receivers has attracted a lot of attention recently. It is usual to use the Signal to Interference plus Noise Ratio (SINR) as a measure of performance. Most frequently, the SINR analytical expressions are difficult to interpret because they depend in a complex manner on the spreading codes allocated to the various users. To overcome this difficulty, it was proposed in [24] to model the code matrix as a realization of a random matrix with independent identically distributed (i.i.d) entries and to evaluate the limit SINR in a certain large system regime. The large-system (asymptotic) regime is obtained by letting the spreading factor N and the number of users K both tend to infinity, while their ratio remains constant. Under these conditions, the SINR of the optimum MMSE receiver for unfaded CDMA was shown to converge to a deterministic limit β independent of the particular entries of the code matrix [24]. In the same context, the reduced-rank optimum MMSE receiver was analyzed in [11] where it was shown that the convergence of the reduced-rank SINR β_n (where n represents the rank of the filter) to the full rank SINR β is very fast, *i.e.*, $\beta_n \simeq \beta$ for small values of n.

In a more general context, the influence of the rank n on the performance of a reduced-rank MMSE receiver was analyzed in [19]. Under the hypothesis that successive powers of the covariance matrix post multiplied by the desired user signature and premultiplied by its conjugate tend to a finite limit s_k when N and K tend to $+\infty$ with constant ratio, it was shown there that finite size SINRs converge to finite limits β and β_n respectively. More importantly, the convergence speed of β_n towards β can be evaluated using the properties of certain orthogonal polynomials.

In long spreading code downlink CDMA systems the covariance matrix of the observation is unknown. Due to the presence of long codes, the covariance matrix is time-varying, and therefore cannot be estimated at the receiver side. This prevents the use of conventional MMSE receivers in this context. Chip-rate MMSE equalization followed by despreading was proposed as an alternative ([7], [15], [16], [13]). The corresponding receiver is usually called the suboptimum Wiener receiver because it has no knowledge of the interfering users codes, unlike the optimum Wiener receiver. As in the short code case, and for the same reasons, a reduced rank version of the chip rate MMSE equalizer can be implemented here.

Since the suboptimum Wiener receivers are more recent (in both their full rank and reduced rank versions) than their optimum counterparts, their large system asymptotic performance analysis has received much less attention. The suboptimum full-rank MMSE receiver was analyzed in [4] in the context of certain random orthogonal code matrices. However, to our knowledge, the asymptotic performance of reduced-rank suboptimum receivers has not been studied yet.

In this paper, we consider the asymptotic performance of suboptimum reduced rank receivers based on reduced rank chip rate Wiener equalizers followed by a despreading. Motivated by the specifications of the UMTS system, we model the code matrices as Walsh-Hadamard matrices scrambled by a realization of an i.i.d sequence. The code matrices can thus be interpreted as realizations of particular random matrices. We prove that when the number of users and the spreading factor grow without bound while their ratio remains constant, the SINR of full and reduced-rank suboptimum receivers tend to deterministic limits, and these are independent of the particular realization of the scrambling code. Using the results of [19], we characterize the convergence speed of the reduced-rank receiver SINR towards the full rank receiver SINR and draw some conclusions about the effect of the channel and the load factor on the convergence speed.

The rest of the paper is organized as follows. We first summarize some previous contributions in section II. In section III, we present the main results of [19]. In section IV, we present the downlink CDMA system model as well as the reduced-rank Wiener equalizers under consideration. In section V, we study the performance of the aforementioned receivers in the asymptotic regime. We show that the hypotheses formulated in section III are valid, and deduce the convergence speed of the reduced-rank chip-rate equalizer receivers. Finally, we compare in section VI our asymptotic predictions with the empirical performance obtained through numerical simulations. The asymptotic and empirical performance are quite close to each other for both the standard Vehicular A channel (short delay spread) and the Vehicular B channel (long delay spread) even for short spreading factors.

Frequently used notations.

N : spreading factor, K : number of users

n : order of the reduced rank receiver

m : time symbol index, i : time chip index

$b(m) = (b_1(m), \ldots, b_K(m))^T$: vector of transmitted symbols at time m, $(d(i))_{i\in\mathbb{Z}}$: chip sequence

$\mathbf{S}(m)$: diagonal $N \times N$ matrix whose entries represent the scrambling code, $\mathbf{C} = (\mathbf{c}_1, \mathbf{C}_2)$: $N \times K$ Walsh-Hadamard matrix

$\mathbf{W}(m) = \mathbf{S}(m)\mathbf{C}$: $N \times K$ code matrix allocated to the various users

$h(z)$: transfer function of the channel, $\mathcal{H}$: $2N \times 3N$ Toeplitz matrix associated with $h(z)$

$g_n(z)$: transfer function of the rank n Wiener filter, $\mathcal{G}_n$: $N \times 3N$ Toeplitz matrix associated with $g_n(z)$

$f_n(z) = g_n(z)h(z)$, $\mathcal{F}_n$: $N \times 4N$ Toeplitz matrix associated with $f_n(z)$

II. Background on reduced-rank Wiener filters.

To fix our ideas, let us begin with the general signal model

$$\mathbf{y} = \mathbf{h}b + \mathbf{x} \tag{1}$$

where $\mathbf{y}$ is the $N \times 1$ received signal, b is a unit-variance scalar signal to be estimated and $\mathbf{x}$ is a signal uncorrelated with b that accounts for interference and/or noise. The $N \times N$ covariance matrix of $\mathbf{x}$ is denoted $\mathbf{R}_I$ and will be assumed invertible.

We consider the problem of estimating the scalar b from the received signal $\mathbf{y}$ using a $1 \times N$ linear receiver $\mathbf{g}$. The soft estimate $\tilde{b}$ is given by:

$$\tilde{b} = \mathbf{g}\mathbf{y} \tag{2}$$

The MMSE (Wiener) receiver minimizes the Mean-Squared estimation Error (MSE), and is of course given by:

$$\mathbf{g}\mathbf{R} = \mathbf{h}^H \tag{3}$$

where $\mathbf{R} = \mathbf{h}\mathbf{h}^H + \mathbf{R}_I$ is the received signal $\mathbf{y}$ covariance matrix. This receiver will be called in the sequel the full rank MMSE receiver. Its output SINR that we index by the number of dimensions of the received signal is given by the standard expression:

$$\beta^{(N)} = \frac{\eta^{(N)}}{1 - \eta^{(N)}} \tag{4}$$

where $\eta^{(N)}$ is defined by

$$\eta^{(N)} = \mathbf{h}^H \mathbf{R}^{-1} \mathbf{h} \ . \tag{5}$$

In the context of reduced-rank methods considered in this paper, filter $\mathbf{g}$ is constrained to lie in the n–dimensional Krylov subspace associated with the pair $(\mathbf{R}, \mathbf{h})$, i.e. the subspace generated by the n-columns of the Krylov matrix

$$\mathbf{K}_n = [\mathbf{h}, \mathbf{R}\mathbf{h}, \ldots, \mathbf{R}^{n-1}\mathbf{h}]$$

The motivation behind choosing the Krylov subspaces and the implementation of the subsequent filters are discussed in a number of works (see e.g. [2], [11] and [9]). The corresponding reduced-rank Wiener filter $\mathbf{g}_n$ is defined by $\mathbf{g}_n = \mathbf{f}_n \mathbf{K}_n^H$ where the n-dimensional vector $\mathbf{f}_n$ is solution of the linear system:

$$\mathbf{f}_n \mathbf{K}_n^H \mathbf{R} \mathbf{K}_n = \mathbf{h}^H \mathbf{K}_n. \tag{6}$$

Contrary to (3), (6) is a system of n linear equations. Therefore, confining the filtering operation to a low-dimensional subspace leads to a substantial gain in complexity when $n \ll N$.

The output SINR $\beta_n^{(N)}$ of the n-th order reduced-rank Wiener filter is given by

$$\beta_n^{(N)} = \frac{\eta_n^{(N)}}{1 - \eta_n^{(N)}} \tag{7}$$

where $\eta_n^{(N)}$ is now defined by

$$\eta_n^{(N)} = \mathbf{h}^H \mathbf{K}_n \left(\mathbf{K}_n^H \mathbf{R} \mathbf{K}_n\right)^{-1} \mathbf{K}_n^H \mathbf{h} \tag{8}$$

Confining the Wiener filter to a low-dimensional subspace is of course not optimal. Therefore, this operation should decrease the SINR at the receiver output:

$$\beta^{(N)} \geq \beta_n^{(N)} \tag{9}$$

However, the use of reduced-rank Wiener filters is of course attractive if a performance close to the optimum can be achieved for small values of n. In order to make it clear in which contexts this nice condition holds, the convergence speed of $\beta_n^{(N)}$ to $\beta^{(N)}$, or equivalently, of $\eta_n^{(N)}$ to $\eta^{(N)}$ when n increases has to be studied. This problem has been successfully addressed in recent works in the context of the following simple CDMA transmission model

$$\mathbf{y} = \mathbf{W}\mathbf{b} + \mathbf{v} \tag{10}$$

where $\mathbf{b} = [b_1, \ldots, b_K]^T$ is the $K \times 1$ symbol vector, K is the number of users, N is the spreading factor, $\mathbf{W}$ is the $N \times K$ spreading code matrix, and $\mathbf{v}$ is the classical noise with covariance matrix $\sigma^2 \mathbf{I}$. The purpose is to estimate the symbol b_1. We denote by $\mathbf{w}$ the first column of $\mathbf{W}$ (i.e. the code vector of the user of interest: user 1), and by $\mathbf{U}$ the $N \times (K-1)$ matrix such that

$$\mathbf{W} = [\mathbf{w}, \mathbf{U}]$$

similarly,

$$\mathbf{b} = [b_1 \ \mathbf{b}_I^T]^T$$

where $\mathbf{b}_I$ is the vector of the interferers symbols. By replacing $\mathbf{h}$ by $\mathbf{w}$ and letting

$$\mathbf{x} = \mathbf{U}\mathbf{b}_I + \mathbf{v}$$

model (10) shows to be a particular case of (1).

It is in general very difficult to study the convergence of $\beta_n^{(N)}$ to $\beta^{(N)}$, because both SINRs (reduced-rank and full-rank) depend in a complex way on the spreading codes. A noticeable exception is given by the case where $\mathbf{W}$ is obtained using the same length Gold code [8]. In this context, the authors of [8] showed that a performance equal to the full-rank MMSE receiver is obtained for n as small as 2 if the Gold code is well chosen.

In order to address more general situations, Honig and Xiao [11] followed the attractive approach introduced for the first time in [24]. They modelled the code matrix $\mathbf{W}$ as a realization of a random matrix with centered i.i.d. elements with variance $1/N$ and studied the performance of the reduced-rank filter in the "large system" regime where N and K tend to infinity in such a way that K/N converges towards a constant α. They established that $\beta_n^{(N)}$ and $\beta^{(N)}$ converge to finite limits β_n and β independent of the particular realizations of the code matrix. Therefore, at least for N and K large enough, it is possible to replace the study of the speed of convergence of $\beta_n^{(N)}$ towards $\beta^{(N)}$ by that of β_n towards β, which is a much simpler problem. They were able to show that β is a continued fraction expansion whose order n truncation coincides with β_n. They obtained the following recurrence equation :

$$\beta_{n+1} = \frac{1}{\sigma^2 + \alpha \frac{1}{1+\beta_n}} \tag{11}$$

From this expression, they concluded for the rapid convergence of this SINR towards the full rank SINR using some numerical results.

Partial results have been obtained in the large system regime for models more general than (10) (see e.g. [5] and [17]). In these works, the convergence of $\beta_n^{(N)}$ towards β_n is established. However, the convergence speed of β_n towards β is not addressed.

In [19], the influence of n on the performance of the receiver in the asymptotic regime when $N \to +\infty$ was studied in the much more general context defined by model (1). Under the hypothesis that for each integer k, $s_k^{(N)} = \mathbf{h}^H \mathbf{R}_I^k \mathbf{h}$ converges when $N \to +\infty$ to a finite limit s_k, it was shown in this contribution that $\beta^{(N)}$ and $\beta_n^{(N)}$ also converge to certain finite limits β and β_n respectively. More importantly, the convergence speed of β_n towards β can be evaluated using properties of certain orthogonal polynomials. We give here a brief review of the main results of [19] that will be used later to derive the asymptotic performance of reduced-rank equalization.

III. Review of the main results of [19]

We still consider the general model (1). It will be more convenient work on the asymptotic behaviour of $\eta_n^{(N)}$ and $\eta^{(N)}$. The reader may check that if one replaces matrix $\mathbf{R}_I$ by matrix $\mathbf{R}$ in [19], then all the results of sections 2 and 3 of [19] that concern $\beta_n^{(N)}$ and $\beta^{(N)}$ can be adapted to $\eta_n^{(N)}$ and $\eta^{(N)}$. In this section, we review the corresponding results.

Assumption 1: We assume that for each k, $s_k^{(N)} = \mathbf{h}^H \mathbf{R}^k \mathbf{h}$ converges when $N \to +\infty$ to a finite limit s_k, and that $s_0 = 1$.

It is easily seen that $\eta_n^{(N)}$ is equal to

$$(s_0^{(N)}, \ldots, s_{n-1}^{(N)}) \begin{pmatrix} s_1^{(N)} & s_2^{(N)} & \ldots & s_n^{(N)} \\ s_2^{(N)} & s_3^{(N)} & \ldots & s_{n+1}^{(N)} \\ \vdots & \vdots & \vdots & \vdots \\ s_n^{(N)} & s_{n+1}^{(N)} & \ldots & s_{2n-1}^{(N)} \end{pmatrix}^{-1} \begin{pmatrix} s_0^{(N)} \\ \vdots \\ s_{n-1}^{(N)} \end{pmatrix} \tag{12}$$

Assumption 1 thus implies that for each n, $\eta_n^{(N)}$ converges to the quantity η_n obtained by replacing $(s_k^{(N)})_{k=0,2n-1}$ in (12) by sequence $(s_k)_{k=0,2n-1}$. Moreover, $\mathbf{K}_n^H \mathbf{K}_n$ (recall that $\mathbf{K}_n$ represents the n-th order Krylov matrix of $(\mathbf{R}, \mathbf{h})$) and $\mathbf{K}_n^H \mathbf{R} \mathbf{K}_n$ are positive Hankel matrices converging to the Hankel matrices $(s_{k+l})_{(k,l)=0,\ldots,n-1}$ and $(s_{k+l+1})_{(k,l)=0,\ldots,n-1}$. Therefore, matrices $(s_{k+l})_{(k,l)=0,\ldots,n-1}$ and $(s_{k+l+1})_{(k,l)=0,\ldots,n-1}$ are also positive. Using well known results (see e.g. [1]), there exists a probability measure ν such that

$$s_k = \int_0^\infty \lambda^k d\nu(\lambda)$$

Assumption 2: Measure ν is carried by an interval $[\delta_1, \delta_2]$, $\delta_1 > 0$, $\delta_2 < +\infty$, and is thus uniquely defined ([1]). Moreover, ν is absolutely continuous, and its probability density is almost surely strictly positive on $[\delta_1, \delta_2]$.

Assumption 3: there exist $A > 0$ and $B > 0$ such that $\sup_N \|\mathbf{R}^{-1}\| \leq A$ and $\sup_N \|\mathbf{R}\| \leq B$. Under the above assumptions, $\eta^{(N)} = \mathbf{h}\mathbf{R}^{-1}\mathbf{h}$ can be shown to converge to $\eta = \int_{\delta_1}^{\delta_2} \frac{1}{\lambda} d\nu(\lambda)$. Therefore, if N is large enough, it is relevant to study the convergence speed of η_n towards η, a simpler problem. For this, we have to evaluate the convergence speed of

$$\eta_n = (s_0, \ldots, s_{n-1}) \begin{pmatrix} s_1 & s_2 & \ldots & s_n \\ s_2 & s_3 & \ldots & s_{n+1} \\ \vdots & \vdots & \vdots & \vdots \\ s_n & s_{n+1} & \ldots & s_{2n-1} \end{pmatrix}^{-1} \begin{pmatrix} s_0 \\ \vdots \\ s_{n-1} \end{pmatrix}$$

towards $\eta = \int_{\delta_1}^{\delta_2} \frac{1}{\lambda} d\nu(\lambda)$. The main result of [19] is the following theorem.

Theorem 1: Let $\mu > 1$ and $\phi < 1$ be defined by $\mu = \frac{1+\frac{\delta_1}{\delta_2}}{1-\frac{\delta_1}{\delta_2}}$ and $\phi = \frac{1}{\mu+\sqrt{\mu^2-1}}$. Then, there exist two strictly positive constants C and D such that

$$C\phi^{2n} \leq (\eta - \eta_n) \leq D\phi^{2n} \tag{13}$$

for n large enough.

This result implies that the convergence of η_n towards η is locally exponential, and that its rate only depends on the ratio $\frac{\delta_1}{\delta_2}$, and not on the particular form of measure ν. In particular,if $\frac{\delta_1}{\delta_2}$ is close to 0, then μ is close to 1, and the convergence is slow. If however $\frac{\delta_1}{\delta_2}$ is close to 1, then μ is large, and the convergence is fast.

IV. Reduced-Rank Equalization for CDMA Downlink

From now on, we consider a downlink CDMA system. A base station transmits K symbol sequences $(b_k)_{k=1,\ldots,K}$ to K mobile units of the corresponding cell. It is assumed that the number of users K is smaller than the spreading factor N. Motivated by the specifications of the downlink of the Third Generation (3G) mobile communication systems (UMTS) [6], we assume that the spreading codes change from one symbol to another, and that code matrix $\mathbf{W}(m)$ at time m is given by

$$\mathbf{W}(m) = \mathbf{S}(m)\mathbf{C} \tag{14}$$

where:

- $\mathbf{C}$ is a time-invariant orthogonal $N \times K$ matrix obtained by extracting K columns from a $N \times N$ Walsh-Hadamard matrix (this implies that each entry of $\mathbf{C}$ is equal to $\pm\frac{1}{\sqrt{N}}$),
- $\mathbf{S}(m) = \text{diag}(s_1(m), \ldots, s_N(m))$ is a time-varying diagonal matrix whose entries $(s_l(m))_{l=1,\ldots,N}$ are QAM4 distributed ($s_l(m) \in \{\pm\frac{1}{\sqrt{2}} + \pm i\frac{1}{\sqrt{2}}\}$) and represent the long scrambling code of the cell under consideration.

We notice that $\mathbf{W}(m)^H\mathbf{W}(m) = \mathbf{I}$ for each m. Of course, we take into account the effect of the propagation channel between the base station and the mobile of interest (say mobile 1), and we denote by

$$h(z) = \sum_{l=0}^{L-1} h_l z^{-l}$$

its chip rate discrete-time equivalent transfer function. $h(z)$ is assumed to be known at the receiver side, and is normalized in such a way that $\sum_{l=0}^{L-1} |h_l|^2 = 1$. $(d(i))_{i\in\mathbb{Z}}$ represents the chip sequence transmitted by the base station. Therefore, the received signal $(y(i))_{i\in\mathbb{Z}}$ sampled at the chip rate can be written as

$$y(i) = \sum_{l=0}^{L-1} h_l d(i-l) + v(i) \tag{15}$$

where v is an additive white noise of variance σ^2. It is more convenient to express this in a matrix form. Let

$$\mathbf{d}(m) = \big(d(mN), d(mN+1), ..., d(mN+N-1)\big)^T$$

be the transmitted chip-vector sequence at symbol instant m. $\mathbf{d}(m)$ is of course given by

$$\mathbf{d}(m) = \mathbf{W}(m)\mathbf{b}(m) \tag{16}$$

where $\mathbf{b}(m) = (b_1(m), \ldots, b_K(m))^T$ represents the K symbols transmitted at time m by the base station. We put $\mathbf{y}(m) = \big(y(mN), y(mN+1), ..., y(mN+N-1)\big)^T$. Then, (15) is equivalent to

$$\mathbf{y}(m) = \mathbf{H}_0\mathbf{d}(m) + \mathbf{H}_1\mathbf{d}(m-1) + \mathbf{v}(m) \tag{17}$$

where

$$\mathbf{H}_0 = \begin{bmatrix} h[0] & 0 & & 0 \\ \vdots & h[0] & & \\ h[L-1] & & & \\ & \ddots & \ddots & \\ & & & \\ 0 & & h[L-1] & h[0] \end{bmatrix}$$

and

$$\mathbf{H}_1 = \begin{bmatrix} & h[L-1] & \ldots & h[1] \\ & & \ddots & \vdots \\ & & & h[L-1] \\ 0 & & & \end{bmatrix}$$

In this paper, we study the performance of chip-rate equalizers followed by despreading. We consider non causal FIR chip rate (reduced-rank) MMSE equalizers with transfer functions

$g(z) = \sum_{k=-(N-1)}^{N} g_k z^{-k}$, the coefficients of which are designed as if the chip sequence $(d(i))_{i\in\mathbb{Z}}$ were a decorrelated sequence with variance $\frac{K}{N}$. This property is of course not verified because (16) implies that the covariance matrix of $\mathbf{d}(m)$ is rank deficient. The variance $\frac{K}{N}$ is justified by the fact that as $\mathbf{W}(m)^H\mathbf{W}(m) = \mathbf{I}$, then $\mathbb{E}\|\mathbf{d}(m)\|^2 = \mathbb{E}\|\mathbf{b}(m)\|^2 = K$. If $(d(i))_{i\in\mathbb{Z}}$ were an i.i.d. sequence, its variance would therefore be equal to $\frac{K}{N}$. In the following, we collect the coefficients of any of the above equalizers $g(z)$ into the $2N$–dimensional row vector $\mathbf{g} = (g_N, \ldots, g_0, g_{-1}, \ldots, g_{-(N-1)})$. The plain MMSE chip rate equalizer thus corresponds to row vector $\mathbf{g}_{2N}$ given by

$$\mathbf{g}_{2N} = \mathbf{h}^H \left(\mathcal{H}\mathcal{H}^H + \frac{\sigma^2}{K/N}\mathbf{I}\right)^{-1} \tag{18}$$

where $\mathbf{h}$ is the $2N$–dimensional vector defined by $\mathbf{h} = (0, \ldots, 0, h_0, \ldots, h_L, 0, \ldots, 0)^T$ and where $\mathcal{H}$ is the $2N \times 3N$ Sylvester matrix given by

$$\mathcal{H} = \begin{bmatrix} \mathbf{H}_1 & \mathbf{H}_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_1 & \mathbf{H}_0 \end{bmatrix} \tag{19}$$

In the following, we denote by $\mathbf{R}$ the $2N \times 2N$ matrix

$$\mathbf{R} = \mathcal{H}\mathcal{H}^H + \frac{\sigma^2}{K/N}\mathbf{I} \tag{20}$$

and by $\mathbf{K}_n$ the $n \times 2N$ Krylov matrix associated with the pair $(\mathbf{R}, \mathbf{h})$, i.e.

$$\mathbf{K}_n = \left[\mathbf{h}, \mathbf{R}\mathbf{h}, \ldots, \mathbf{R}^{n-1}\mathbf{h}\right]$$

The n-th stage reduced-rank Wiener equalizer corresponds to vector $\mathbf{g}_n$ given by

$$\mathbf{g}_n = \mathbf{h}^H\mathbf{K}_n^H \left(\mathbf{K}_n^H\mathbf{R}\mathbf{K}_n\right)^{-1} \mathbf{K}_n^H \tag{21}$$

We denote by $g_n(z)$ the transfer function associated with vector $\mathbf{g}_n$ and define $\hat{d}_n(i)$ as the corresponding estimated chip sequence $\hat{d}_n(i) = [g_n(z)]y(i)$. We propose to study the effect of n on the performance of the estimator of symbol $b_1(m)$ defined by

$$\hat{b}_{1,n}(m) = \mathbf{w}_1^H(m)\hat{\mathbf{d}}_n(m) \tag{22}$$

where $\hat{\mathbf{d}}_n(m) = (\hat{d}_n(mN), \ldots, \hat{d}_n(mN+N-1))^T$ and where $\mathbf{w}_1(m)$ represents the first column of $\mathbf{W}(m)$, i.e. the code allocated to the users of interest.

V. Asymptotic analysis of reduced-rank equalizers.

From now on, we formulate the following realistic assumption:

Assumption 4: The long code sequence is a realization of a QAM4 i.i.d. sequence.

Therefore, due to the presence of the matrix $\mathbf{S}(m)$, matrix $\mathbf{W}(m)$ can be seen as the realization of a quite particular random matrix. In the following, we study the performance of the above reduced-rank receivers in the asymptotic regime N and K tend to $+\infty$ in such a way that $\frac{K}{N} \to \alpha$ where $0 < \alpha < 1$. For the sake of simplicity, we also assume that the length L of the impulse response of the channel is assumed to be kept constant. However, we conjecture that our results can be extended if L also converges to ∞ in such a way that $L < N$ provided that $\sup_N \sum_{l=0}^{L-1} |h_l| < +\infty$. As the proofs of the main results are more technical in this context, we do not address this case. However, some simulations are given to support this claim.

As $\frac{K}{N} \to \alpha$, we replace factor $\frac{K}{N}$ by α in definition (20) of matrix $\mathbf{R}$ in order to simplify the exposition. This, of course, modifies the expressions of matrices $\mathbf{K}_n$ and of vectors $\mathbf{g}_n$.

In order to characterize the performance of receiver (22), we first evaluate its output SINR. For this, we consider the filter $f_n(z) = \sum_{l=-(N-1)}^{N+L} f_{n,l} z^{-l} = g_n(z)h(z)$, and notice that the estimated chip sequence $\hat{d}_n(i)$ is given by

$$\hat{d}_n(i) = [f_n(z)]d(i) + [g_n(z)]v(i) \tag{23}$$

Vector $\hat{\mathbf{d}}_n(m)$ can thus be written as

$$\hat{\mathbf{d}}_n(m) = \mathcal{F}_n \begin{bmatrix} \mathbf{d}(m-2) \\ \mathbf{d}(m-1) \\ \mathbf{d}(m) \\ \mathbf{d}(m+1) \end{bmatrix} + \mathcal{G}_n \begin{bmatrix} \mathbf{v}(m-1) \\ \mathbf{v}(m) \\ \mathbf{v}(m+1) \end{bmatrix} \tag{24}$$

Here, matrix $\mathcal{G}_n$ is the $N \times 3N$ Sylvester matrix associated with the filter $g_n(z)$, i.e.

$$\mathcal{G}_n = \begin{pmatrix} g_{n,N} & \dots & g_{n,0} & \dots & g_{n,-(N-1)} & 0 & \dots & \dots & 0 \\ 0 & g_{n,N} & \dots & g_{n,0} & \dots & g_{n,-(N-1)} & 0 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & g_{n,N} & \dots & g_{n,0} & \dots & g_{n,-(N-1)} & 0 \end{pmatrix}$$

and $\mathcal{F}_n$ is the $N \times 4N$ Sylvester matrix associated with $f_n(z)$ defined as $\mathcal{G}_n$ from $3N$–dimensional vector $\mathbf{f}_n = (0, \ldots, 0, f_{n,N+L}, \ldots, f_{n,0}, f_{n,-1}, \ldots, f_{n,-(N-1)})^T$. As $f_n(z) = g_n(z)h(z)$, vector $\mathbf{f}_n$ is equal to $\mathbf{g}_n\mathcal{H}$ and matrix $\mathcal{F}_n$ can be written as

$$\mathcal{F}_n = \mathcal{G}_n\mathcal{H}_{3N}$$

where $\mathcal{H}_{3N}$ is the $3N \times 4N$ Sylvester matrix defined in the same way that matrix $\mathcal{H}$ (see eq. (19)). For convenience, we partition $\mathcal{F}_n$ as $\mathcal{F}_n = (\mathbf{F}_{n,2}, \mathbf{F}_{n,1}, \mathbf{F}_{n,0}, \mathbf{F}_{n,-1})$ where the 4 blocks are $N \times N$.

From now on, $\mathbf{c}_1$ represents the first column of matrix $\mathbf{C}$, and $\mathbf{C}$ is partitioned as $\mathbf{C} = (\mathbf{c}_1, \mathbf{C}_2)$.

In order to express the output SINR provided by receiver (22), it is necessary to identify in (22) the contribution of symbol $b_1(m)$, of symbols $(b_j(m))_{j=2,\ldots,K}$ and symbols $(b_j(m-k))_{j=1,\ldots,K,k=-1,1,2}$, and of the noise. After some straightforward calculations, we get that the output SINR at time m, denoted $\tilde{\beta}_n^{(N)}(m)$, is given by

$$\tilde{\beta}_n^{(N)}(m) = \frac{|\mathbf{c}_1^H \mathbf{S}(m)^H \mathbf{F}_{n,0} \mathbf{S}(m)\mathbf{c}_1|^2}{\sum_{k=-1}^{2} T_{n,k} + \sigma^2 \mathbf{c}_1^H \mathbf{S}(m)^H \mathcal{G}_n \mathcal{G}_n^H \mathbf{S}(m)\mathbf{c}_1} \tag{25}$$

where the terms $(T_{n,k})_{k=-1,\ldots,2}$ are defined by

$$T_{n,0} = \mathbf{c}_1^H \mathbf{S}(m)^H \mathbf{F}_{n,0} \mathbf{S}(m) \mathbf{C}_2 \mathbf{C}_2^H \mathbf{S}(m)^H \mathbf{F}_{n,0}^H \mathbf{S}(m)\mathbf{c}_1 \tag{26}$$

$$T_{n,k} = \mathbf{c}_1^H \mathbf{S}(m)^H \mathbf{F}_{n,k} \mathbf{S}(m-k) \mathbf{C}\mathbf{C}^H \mathbf{S}(m-k)^H \mathbf{F}_{n,k}^H \mathbf{S}(m)\mathbf{c}_1 \text{ for } k \neq 0 \tag{27}$$

In order to simplify the notations, the SINR $\tilde{\beta}_N^{(N)}(m)$ of the plain MMSE receiver (i.e. $n = N$) is denoted $\tilde{\beta}^{(N)}(m)$.

The expression (25) is quite complicated, and does not allow to obtain any insight on the performance of the reduced-rank receivers, and in particular on the influence of n on the SINR. We also note that, considered as symbol rate receivers, the chip rate (reduced-rank) Wiener equalizers followed by a despreading **are not Wiener filters** in the classical sense. This explains why $\tilde{\beta}^{(N)}(m)$ and $\tilde{\beta}_n^{(N)}(m)$ are not given by expressions similar to (4) and (7). Therefore, some work is needed in order to be able to use the results of [19].

$\tilde{\beta}_n^{(N)}(m)$ depends on the values of the scrambling code. It can thus be interpreted as a random variable. The key point of this paper is the following result, which states that as N and K converge to $+\infty$ in such a way that $\frac{K}{N} \to \alpha$, then $\tilde{\beta}_n^{(N)}(m)$ has the same behaviour as a certain deterministic quantity which does not depend on the entries of the code matrix $\mathbf{W}(m)$. Neither the particular values of the scrambling code, nor the entries of the Walsh-Hadamard part $\mathbf{C}$ of $\mathbf{W}(m)$ have an influence on this deterministic limit. Before stating our main result, we introduce the following notation:

Definition 1: For each $n \leq N$, we define $\eta_n^{(N)}$ as

$$\eta_n^{(N)} = \mathbf{h}^H \mathbf{K}_n (\mathbf{K}_n^H \mathbf{R} \mathbf{K}_n)^{-1} \mathbf{K}_n^H \mathbf{h} \tag{28}$$

and denote $\eta_N^{(N)} = \mathbf{h}^H \mathbf{R}^{-1} \mathbf{h}$ by $\eta^{(N)}$ in order to simplify the notations.

Remark 1: We notice that the term $\eta_n^{(N)}$ coincides with the coefficient $f_{n,0}$ of the transfer function $f_n(z) = g_n(z)h(z) = \sum_{l=-(N-1)}^{n,N+L} f_{n,l} z^{-l}$. In fact, $f_{n,0}$ is equal to $f_{n,0} = \mathbf{g}_n \mathbf{h}$. The expression (21) provides immediately $f_{n,0} = \eta_n^{(N)}$.

We are now in a position to state the main result of this paper.

Theorem 2: For any fixed n,

$$\lim_{N \to +\infty, K/N \to \alpha} \tilde{\beta}_n^{(N)}(m) - \frac{1}{\alpha} \frac{\eta_n^{(N)}}{(1 - \eta_n^{(N)})} = 0 \tag{29}$$

where the convergence stands for the convergence in probability.

Moreover,

$$\lim_{N \to +\infty, K/N \to \alpha} \left(\tilde{\beta}^{(N)}(m) - \frac{1}{\alpha} \frac{\eta^{(N)}}{(1 - \eta^{(N)})} \right) = 0 \tag{30}$$

The proof of Theorem 2 is somewhat technical, and is given in the appendix.

Remark 2: Expressions at the right hand side of(29) and (30) have a simple interpretation. In fact, it is easy to check that $\frac{1}{\alpha} \frac{\eta_n^{(N)}}{(1-\eta_n^{(N)})}$ coincides with the SINR provided by the rank n Wiener filter $g_n(z)$ (23) *if the chip sequence* $(d(i))_{i \in \mathbb{Z}}$ *in (23) were an i.i.d. sequence of variance* α. (30) can be interpreted similarly. We now explain roughly the reason for which this surprising result holds. We first remark that (24) can be written as

$$\hat{\mathbf{d}}(m) = \eta_n^{(N)} \mathbf{d}(m) + \mathbf{u}(m) \tag{31}$$

where the actual analytical expression of $\mathbf{u}(m)$ is not important at this stage. It is easy to check that if $(d(i))_{i \in \mathbb{Z}}$ were an i.i.d. sequence of variance α, then, the covariance matrix $\Gamma_{u,iid}$ of $\mathbf{u}(m)$

would be Toeplitz with main diagonal term $\alpha\eta_n^{(N)}(1-\eta_n^{(N)})$. This would imply that the SINR associated with each component of $\hat{\mathbf{d}}(m)$ would coincide with $\frac{1}{\alpha}\frac{\eta_n^{(N)}}{(1-\eta_n^{(N)})}$.

However, $(d(i))_{i\in\mathbb{Z}}$ is not an i.i.d. sequence of variance α. In particular, $\mathbf{d}(m)$ is given by $\mathbf{d}(m) = \mathbf{S}(m)\mathbf{C}\mathbf{b}(m)$, and the covariance matrix $\Gamma_u(m)$ of $\mathbf{u}(m)$ (which depends of m via matrix $\mathbf{S}(m)$) does not coincides with $\Gamma_{u,iid}$. Plugging $\mathbf{d}(m) = \mathbf{S}(m)\mathbf{C}\mathbf{b}(m)$ in (31), we get immediately that

$$\hat{b}_1(m) = \mathbf{c}_1^H\mathbf{S}(m)^H\hat{\mathbf{d}}(m) = \eta_n^{(N)}b_1(m) + \mathbf{c}_1^H\mathbf{S}(m)^H\mathbf{u}(m) \tag{32}$$

The variance of the interference + noise term $\mathbf{c}_1^H\mathbf{S}(m)^H\mathbf{u}(m)$ is equal to $\mathbf{c}_1^H\mathbf{S}(m)^H\Gamma_u(m)\mathbf{S}(m)\mathbf{c}_1$. (29) holds because if $N \to +\infty$ and $\frac{K}{N} \to \alpha$, it is possible to replace $\Gamma_u(m)$ by $\Gamma_{u,iid}$ in the variance of the interference + noise term without modifying its asymptotic behaviour. As the components of vector $\mathbf{S}(m)\mathbf{c}_1$ are i.i.d. with variance $\frac{1}{N}$, it is easily seen that $\mathbf{c}_1^H\mathbf{S}(m)^H\Gamma_{u,iid}\mathbf{S}(m)\mathbf{c}_1$ has the same behaviour as $\frac{1}{N}\mathrm{Tr}(\Gamma_{u,iid}) = \alpha\eta_n^{(N)}(1-\eta_n^{(N)})$ (see Lemma 1 in Appendix for more details). This, in turn, explains (29). We stress on the fact that the replacement of $\Gamma_u(m)$ by $\Gamma_{u,iid}$ is not obvious. In the appendix, we implicitly justify this replacement.

Remark 3: It is interesting to notice that (30) coincides with the asymptotic SINR found in [4] in the case where the code matrix $\mathbf{W}$ is obtained by extracting K columns from a Haar distributed random unitary matrix [1]. This is a rather surprising result because our actual code matrix model ((14) and assumption 4) looks very different from a Haar distributed matrix.

Theorem 2 is important in that it allows us to use the results of [19] (see section III) in order to obtain insights on the convergence speed of $\tilde{\beta}_n^{(N)}$ towards $\tilde{\beta}^{(N)}$ when N and K are large enough. Relation (29) implies that it is sufficient to evaluate the convergence speed of $\eta_n^{(N)}$ towards $\eta^{(N)}$ when $N \to +\infty, K/N \to \alpha$, a simpler problem. For this, it is possible to use the results of [19] recalled in section III. Formula (28) coincides with (8) when N is exchanged with $2N$. In our context, matrix $\mathbf{R}$ is $\mathbf{R} = \mathcal{H}\mathcal{H}^H + \frac{\sigma^2}{\alpha}\mathbf{I}$ while vector $\mathbf{h}$ is $\mathbf{h} = (0,\ldots,0,h_0,\ldots,h_L,0,\ldots,0)^T$. We have thus only to check that Assumptions 1, 2, 3 hold.

[1] A random unitary matrix $\mathbf{U}$ is said to be Haar distributed if for each deterministic unitary matrix $\mathbf{Q}$, the distribution of $\mathbf{U}$ coincides with the distribution of $\mathbf{UQ}$

As $\mathbf{R}$ is a Toeplitz matrix associated with the spectral density $|h(e^{2i\pi f})|^2 + \frac{\sigma^2}{\alpha}$, the term $s_k^{(2N)} = \mathbf{h}^H \mathbf{R}^k \mathbf{h}$ defined in Assumption 1 is easily seen to converge towards s_k defined by

$$s_k = \int_0^1 |h(e^{2i\pi f})|^2 (|h(e^{2i\pi f})|^2 + \frac{\sigma^2}{\alpha})^k \, df$$

when $N \to +\infty$. We put $\delta_1 = |h_{min}|^2 + \frac{\sigma^2}{\alpha}$ and $\delta_2 = |h_{max}|^2 + \frac{\sigma^2}{\alpha}$ where $|h_{min}| = \min_f |h(e^{2i\pi f})|$ and $|h_{max}| = \max_f |h(e^{2i\pi f})|$. Then, it is easy to check that s_k can be written as

$$s_k = \int_{\delta_1}^{\delta_2} \lambda^k \, d\nu(\lambda)$$

where ν is the probability measure supported by $[\delta_1, \delta_2]$ defined by

$$\int_{\delta_1}^{\delta_2} \phi(\lambda) \, d\nu(\lambda) = \int_0^1 |h(e^{2i\pi f})|^2 \ \phi\big(|h(e^{2i\pi f})|^2 + \frac{\sigma^2}{\alpha}\big) \, df$$

for each continuous function ϕ. Measure ν is easily seen to be absolutely continuous and to have a strictly positive density on $[\delta_1, \delta_2]$. Thus Assumptions 1 and 2 hold. As for Assumption 3, we remark that as $\mathbf{R}$ is a Toeplitz matrix associated with the spectral density $|h(e^{2i\pi f})|^2 + \frac{\sigma^2}{\alpha}$, then,

$$\begin{aligned} ||\mathbf{R}|| \leq \max_f (|h(e^{2i\pi f})|^2 + \frac{\sigma^2}{\alpha}) &= \delta_2 \\ ||\mathbf{R}^{-1}|| \leq \left(\min_f (|h(e^{2i\pi f})|^2 + \frac{\sigma^2}{\alpha}) \right)^{-1} &= \frac{1}{\delta_1} \end{aligned} \tag{33}$$

Theorem 1 thus shows that $\eta_n^{(N)}$ and $\eta^{(N)}$ converge towards η_n and η defined in section III. Hence, $\beta_n^{(N)}$ and $\beta^{(N)}$ converge towards β_n and β defined by

$$\begin{aligned} \beta_n &= \frac{1}{\alpha} \frac{\eta_n}{1 - \eta_n} \\ \beta &= \frac{1}{\alpha} \frac{\eta}{1 - \eta} \end{aligned} \tag{34}$$

Moreover, the convergence speed of η_n and β_n towards η and β is locally exponential, and the rate of convergence essentially depends on the ratio $\mu = \frac{1 + \frac{\delta_1}{\delta_2}}{1 - \frac{\delta_1}{\delta_2}}$. If μ is close to 1, or equivalently if $\frac{\delta_1}{\delta_2} << 1$, the convergence speed is small. Using standard results on Toeplitz matrices, the smallest and the largest eigenvalue of $\mathbf{R}$ converge to δ_1 and δ_2 respectively when $N \to \infty$. Therefore, the ratio $\frac{\delta_1}{\delta_2}$ is for N large enough nearly equal to the condition number of matrix $\mathbf{R}$. Thus, it can be seen that for N large enough the convergence rate is poor if $\mathbf{R}$ is ill conditioned and vice versa. Our result also allows to evaluate the influence of the load of the cell, i.e. parameter α. Indeed, $\frac{\delta_1}{\delta_2}$ can be written as

$$\frac{\delta_1}{\delta_2} = \frac{\sigma^2 + \alpha|h_{min}|^2}{\sigma^2 + \alpha|h_{max}|^2}$$

Therefore, the smaller α is, the better the convergence rate is.

VI. Simulation results

In this section, we first verify that our asymptotic SINR evaluations allow us to predict the empirical performance of the studied receivers. We have implemented the physical layer of the downlink of the UMTS-FDD and have compared the measured bit error rate with its asymptotic evaluation given by $Q(\sqrt{\beta_n})$. The results are presented in Figure 1. Here, the propagation channel is the so-called Vehicular A. The profile of the Vehicular A channel (i.e. the location of each path with the corresponding path average power) is given in Table 1 (on each frame, a different realization of the channel is generated). Note that the chip period T_c is equal to $T_c = 260nsec$. The Signal to Noise Ratio (for each user) $\frac{E_b}{N_0}$ is equal to 10 dB and the load factor α is equal to $\frac{1}{2}$. Figure 1 shows that our asymptotic evaluations allow to predict rather accurately the performance of the true system even for spreading factors as low as $N = 16$. [2]

Vehicular A channel		Vehicular B channel	
Path Delay (nsec)	Path average Power (dB)	Path Delay (nsec)	Path average Power (dB)
0	0	0	-2.5
310	-1	300	0
710	-9	8900	-12.8
1090	-10	12900	-10
1730	-15	17100	-25.2
2510	-20	20000	-16

Table 1. The Vehicular A and Vehicular B channel profiles.

[2] This means that this asymptotic analysis can be used to study the reduced-rank equalizer raw BER performance in the context of the very recent High-Speed Downlink Packet Access (HSDPA) mode of the UMTS in which many spreading codes of length 16 are allocated to the same user.

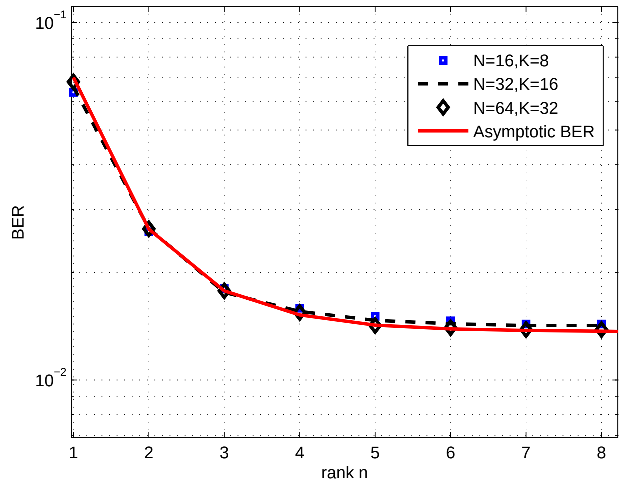

Fig. 1. Comparison of empirical and theoretical BER for the Vehicular A channel.

In section V, we claimed that the results remain valid even for channels with very long delay spread (comparable to N). To verify this, we consider the Vehicular B channel (see Table 1). The delay spread in this case is roughly equal to $80T_c$. We consider the case $N = 128$ and $\alpha = \frac{1}{2}$. The SNR $\frac{E_b}{N_0}$ is equal to 10 dB. The results are given in Figure 2. We notice that the fit is as good as in the Vehicular A case. Thus the results remain valid for channels with delay spreads growing with the spreading factor (provided that $L < N$). As we have verified that β_n and β allow to predict accurately the performance of the above practical system, we next study the influence of various parameters on the convergence speed of β_n towards β. For this, we represent in the following figures the relative SINR defined as the ratio$\frac{\beta_n}{\beta}$ as a function of the rank n. In Figure 3, we first study the influence of α on the convergence speed of the relative SINR towards 1. Here, the propagation channel is the Vehicular A channel, and the ratio $\frac{E_b}{N_0}$ is equal to 7 dB. This figure confirms that the convergence speed of the reduced rank receivers depends crucially on the load factor.

In Figure 4, we study the effect of the channel on the convergence speed of β_n towards β. For this, we consider a 2 taps channel with transfer function $h(z) = h_0 + h_1 z^{-1}$. In this case, the ratio $\frac{\delta_1}{\delta_2}$ is minimum if $|h_0| = |h_1|$ and is equal to $\frac{\sigma^2/\alpha}{2+\sigma^2/\alpha}$: $h(z)$ has a zero on the unit circle, so that $|h_{min}| = 0$, while $|h_{max}| = 2|h_0| = \sqrt{2}$ (because $|h_0|^2 + |h_1|^2 = 1$). Therefore, if $|h_0| = |h_1|$,

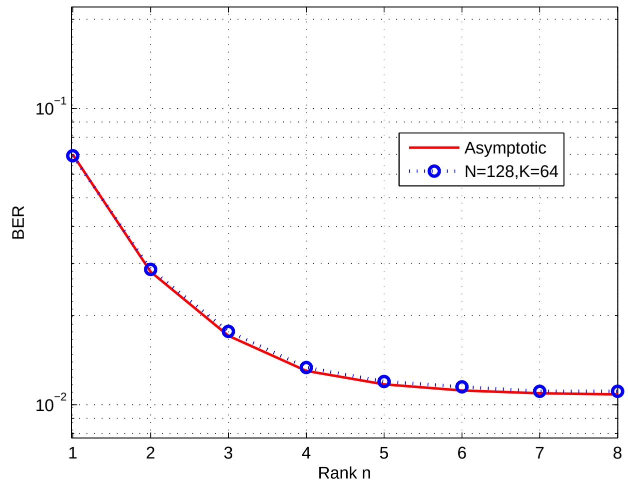

Fig. 2. Comparison of empirical and theoretical BER for the Vehicular B channel

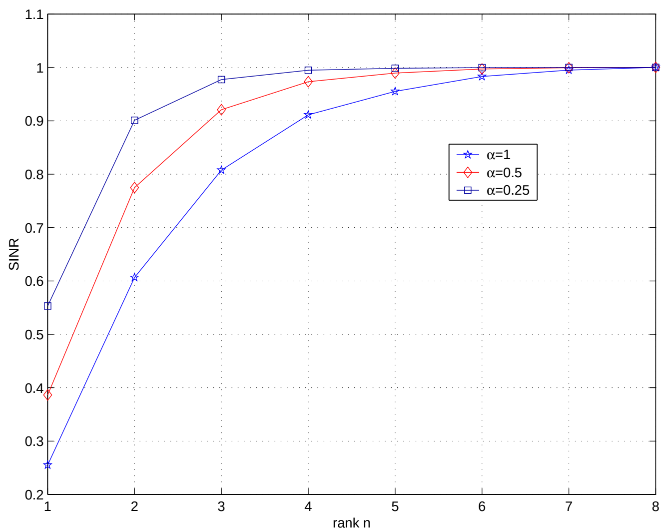

Fig. 3. Influence of α on the convergence of the relative SINR

the convergence speed of β_n towards β is expected to be minimum. This is confirmed by Figure 4 obtained for $\alpha = \frac{1}{2}$ and $\frac{E_b}{N_0} = 7dB$.

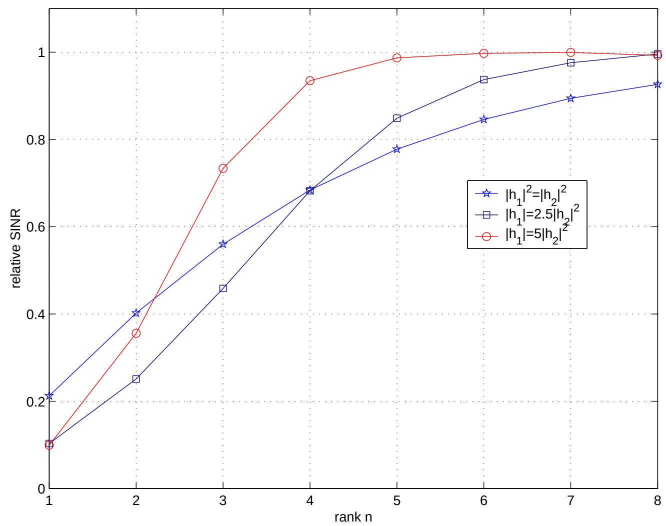

Fig. 4. Influence of the channel on the convergence of the relative SINR

VII. Conclusion

In this paper, we have addressed the performance of downlink CDMA receivers consisting of reduced rank Wiener equalizers followed by despreading. We have studied the convergence speed of their SINR versus their order in the asymptotic regime $N \to +\infty, K/N \to \alpha$. In this context, we have shown that for each n, the SINR provided by the rank n receiver converges to a deterministic term β_n, and that the convergence of β_n when n increases is locally exponential. We have evaluated the corresponding rate which only depends on the condition number of the covariance matrix to be inverted in order to calculate the full rank receiver. Simulation results have shown that our asymptotic results allow to predict the performance of finite dimension CDMA system even for very short spreading factors.

Appendix I
Outline of the proof.

The proof of Theorem 2 is quite technical. In order to improve its readability, we first outline the main steps of (29) and provide more details in the next sections of the appendix. We finally briefly justify (30).

In order to study the asymptotic behaviour of $\tilde{\beta}_n^{(N)}(m)$, it is necessary to study separately the various terms of the right hand side of (25).

First step: study of $|\mathbf{c}_1^H \mathbf{S}(m)^H \mathbf{F}_{n,0} \mathbf{S}(m)\mathbf{c}_1|^2$ and $\mathbf{c}_1^H \mathbf{S}(m)^H \mathcal{G}_n \mathcal{G}_n^H \mathbf{S}(m)\mathbf{c}_1$.

The above terms can be studied by using the following useful lemma.

Lemma 1: Let $\mathbf{B}_N$ be a deterministic $N \times N$ uniformly bounded matrix, that is $\sup_N \|\mathbf{B}_N\| < +\infty$. Then,

$$\lim_{N \to +\infty} \mathbb{E} \left| \mathbf{c}_1^H \mathbf{S}(m)^H \mathbf{B}_N \mathbf{S}(m) \mathbf{c}_1 - \frac{1}{N} \mathrm{Trace}(\mathbf{B}_N) \right|^2 = 0 \tag{35}$$

This result is an immediate consequence of a classical result extensively used in the performance evaluation of large dimension communication system (see [24]).

In order to be able to use Lemma 1, we need to verify that matrices $\mathbf{F}_{n,0}$ and $\mathcal{G}_n \mathcal{G}_n^H$, or equivalently $\mathcal{G}_n$, are uniformly bounded.

Lemma 2: For each n fixed, matrix $\mathcal{G}_n$ is uniformly bounded, i.e. $\sup_N \|\mathcal{G}_n\| < +\infty$.

The proof is given in Appendix II. Matrix $\mathcal{F}_n$ is given by $\mathcal{F}_n = \mathcal{G}_n \mathcal{H}_{3N}$. Matrix $\mathcal{H}_{3N}$ is a Toeplitz matrix associated with the filter $h(z) = \sum_{l=0}^{L} h_l z^{-L}$. Therefore, for each N, $\|\mathcal{H}_{3N}\| \leq \|h\|_\infty = \sup_f |h(e^{2i\pi f})|$. This shows that $\mathcal{H}_{3N}$ is uniformly bounded. As $\|\mathcal{F}_n\| \leq \|\mathcal{G}_n\| \|\mathcal{H}_{3N}\|$, Lemma 2 implies that $\mathcal{F}_n$, and thus matrices $(\mathbf{F}_{n,k})_{k=-1,\ldots,2}$ are uniformly bounded.

Lemma 1 and the above discussion imply the following corollary:

Corollary 1:

$$\mathbf{c}_1^H \mathbf{S}(m)^H \mathbf{F}_{n,0} \mathbf{S}(m) \mathbf{c}_1 - \eta_n^{(N)} \quad \to \quad 0 \tag{36}$$

$$\mathbf{c}_1^H \mathbf{S}(m)^H \mathcal{G}_n \mathcal{G}_n^H \mathbf{S}(m) \mathbf{c}_1 - \|\mathbf{g}_n\|^2 \quad \to \quad 0 \tag{37}$$

where the convergence stands for the convergence in probability.

Proof. In order to prove the first statement of Corollary 1, we remark that Lemma 1 and $\mathbf{F}_{n,0}$ uniformly bounded imply that

$$\mathbf{c}_1^H \mathbf{S}(m)^H \mathbf{F}_{n,0} \mathbf{S}(m) \mathbf{c}_1 - \frac{1}{N} \mathrm{Trace}(\mathbf{F}_{n,0})$$

converges in the mean-square sense, and thus in probability, to 0. As $\mathbf{F}_{n,0}$ is a Toeplitz matrix, its normalized trace coincides with the constant term $f_{n,0}$ of transfer function $f_n(z) = g_n(z)h(z)$, which is equal to $f_{n,0} = \mathbf{g}_n \mathbf{h} = \eta_n^{(N)}$. The second statement of Corollary 1 follows directly from

Lemmas 1 and 2 and from the observation that $\frac{1}{N}\text{Trace}(\mathcal{G}_n\mathcal{G}_n^H) = \|\mathbf{g}_n\|^2$.

Second step: study of $T_{n,0}$.

The asymptotic behaviour of $T_{n,0} = \mathbf{c}_1^H\mathbf{S}(m)^H\mathbf{F}_{n,0}\mathbf{S}(m)\mathbf{C}_2\mathbf{C}_2^H\mathbf{S}(m)^H\mathbf{F}_{n,0}^H\mathbf{S}(m)\mathbf{c}_1$ is a straightforward consequence of the following Lemma.

Lemma 3: Let $\mathbf{B}_N$ be a $N \times N$ uniformly bounded Toeplitz matrix, i.e. $\sup_N \|\mathbf{B}_N\| < +\infty$. Then,

$$\lim_{N\to+\infty, \frac{K}{N}\to\alpha} \mathbb{E}\left|\mathbf{c}_1^H\mathbf{S}(m)^H\mathbf{B}_N\mathbf{S}(m)\mathbf{C}_2\mathbf{C}_2^H\mathbf{S}(m)^H\mathbf{B}_N^H\mathbf{S}(m)\mathbf{c}_1 - \alpha\left(\frac{1}{N}\text{Trace}(\mathbf{B}_N\mathbf{B}_N^H) - |\frac{1}{N}\text{Trace}(\mathbf{B}_N)|^2\right)\right|^2 = 0 \tag{38}$$

Proof. See Appendix III.

Lemma 2 implies that matrix $\mathbf{F}_{n,0}$ is uniformly bounded. As the mean-square convergence implies the convergence in probability, Lemma 3 shows that $T_{n,0}$ converges in probability to $\alpha\left(\frac{1}{N}\text{Trace}(\mathbf{F}_{n,0}\mathbf{F}_{n,0}^H) - |\frac{1}{N}\text{Trace}(\mathbf{F}_{n,0})|^2\right)$. As $\frac{1}{N}\text{Trace}(\mathbf{F}_{n,0}) = \eta_n^{(N)}$, we get immediately the following Corollary.

Corollary 2:

$$T_{n,0} \to \alpha\left(\frac{1}{N}(\text{Trace}(\mathbf{F}_{n,0}\mathbf{F}_{n,0}^H) - (\eta_n^{(N)})^2\right) \tag{39}$$

where the convergence stands for the convergence in probability.

Third step: study of $T_{n,k}$ for $k = -1, 1, 2$.

The following lemma allows to precise the behaviour of $T_{n,k}$ for $k = -1, 1, 2$.

Lemma 4: Let $\mathbf{B}_N$ be a uniformly bounded $N \times N$ matrix. Then, for $k = -1, 1, 2$,

$$\lim_{N\to+\infty, \frac{K}{N}\to\alpha} \mathbb{E}\left|\mathbf{c}_1^H\mathbf{S}(m)^H\mathbf{B}_N\mathbf{S}(m-k)\mathbf{C}\mathbf{C}^H\mathbf{S}(m-k)^H\mathbf{B}_N^H\mathbf{S}(m)\mathbf{c}_1 - \alpha\frac{1}{N}\text{Trace}(\mathbf{B}_N\mathbf{B}_N^H)\right|^2 = 0 \tag{40}$$

Proof. See Appendix IV for a sketch of the proof.

Lemma 2 implies that matrices $\mathbf{F}_{n,k}$ are bounded. As the mean-square convergence implies the convergence in probability, Lemma 4 shows that

$$T_{n,k} \to \alpha \frac{1}{N} \text{Trace}(\mathbf{F}_{n,k}\mathbf{F}_{n,k}^H) \tag{41}$$

where the convergence stands for the convergence in probability.

Fourth step: proof of (29)

We are now in position to complete the proof of (29). From the above discussions, we get that

$$\tilde{\beta}_n^{(N)} - \frac{(\eta_n^{(N)})^2}{\alpha\left(\sum_{k=-1}^{2} \frac{1}{N}\text{Trace}(\mathbf{F}_{n,k}\mathbf{F}_{n,k}^H) - (\eta_n^{(N)})^2\right) + \sigma^2\|\mathbf{g}_n\|^2}$$

converges to 0 in probability. We remark that

$$\sum_{k=-1}^{2} \frac{1}{N}\text{Trace}(\mathbf{F}_{n,k}\mathbf{F}_{n,k}^H) = \frac{1}{N}\text{Trace}(\mathcal{F}_n\mathcal{F}_n^H)$$

As $\mathcal{F}_n\mathcal{F}_n^H$ is a $N \times N$ Toeplitz matrix, its normalized trace coincides with its diagonal term which is equal to $\|\mathbf{f}_n\|^2$. As $\mathbf{f}_n = \mathbf{g}_n\mathcal{H}$, we get that

$$\frac{1}{N}\text{Trace}(\mathcal{F}_n\mathcal{F}_n^H) = \mathbf{g}_n\mathcal{H}\mathcal{H}^H\mathbf{g}_n^H$$

and that

$$\alpha \sum_{k=-1}^{2} \frac{1}{N}\text{Trace}(\mathbf{F}_{n,k}\mathbf{F}_{n,k}^H) + \sigma^2\|\mathbf{g}_n\|^2 = \alpha\mathbf{g}_n\left(\mathcal{H}\mathcal{H}^H + \frac{\sigma^2}{\alpha}\mathbf{I}_{2N}\right)\mathbf{g}_n^H = \alpha\mathbf{g}_n\mathbf{R}\mathbf{g}_n^H$$

But, as $\mathbf{g}_n$ is given by (21), $\mathbf{g}_n\mathbf{R}\mathbf{g}_n^H$ coincides with $\mathbf{h}^H\mathbf{K}_n^H\left(\mathbf{K}_n\mathbf{R}\mathbf{K}_n^H\right)^{-1}\mathbf{K}_n\mathbf{h}$, i.e. with $\eta_n^{(N)}$. Putting all pieces together, we get that

$$\frac{(\eta_n^{(N)})^2}{\alpha\left(\sum_{k=-1}^{2} \frac{1}{N}\text{Trace}(\mathbf{F}_{n,k}\mathbf{F}_{n,k}^H) - (\eta_n^{(N)})^2\right) + \sigma^2\|\mathbf{g}_n\|^2} = \frac{1}{\alpha}\frac{\eta_n^{(N)}}{1-\eta_n^{(N)}}$$

which, eventually, proves (29).

We finally justify (30). For this, we just mention that, as the full rank Wiener filter $g_N(z)$ converges when $N \to +\infty$ to the usual non causal filter $g_\infty(z) = \frac{h^*(z^{-1}}{h(z)h^*(z^{-1})+\frac{\sigma^2}{\alpha}}$, which verifies $\|g_\infty\|_\infty < +\infty$, then $\sup_N \|g_N\|_\infty < +\infty$. Therefore, matrices $\mathcal{G}_N$ and $\mathcal{F}_N$ are uniformly bounded. The reader may check that this allows to generalize the above arguments to the case where $n = N$.

APPENDIX II
PROOF OF LEMMA 2

We show that $\sup_N \|\mathcal{G}_n\| < +\infty$. For this, we note that matrix $\mathcal{G}_n$ is a Toeplitz matrix associated with the transfer function $g_n(z)$. Therefore, for each N, $\|\mathcal{G}_n\| < \|g_n\|_\infty = \sup_f |g_n(e^{2i\pi f})|$. Hence,

$$\sup_N \|\mathcal{G}_n\| < \sup_N \|g_n\|_\infty$$

We now prove that $\sup_N \|g_n\|_\infty < +\infty$. As $h(z)$ is a degree L FIR filter, we claim that if N is large enough, then the number of non zero coefficients of $g_n(z)$ is less than $(2n-1)L$, and thus remains finite when $N \to +\infty$. In fact, row vector $\mathbf{g}_n$ is a linear combination of the rows $(\mathbf{h}^H, \mathbf{h}^H\mathbf{R}, \ldots, \mathbf{h}^H\mathbf{R}^{n-1})$ of matrix $\mathbf{K}_n^H$. If N is large enough, for each $1 \leq k \leq (n-1)$, $\mathbf{R}^k$ is a band matrix whose entries $(\mathbf{R}^k)_{i,j}$ are zero if $|i-j| > kL$. It is therefore easy to check that components 1 to $N-kL-1$ and $N+(k+1)L+1$ to $2N$ of vector $\mathbf{h}^H\mathbf{R}^k$ are zero. This implies that components 1 to $N-(n-1)L-1$ and $N+nL+1$ to $2N$ of any linear combination of the rows of $\mathbf{K}_n^H$ are zero if N is large enough. In order to establish that $\sup_N \|g_n\|_\infty < +\infty$, it is therefore sufficient to show that the Euclidean norm $\|\mathbf{g}_n\|$ of vector $\mathbf{g}_n$ remains bounded when N increases. For this, we remark that

$$\|\mathbf{g}_n\|^2 = \mathbf{h}^H\mathbf{K}_n(\mathbf{K}_n^H\mathbf{R}\mathbf{K}_n)^{-1}\mathbf{K}_n^H\mathbf{K}_n(\mathbf{K}_n^H\mathbf{R}\mathbf{K}_n)^{-1}\mathbf{K}_n^H\mathbf{h}.$$

As $\mathbf{R} \geq \frac{\sigma^2}{\alpha}\mathbf{I}$, it is clear that $(\mathbf{K}_n^H\mathbf{R}\mathbf{K}_n)^{-1} \leq \alpha\sigma^{-2}(\mathbf{K}_n^H\mathbf{K}_n)^{-1}$, and that $\mathbf{K}_n(\mathbf{K}_n^H\mathbf{R}\mathbf{K}_n)^{-1}\mathbf{K}_n^H \leq \alpha\sigma^{-2}\mathbf{K}_n(\mathbf{K}_n^H\mathbf{K}_n)^{-1}\mathbf{K}_n^H$, which is itself less than $\alpha\sigma^{-2}\mathbf{I}$. This, in turn, shows that $\|\mathbf{g}_n\|^2 \leq \frac{\alpha\|\mathbf{h}\|^2}{\sigma^2}$, and that the norm $\|\mathbf{g}_n\|$ remains bounded when N increases.

APPENDIX III
PROOF OF LEMMA 3.

The proof of Lemma 3 needs some work. In order to make the proof easier to follow, we simplify the notations: As the parameter m is irrelevant here, $\mathbf{S}(m)$ is denoted $\mathbf{S}$. Matrix $\mathbf{B}_N$ is denoted $\mathbf{B}$. We denote by b_0 the diagonal term of $\mathbf{B}$ (we recall that $\mathbf{B}$ is Toeplitz), and put $\mathbf{A} = \mathbf{B} - b_0\mathbf{I}$ and

$$T_N = \mathbf{c}_1^H\mathbf{S}^H\mathbf{B}\mathbf{S}\mathbf{C}_2\mathbf{C}_2^H\mathbf{S}^H\mathbf{B}^H\mathbf{S}\mathbf{c}_1$$

We remark that, as the entries of matrix $\mathbf{C}$ are equal to $\pm\frac{1}{\sqrt{N}}$, then, the diagonal entries of $\mathbf{C}_2\mathbf{C}_2^H$ are equal to $\frac{K-1}{N}$. We denote by $\mathbf{D}$ the matrix $\mathbf{D} = \mathbf{C}_2\mathbf{C}_2^H - \frac{K-1}{N}\mathbf{I}$. The diagonal entries of $\mathbf{A}$ and $\mathbf{D}$ are of course zero. The main steps of the proof of Lemma 3 T_N are the following:

First step. Observe that T_N can be written as

$$T_N = \frac{K-1}{N}\mathbf{c}_1^H\mathbf{S}^H\mathbf{A}\mathbf{A}^H\mathbf{S}\mathbf{c}_1 + \mathbf{c}_1^H\mathbf{S}^H\mathbf{A}\mathbf{S}\mathbf{D}\mathbf{S}^H\mathbf{A}^H\mathbf{S}\mathbf{c}_1 \tag{42}$$

Proof of the first step. T_N is given by

$$T_N = \mathbf{c}_1^H\mathbf{S}^H(\mathbf{A}+b_0\mathbf{I})\mathbf{S}\mathbf{C}_2\mathbf{C}_2^H\mathbf{S}^H(\mathbf{A}+b_0\mathbf{I})^H\mathbf{S}\mathbf{c}_1$$

As $\mathbf{c}_1^H\mathbf{C}_2 = 0$ and $\mathbf{S}$ is unitary, this reduces to

$$T_N = \mathbf{c}_1^H\mathbf{S}^H\mathbf{A}\mathbf{S}\mathbf{C}_2\mathbf{C}_2^H\mathbf{S}^H\mathbf{A}^H\mathbf{S}\mathbf{c}_1$$

Writing $\mathbf{C}_2\mathbf{C}_2^H$ as $\mathbf{D} + \frac{K-1}{N}\mathbf{I}$, we get immediately that T_N is given by (42).

Second step. Establish that

$$\lim_{N\to+\infty, \frac{K}{N}\to\alpha} \frac{K-1}{N}\mathbf{c}_1^H\mathbf{S}^H\mathbf{A}\mathbf{A}^H\mathbf{S}\mathbf{c}_1 - \alpha\left(\frac{1}{N}\text{Trace}(\mathbf{B}\mathbf{B}^H) - \left|\frac{1}{N}\text{Trace}(\mathbf{B})\right|^2\right) = 0 \tag{43}$$

where the convergence stands for the convergence in probability.

Proof of step 2. $\mathbf{B}$ uniformly bounded implies that $\mathbf{A}\mathbf{A}^H$ is uniformly bounded. Therefore, Lemma 1 implies that $\frac{K-1}{N}\mathbf{c}_1^H\mathbf{S}^H\mathbf{A}\mathbf{A}^H\mathbf{S}\mathbf{c}_1$ converges in quadratic mean to $\alpha\frac{1}{N}\text{Trace}(\mathbf{A}\mathbf{A}^H)$. But, it is easy to check that

$$\frac{1}{N}\text{Trace}(\mathbf{A}\mathbf{A}^H) = \frac{1}{N}\text{Trace}(\mathbf{B}\mathbf{B}^H) - \left|\frac{1}{N}\text{Trace}(\mathbf{B})\right|^2$$

(43) thus follows from the fact that $\frac{K-1}{N}$ converges to α.

Third step. Establish that $\epsilon_N = \mathbf{c}_1^H\mathbf{S}^H\mathbf{A}\mathbf{S}\mathbf{D}\mathbf{S}^H\mathbf{A}^H\mathbf{S}\mathbf{c}_1$ converges in the least-squares sense to 0, i.e. that

$$\lim_{N\to+\infty, \frac{K}{N}\to\alpha} \mathbb{E}(\epsilon_N^2) = 0 \tag{44}$$

because ϵ_N is real.

Proof of step 3. In order to establish (44), we express $\mathbb{E}(\epsilon_N^2)$ by taking benefit that the entries $(s_i)_{i=1,\ldots,N}$ of $\mathbf{S}$ are independent QAM4 sequences and that the diagonal entries of $\mathbf{D}$ and $\mathbf{A}$

are zero. A straightforward, but quite tedious, analysis of the various terms of the corresponding expression gives (44). ϵ_N can be written as

$$\epsilon_N = \sum_{i_1,j_1,i_2,j_2} \mathbf{c}_{i_1,1} s_{i_1}^* \mathbf{A}_{i_1,j_1} s_{j_1} \mathbf{D}_{j_1,i_2} s_{i_2}^* (\mathbf{A}^H)_{i_2,j_2} s_{j_2} c_{j_2,1}$$

Hence, $\mathbb{E}(\epsilon_N^2)$ is equal to

$$\sum_{(i_1,i_2,i_3,i_4),(j_1,j_2,j_3,j_4)} \mathbf{c}_{i_1,1} \mathbf{A}_{i_1,j_1} \mathbf{D}_{j_1,i_2} (\mathbf{A}^H)_{i_2,j_2} \mathbf{c}_{j_2,1} \mathbf{c}_{i_3,1} \mathbf{A}_{i_3,j_3} \mathbf{D}_{j_3,i_4} (\mathbf{A}^H)_{i_4,j_4} \mathbf{c}_{j_4,1} \mathbb{E}(s_{i_1}^* s_{j_1} s_{i_2}^* s_{j_2} s_{i_3}^* s_{j_3} s_{i_4}^* s_{j_4})$$

As $(s_i)_{i=1,\dots,N}$ is an independent QAM4 sequence, the term $\mathbb{E}(s_{i_1}^* s_{j_1} s_{i_2}^* s_{j_2} s_{i_3}^* s_{j_3} s_{i_4}^* s_{j_4})$ is non zero if and only if it exists a permutation π (depending on the multi-index (i_1, i_2, i_3, i_4)) from the set $\{1, 2, 3, 4\}$ for which $j_k = i_{\pi(k)}$ for each $k \in \{1, 2, 3, 4\}$. In this case, $\mathbb{E}(s_{i_1}^* s_{j_1} s_{i_2}^* s_{j_2} s_{i_3}^* s_{j_3} s_{i_4}^* s_{j_4})$ is equal to 1. As the diagonal entries of $\mathbf{A}$ and $\mathbf{D}$ are zero, coeffi cient

$$\mathbf{c}_{i_1,1} \mathbf{A}_{i_1,j_1} \mathbf{D}_{j_1,i_2} (\mathbf{A}^H)_{i_2,j_2} \mathbf{c}_{j_2,1} \mathbf{c}_{i_3,1} \mathbf{A}_{i_3,j_3} \mathbf{D}_{j_3,i_4} (\mathbf{A}^H)_{i_4,j_4} \mathbf{c}_{j_4,1}$$

is possibly non zero only if $j_k \neq i_k$ for $k \in \{1, 2, 3, 4\}$ and $j_{k-1} \neq i_k$ for $k \in \{2, 4\}$, that is if

$$\pi(1) \neq 1, \pi(1) \neq 2, \pi(2) \neq 2, \pi(3) \neq 3, \pi(3) \neq 4, \pi(4) \neq 4$$

Therefore, a permutation π corresponds to a possibly non zero term if

$$\pi(1) \in \{3, 4\}, \pi(2) \in \{1, 3, 4\}, \pi(3) \in \{1, 2\}, \pi(4) \in \{1, 2, 3\}$$

This corresponds to the following 5 possible permutations:

- $\pi(1) = 3, \pi(3) = 1, \pi(2) = 4, \pi(4) = 2$, permutation π_1,
- $\pi(1) = 3, \pi(3) = 2, \pi(2) = 4, \pi(4) = 1$, permutation π_2,
- $\pi(1) = 4, \pi(3) = 1, \pi(2) = 3, \pi(4) = 2$, permutation π_3,
- $\pi(1) = 4, \pi(3) = 2, \pi(2) = 1, \pi(4) = 3$, permutation π_4,
- $\pi(1) = 4, \pi(3) = 2, \pi(2) = 3, \pi(4) = 1$, permutation π_5.

In the following, we denote by $\mathbf{i} = (i_1, i_2, i_3, i_4)$ a four-dimensional multi-index, and for each $k = 1, \dots, 5$, by $\pi_k(\mathbf{i})$ the multi-index $(i_{\pi_k(1)}, i_{\pi_k(2)}, i_{\pi_k(3)}, i_{\pi_k(4)})$. We will show below that for each $k = 1, 2, \dots, 5$, then,

$$\sum_{\mathbf{i}} \sum_{\mathbf{j}=\pi_k(\mathbf{i})} \mathbf{c}_{i_1,1} \mathbf{A}_{i_1,j_1} \mathbf{D}_{j_1,i_2} (\mathbf{A}^H)_{i_2,j_2} \mathbf{c}_{j_2,1} \mathbf{c}_{i_3,1} \mathbf{A}_{i_3,j_3} \mathbf{D}_{j_3,i_4} (\mathbf{A}^H)_{i_4,j_4} \mathbf{c}_{j_4,1} \to 0 \quad (45)$$

Unfortunately, this does not show that $\mathbb{E}(\epsilon_N^2)$ converges to 0 because

$$\mathbb{E}(\epsilon_N^2) \neq \sum_{k=1}^{5} \sum_{\mathbf{i}} \sum_{\mathbf{j}=\pi_k(\mathbf{i})} \mathbf{c}_{i_1,1}\mathbf{A}_{i_1,j_1}\mathbf{D}_{j_1,i_2}(\mathbf{A}^H)_{i_2,j_2}\mathbf{c}_{j_2,1}\mathbf{c}_{i_3,1}\mathbf{A}_{i_3,j_3}\mathbf{D}_{j_3,i_4}(\mathbf{A}^H)_{i_4,j_4}\mathbf{c}_{j_4,1} \tag{46}$$

This is because, for certain multi indices $\mathbf{i}$ having identical components, it may exist $k \neq l$ for which $\pi_k(\mathbf{i}) = \pi_l(\mathbf{i})$. For example, if $i_1 = i_2$, then $\pi_1(\mathbf{i}) = \pi_2(\mathbf{i})$. These multi indices are thus taken into account at least two times in the right hand side of equation (46). In order to show that $\mathbb{E}(\epsilon_N^2)$ converges towards 0, the reader may check that it is suffi cient to prove (45) for $k = 1, \ldots, 5$, as well as (45) but in which the summation over $\mathbf{i}$ is restricted to indices for which $(i_1 = i_2)$, $(i_1 = i_3)$, $(i_3 = i_4)$, $(i_1 = i_2)$ and $(i_3 = i_4)$.

We now prove (45) for $k = 1$, i.e. that

$$\sum_{\mathbf{i}} \mathbf{c}_{i_1,1}\mathbf{A}_{i_1,i_3}\mathbf{D}_{i_3,i_2}(\mathbf{A}^H)_{i_2,i_4}\mathbf{c}_{i_4,1}\mathbf{c}_{i_3,1}\mathbf{A}_{i_3,i_1}\mathbf{D}_{i_1,i_4}(\mathbf{A}^H)_{i_4,i_2}\mathbf{c}_{i_2,1} \rightarrow 0 \tag{47}$$

For this, we replace $\mathbf{D}$ by $\mathbf{C}_2\mathbf{C}_2^H - \frac{K-1}{N}\mathbf{I}$, and verify that

$$\sum_{\mathbf{i}} \mathbf{c}_{i_1,1}\mathbf{A}_{i_1,i_3}(\mathbf{C}_2\mathbf{C}_2^H)_{i_3,i_2}(\mathbf{A}^H)_{i_2,i_4}\mathbf{c}_{i_4,1}\mathbf{c}_{i_3,1}\mathbf{A}_{i_3,i_1}(\mathbf{C}_2\mathbf{C}_2^H)_{i_1,i_4}(\mathbf{A}^H)_{i_4,i_2}\mathbf{c}_{i_2,1} \rightarrow 0, \tag{48}$$

and

$$\begin{array}{l} \sum_{\mathbf{i}} \mathbf{c}_{i_1,1}\mathbf{A}_{i_1,i_3}\delta_{i_3-i_2}(\mathbf{A}^H)_{i_2,i_4}\mathbf{c}_{i_4,1}\mathbf{c}_{i_3,1}\mathbf{A}_{i_3,i_1}(\mathbf{C}_2\mathbf{C}_2^H)_{i_1,i_4}(\mathbf{A}^H)_{i_4,i_2}\mathbf{c}_{i_2,1} \rightarrow 0 \\ \sum_{\mathbf{i}} \mathbf{c}_{i_1,1}\mathbf{A}_{i_1,i_3}(\mathbf{C}_2\mathbf{C}_2^H)_{i_3,i_2}(\mathbf{A}^H)_{i_2,i_4}\mathbf{c}_{i_4,1}\mathbf{c}_{i_3,1}\mathbf{A}_{i_3,i_1}\delta_{i_1-i_4}(\mathbf{A}^H)_{i_4,i_2}\mathbf{c}_{i_2,1} \rightarrow 0 \end{array} \tag{49}$$

as well as

$$\sum_{\mathbf{i}} \mathbf{c}_{i_1,1}\mathbf{A}_{i_1,i_3}\delta_{i_3-i_2}(\mathbf{A}^H)_{i_2,i_4}\mathbf{c}_{i_4,1}\mathbf{c}_{i_3,1}\mathbf{A}_{i_3,i_1}\delta_{i_1-i_4}(\mathbf{A}^H)_{i_4,i_2}\mathbf{c}_{i_2,1} \rightarrow 0 \tag{50}$$

We fi rst check (48). We recall that matrix $(\mathbf{C}_{i,k})_{i=1,\ldots,N,k=1,\ldots,K}$ is obtained by extracting K columns from a $N \times N$ (unitary) Walsh-Hadamard matrix. In order to simplify the notations, we denote by $(\mathbf{c}_k)_{k=1,\ldots,N}$ the columns of this unitary matrix, and by $(\mathbf{c}_{i,k})_{i=1,\ldots,N}$ the components of vector $\mathbf{c}_k$. In particular, matrix $\mathbf{C}_2$ is equal to $\mathbf{C}_2 = (\mathbf{c}_2, \ldots, \mathbf{c}_K)$. The term to be studied, denoted $u_{1,N}$, is equal to

$$u_{1,N} = \sum_{k=2}^{K} \sum_{l=2}^{K} \sum_{(i_1,i_2,i_3,i_4)} \mathbf{c}_{i_1,1}\mathbf{A}_{i_1,i_3}\mathbf{c}_{i_3,k}\mathbf{c}_{i_2,k}(\mathbf{A}^H)_{i_2,i_4}\mathbf{c}_{i_4,1}\mathbf{c}_{i_3,1}\mathbf{A}_{i_3,i_1}\mathbf{c}_{i_1,l}\mathbf{c}_{i_4,l}(\mathbf{A}^H)_{i_4,i_2}\mathbf{c}_{i_2,1}$$

It can also be written as

$$u_{1,N} = \sum_{k=2}^{K} \sum_{l=2}^{K} \left| \sum_{i_1,i_3} \mathbf{c}_{i_1,1}\mathbf{c}_{i_3,1}\mathbf{c}_{i_1,l}\mathbf{c}_{i_3,k}\mathbf{A}_{i_1,i_3}\mathbf{A}_{i_3,i_1} \right|^2$$

It is clear that $u_{1,N}$ is smaller than the term $v_{1,N}$ defined by

$$v_{1,N} = \sum_{k=1}^{N} \sum_{l=1}^{N} \left| \sum_{i_1,i_3} \mathbf{c}_{i_1,1}\mathbf{c}_{i_3,1}\mathbf{c}_{i_1,l}\mathbf{c}_{i_3,k}\mathbf{A}_{i_1,i_3}\mathbf{A}_{i_3,i_1} \right|^2$$

$v_{1,N}$ is equal to

$$v_{1,N} = \sum_{\mathbf{i}} \sum_{k=1}^{N} \sum_{l=1}^{N} \mathbf{c}_{i_1,1}\mathbf{c}_{i_3,1}\mathbf{c}_{i_2,1}\mathbf{c}_{i_4,1}\mathbf{c}_{i_1,l}\mathbf{c}_{i_2,l}\mathbf{c}_{i_3,k}\mathbf{c}_{i_4,k}\mathbf{A}_{i_1,i_3}\mathbf{A}_{i_3,i_1}\mathbf{A}^*_{i_2,i_4}\mathbf{A}^*_{i_4,i_2}$$

As $\sum_{l=1}^{N} \mathbf{c}_{i_1,l}\mathbf{c}_{i_2,l} = \delta_{i_1-i_2}$ and $\sum_{k=1}^{N} \mathbf{c}_{i_3,k}\mathbf{c}_{i_4,k} = \delta_{i_3-i_4}$, we get that

$$v_{1,N} = \sum_{i_1,i_3} (\mathbf{c}_{i_1,1})^2(\mathbf{c}_{i_3,1})^2|\mathbf{A}_{i_1,i_3}|^2|\mathbf{A}_{i_3,i_1}|^2 = \frac{1}{N^2}\sum_{i_1,i_3} |\mathbf{A}_{i_1,i_3}|^2|\mathbf{A}_{i_3,i_1}|^2$$

because the entries of $\mathbf{C}$ are equal to $\pm\frac{1}{\sqrt{N}}$. We finally show that $u_{1,N} \rightarrow 0$, which in turn, implies that $u_{1,N} \rightarrow 0$. For this, we have to check that $\frac{1}{N}\sum_{i_1,i_3} |\mathbf{A}_{i_1,i_3}|^2|\mathbf{A}_{i_3,i_1}|^2$ is bounded. If $\mathbf{E}$ and $\mathbf{F}$ are $N \times N$ matrices, we denote by $\mathbf{E} \bullet \mathbf{F}$ the Schur-Hadamard product of $\mathbf{E}$ and $\mathbf{F}$ defined by $(\mathbf{E} \bullet \mathbf{F})_{k,l} = \mathbf{E}_{k,l}\mathbf{F}_{k,l}$. It is easily seen that $\|\mathbf{E} \bullet \mathbf{F}\| \leq \|\mathbf{E}\|\|\mathbf{F}\|$. We remark that

$$\frac{1}{N}\sum_{i_1,i_3} |\mathbf{A}_{i_1,i_3}|^2|\mathbf{A}_{i_3,i_1}|^2 = \frac{1}{N}\text{Trace}(\mathbf{A} \bullet \mathbf{A}^T)(\mathbf{A} \bullet \mathbf{A}^T)^H$$

and is thus upper bounded by $\|\mathbf{A} \bullet \mathbf{A}^T\|^2 \leq \|\mathbf{A}\|^4$. As $\mathbf{A}$ is uniformly bounded,

$$\sup_N \frac{1}{N}\text{Trace}(\mathbf{A} \bullet \mathbf{A}^T)(\mathbf{A} \bullet \mathbf{A}^T)^H < +\infty$$

This shows that $v_{1,N}$, and thus $u_{1,N}$ converges to 0.

We now prove the first part of (49). We put

$$u_{2,N} = \sum_{\mathbf{i}} \mathbf{c}_{i_1,1}\mathbf{A}_{i_1,i_3}\delta_{i_3-i_2}(\mathbf{A}^H)_{i_2,i_4}\mathbf{c}_{i_4,1}\mathbf{c}_{i_3,1}\mathbf{A}_{i_3,i_1}(\mathbf{C}_2\mathbf{C}_2^H)_{i_1,i_4}(\mathbf{A}^H)_{i_4,i_2}\mathbf{c}_{i_2,1}$$

Using that $(\mathbf{c}_{i_3,1})^2 = \frac{1}{N}$, we get immediately that

$$u_{2,N} = \frac{1}{N}\sum_{i_1,i_4} \mathbf{c}_{i_1,1}\mathbf{c}_{i_4,1}(\mathbf{C}_2\mathbf{C}_2^H)_{i_1,i_4}\mathbf{E}_{i_1,i_4}$$

where $\mathbf{E}$ is the $N \times N$ matrix defined by

$$\mathbf{E}_{i_1,i_4} = \sum_{i_3} \mathbf{A}_{i_1,i_3}\mathbf{A}_{i_3,i_1}(\mathbf{A}^H)_{i_3,i_4}(\mathbf{A}^H)_{i_4,i_3}$$

It is easy to check that $\mathbf{E} = (\mathbf{A} \bullet \mathbf{A}^T)(\mathbf{A} \bullet \mathbf{A}^T)^H$. Therefore, $u_{2,N}$ can be rewritten as

$$u_{2,N} = \frac{1}{N}\mathbf{c}_1^H \left((\mathbf{C}_2\mathbf{C}_2^H) \bullet \mathbf{E}\right) \mathbf{c}_1$$

As $\mathbf{A}$ and $\mathbf{C}_2\mathbf{C}_2^H$ are uniformly bounded, matrix $(\mathbf{C}_2\mathbf{C}_2^H)\bullet\mathbf{E}$ is uniformly bounded. As $\|\mathbf{c}_1\| = 1$, this implies that

$$\sup_N \mathbf{c}_1^H \left((\mathbf{C}_2\mathbf{C}_2^H) \bullet \mathbf{E}\right) \mathbf{c}_1 < +\infty$$

thus showing that $u_{2,N} \to 0$.

The second part of (49) and (50) are obtained similarly. This establishes (45) for $k = 1$. The proof of (45) for $k \in \{2, 3, 4, 5\}$, and of (45), $k \in \{1, 2, 3, 4, 5\}$ restricted to multi indices satisfying $i_1 = i_2$, $i_1 = i_3$, $i_3 = i_4$, $i_1 = i_2$ and $i_3 = i_4$ are similar, and thus omitted.

APPENDIX IV
PROOF OF LEMMA 4.

As in the proof of Lemma 3, we simplify the notations. We put $\mathbf{B}_N = \mathbf{B}$, $\mathbf{S}(m) = \mathbf{S}$, $\mathbf{S}(m-k) = \mathbf{S}'$, and denote $(s_i)_{i=1,\dots,N}$ and $(s_i')_{i=1,\dots,N}$ their diagonal entries. The diagonal terms of matrix $\mathbf{C}\mathbf{C}^H$ all coincide with $\frac{K}{N}$, and we denote by $\mathbf{D}$ the matrix $\mathbf{D} = \mathbf{C}\mathbf{C}^H - \frac{K}{N}\mathbf{I}$. Finally, we denote by T_N the term to be studied, i.e.

$$T_N = \mathbf{c}_1^H\mathbf{S}^H\mathbf{B}\mathbf{S}'\mathbf{C}\mathbf{C}^H\mathbf{S}'^H\mathbf{B}^H\mathbf{S}\mathbf{c}_1 - \alpha\frac{1}{N}\mathrm{Trace}(\mathbf{B}\mathbf{B}^H)$$

Writing $\mathbf{C}\mathbf{C}^H$ as $\mathbf{D} + \frac{K}{N}\mathbf{I}$ and using that $\mathbf{S}'$ is unitary, we get that

$$T_N = \epsilon_N + \frac{K}{N}\mathbf{c}_1^H\mathbf{S}^H\mathbf{B}\mathbf{B}^H\mathbf{S}\mathbf{c}_1 - \alpha\frac{1}{N}\mathrm{Trace}(\mathbf{B}\mathbf{B}^H)$$

where

$$\epsilon_N = \mathbf{c}_1^H\mathbf{S}^H\mathbf{B}\mathbf{S}'\mathbf{D}\mathbf{S}'^H\mathbf{B}^H\mathbf{S}\mathbf{c}_1$$

As $\mathbf{B}\mathbf{B}^H$ is uniformly bounded, Lemma 1 implies that

$$\mathbf{c}_1^H\mathbf{S}^H\mathbf{B}\mathbf{B}^H\mathbf{S}\mathbf{c}_1 - \frac{1}{N}\mathrm{Trace}(\mathbf{B}\mathbf{B}^H)$$

converges to 0 in the mean square sense. As $\frac{K}{N} \to \alpha$, $\mathbb{E}(\epsilon_N^2) \to 0$ implies that $\mathbb{E}(T_N^2) \to 0$. In the following, we therefore prove that $\mathbb{E}(\epsilon_N^2) \to 0$. For this, we expand $\mathbb{E}(\epsilon_N^2)$ as

$$\sum_{(i_1,i_2,i_3,i_4),(j_1,j_2,j_3,j_4)} \mathbf{c}_{i_1,1}\mathbf{B}_{i_1,j_1}\mathbf{D}_{j_1,i_2}(\mathbf{B}^H)_{i_2,j_2}\mathbf{c}_{j_2,1}\mathbf{c}_{i_3,1}\mathbf{B}_{i_3,j_3}\mathbf{D}_{j_3,i_4}(\mathbf{B}^H)_{i_4,j_4}\mathbf{c}_{j_4,1}\mathbb{E}(s_{i_1}^* s_{j_1}' s_{i_2}'^* s_{j_2} s_{i_3}^* s_{j_3}' s_{i_4}'^* s_{j_4})$$

As sequences $(s_i)_{i=1,\dots,N}$ and $(s_i')_{i=1,\dots,N}$ are independent, it is clear that

$$\mathbb{E}(s_{i_1}^* s_{j_1}' s_{i_2}'^* s_{j_2} s_{i_3}^* s_{j_3}' s_{i_4}'^* s_{j_4}) = \mathbb{E}(s_{i_1}^* s_{j_2} s_{i_3}^* s_{j_4})\mathbb{E}(s_{j_1}' s_{i_2}'^* s_{j_3}' s_{i_4}'^*)$$

But,

$$\begin{aligned}\mathbb{E}(s_{i_1}^* s_{j_2} s_{i_3}^* s_{j_4}) &= \delta_{i_1-j_2}\delta_{i_3-j_4} + \delta_{i_1-j_4}\delta_{j_2-i_3} - \delta_{i_1-j_2}\delta_{i_3-j_4}\delta_{i_1-j_4}\delta_{j_2-i_3} \\ \mathbb{E}(s_{j_1}^{'} s_{i_2}^{'*} s_{j_3}^{'} s_{i_4}^{'*}) &= \delta_{j_1-i_2}\delta_{j_3-j_4} + \delta_{j_1-i_4}\delta_{i_2-j_3} - \delta_{j_1-i_2}\delta_{j_3-j_4}\delta_{j_1-i_4}\delta_{i_2-j_3}\end{aligned}$$

As the diagonal terms of $\mathbf{D}$ are 0, the terms for which $j_1 = i_2$ or $j_3 = i_4$ do not contribute to $\mathbb{E}(\epsilon_N^2)$. Therefore, $\mathbb{E}(\epsilon_N^2)$ reduces to

$$\sum_{(i_1,i_2,i_3,i_4),(j_1,j_2,j_3,j_4)} \mathbf{c}_{i_1,1}\mathbf{B}_{i_1,j_1}\mathbf{D}_{j_1,i_2}(\mathbf{B}^H)_{i_2,j_2}\mathbf{c}_{j_2,1}\mathbf{c}_{i_3,1}\mathbf{B}_{i_3,j_3}\mathbf{D}_{j_3,i_4}(\mathbf{B}^H)_{i_4,j_4}\mathbf{c}_{j_4,1}\mathbb{E}(s_{i_1}^* s_{j_2} s_{i_3}^* s_{j_4})\delta_{j_1-i_4}\delta_{i_2-j_3}$$

Starting from this expression, it is easy to check that $\mathbb{E}(\epsilon_N^2) \to 0$.

REFERENCES

[1] N.I. Akhiezer, *The Classical Moment Problem*, Oliver & Boyd, Edinburgh and London, 1965.

[2] S. Burykh, and K. Abed-Meraim, "Reduced-Rank Adaptive Filtering Using Krylov Subspace", *EURASIP Journal on Advances in Signal Processing (special issue Multiuser Detection and Blind Estimation)*, no. 12, pp. 1387-1400, December 2002.

[3] S. Burykh, and K. Abed-Meraim, "Multi-stage reduced-rank filter with flexible structure", In Proc EUSIPCO-2002, vol. 3, pp. 221-224, Toulouse, France, Sep. 2002.

[4] J. M. Chaufray, W. Hachem, P. Loubaton, "Asymptotic analysis of optimum and sub-optimum CDMA downlink MMSE receivers", *IEEE Trans. on Information Theory*, vol. 50, no. 11, pp. 2620-2638, November 2004.

[5] L. Cottatellucci, R. Muller, " Asymptotic Design and Analysis of Multistage Detectors with Unequal Powers", in Proc. of IEEE Information Theory Workshop (ITW), pp. 167-170, Bangalore, India, October 2002.

[6] "3GPP Technical Specifications 3G TS 25.211, Physical Channels and Mapping of Transport Channels onto Physical Channels (FDD), vol. 1, Release 99,2000.

[7] C.D. Franck, E. Visotsky, "Adaptive interference suppression for CDMA with long spreading codes", in Proc. Allerton Conf., pp. 411-420, Monticello, Illinois, USA, September 1998.

[8] H. Ge, M. Lundberg, L.L. Scharf, "Warp convergence in conjugate gradient Wiener filters", Proc. of the third IEEE Workshop on Sensor Array and Multichannel Signal Processing (SSAP), pp. 109-113, Barcelona, Spain, July 18-21, 2004.

[9] J.S. Goldstein, I.S. Reed, and L.L. Scharf, "A Multistage Representation of the Wiener Filter Based on Orthogonal Projections", *IEEE Trans. on Information Theory*, vol. 44, no. 7, pp. 2943-2959, November 1998.

[10] A.M. Haimovich, A. Shah, W. Xiao, "Reduced-rank array processing for wireless communications with applications to IS-54/IS-136", *IEEE Transactions on Communications*, vol. 48 , no. 5, pp. 743-747.

[11] M.L. Honig and W. Xiao, "Performance of Reduced-Rank Linear Interference Suppression", *IEEE Trans. on Information Theory*, vol. 47, no. 5, pp. 1928-1946, July 2001.

[12] M.L. Honig and W. Xiao, "Adaptive reduced-rank interference suppression with adaptive rank-selection", in Proc. Milcom 2000, vol. 2, pp. 747-751, Los Angeles, CA, Oct. 2000.

[13] K. Hooli, M. Juntti, M. Heikkila, M. Latva-Aho, J. Lilleberg, "Chip-level channel equalization in WCDMA downlink", *EURASIP Journal on Applied Signal Processing*, no. 8, pp. 757-770, August 2002.

[14] A. Klein, ”Data detection algorithms specially designed for the downlink of CDMA mobile radio systems”, in Proc. IEEE 47th Vehicular Technology Conference, vol. 1, pp. 203–207, May 1997.

[15] T.P. Krauss, M. D. Zoltowski and G. Leus,”Simple MMSE equalizer for CDMA downlink to restore chip sequence: Comparison to Zero-Forcing and RAKE”, in Proc. ICASSP-2000, vol. 5, pp. 2865-2868, Istanbul, Turkey, June 2000.

[16] M. Lenardi, A. Medles, D.T.M. Slock, ”A SINR maximizing RAKE receiver for DS-CDMA downlinks”, in Proc. 34th Asilomar Conf., vol. 2, pp. 1283-1287, Pacific Grove, California, October 2000.

[17] L. Li, A. M. Tulino, S. Verdu, ”Design of MMSE multiuser detectors using random matrix techniques”, in Proc. IEEE International Conference on Communications 2003, vol. 4, pp. 2537 - 2541, Anchorage, Alaska, May 2003.

[18] E. Lindskog and C. Tidestav, ”Reduced rank space-time equalization”, in Proc. Personal, Indoor and Mobile Radio Communications, vol. 3 , pp. 1081-1085, September 1998.

[19] P. Loubaton, W. Hachem, ”Asymptotic analysis of reduced-rank Wiener filters”, in Proc. of IEEE Information Theory Workshop (ITW), pp. 328-331, Paris, April 2003.

[20] D. A. Pados and S. N. Batalama, ”Joint Space-Time auxiliary vector filtering for DS/CDMA systems with antenna array”, *IEEE Trans. on Communications*, vol. 47, no. 9, pp. 1406-1415, September 1999.

[21] Y. Saad, ”An overview of Krylov subspace methods with application to control theory”, Report MTNS89, NASA Ames Research Center, Moffet Field, CA, 1989. Available at http://www.cs.umn.edu/ saad/

[22] L. L. Scharf, ”The SVD and reduced-rank signal processing”, *Signal Processing*, vol. 25, no. 2, pp.113-133, 1991.

[23] L.G.F. Trichard, J.S. Evans, and I.B. Collings, ”Optimal linear multistage receivers and the recursive large system SIR”, In Proc. ISIT 2002, p. 21, Lausanne, Switzerland, July 2002.

[24] D. Tse, S. Hanly, ”Linear multiuser receiver: interference, effective bandwidth and user capacity”, *IEEE Trans. on Information Theory*, vol. 45, no. 2, pp. 641-657, March 1999.

[25] S. Sud, W.L. Myrick, J.S. Goldstein, M.D. Zoltowski, ”A reduced-rank MMSE receiver for a DS-CDMA system in frequency selective multipath”, In Proc. MILCOM 2001, vol. 2, pp. 1109-1113, Los Angeles, CA, October 2001.