Asymptotic Analysis of Optimum and Sub-Optimum CDMA Downlink MMSE receivers.

Jean-Marie Chaufray(1), Walid Hachem(2), Philippe Loubaton(3)

Abstract

In this paper, we investigate the performance of two linear receivers for CDMA downlink transmissions over frequency selective channels, the users having possibly different powers. The optimum Minimum Mean Square Error (MMSE) receiver is first considered. Because this receiver requires the knowledge of the code vectors attributed to all the users within the cell when these vectors are time varying, its use may be unrealistic in the forward link. A classical sub-optimum receiver, consisting in a chip rate equalizer followed by a despreading with the code of the user of interest, is therefore studied and compared to the optimum MMSE receiver. Performance of both receivers is assessed through the Signal to Interference plus Noise Ratio (SINR) at their outputs. The analytical expressions of these SINRs depend in a rather non explicit way on the codes allocated to the users of the cell, and are therefore not informative. This difficulty is dealt with by modeling the users code matrix by a random matrix. Because the code matrices used in the forward link are usually isometric, the code matrix is assumed to be extracted from a Haar distributed random unitary matrix. The behavior of the SINRs is studied when the spreading factor and the number of users converge to ∞ at the same rate. Using certain results of the free probability theory, we establish the fact that the SINRs converge almost surely toward quantities that depend only on the

(1) Thalès Communications, 66, rue du Fossé-Blanc, 92231 Gennevilliers, France – Université de Marne-La-Vallée. Phone: +33 1 60 95 72 89, Fax: +33 1 60 95 72 14, chaufray@univ-mlv.fr .

(2) Corresponding author – Supélec, Plateau de Moulon, 91192 Gif-sur-Yvette Cedex, France. Phone: +33 1 69 85 14 52, Fax: +33 1 69 85 14 69, Walid.Hachem@supelec.fr .

(3) Université de Marne-La-Vallée, 5, Bld. Descartes, 77454 Marne-La-Vallée Cedex 2, France. Phone: +33 1 60 95 72 93, Fax: +33 1 60 95 72 14, loubaton@univ-mlv.fr .

complex amplitudes of propagation channel paths. We then put into profit the expressions of these SINR limits to discuss the influence of the various parameters on the performance of the receivers.

Keywords

Linear CDMA Receivers, CDMA Downlink Performance, Large Matrix Analysis, Free Probability

I. Introduction

The design of receivers for Code Division Multiple Access systems has received considerable attention recently. In particular, performance evaluation of the existing detectors became a major related concern. In this course, several works were devoted to the performance study of linear detectors such as the conventional matched filter, the decorrelator, the MMSE receiver, and various kinds of linear interference cancelers (see e.g. [1], [2]). For this study to be done, one commonly uses the observation that the multi-user interference at the output of these receivers can be approximated by a Gaussian distribution, a claim which was thoroughly justified in [3] and quite recently in [4]. Therefore, performance of these linear detectors can be completely characterized by their SINR. As mentioned in [5], the SINR analytical expressions depend on several parameters such as the received powers and the code sequences allocated to the users. In particular, no clear insight on the compared performance of the detectors can be obtained directly from the SINR formulas. To overcome this conceptual difficulty, it is now classical to model the code sequences as random sequences. The various SINRs can in this situation be interpreted as random variables, and it has been shown that, under certain conditions, they converge almost surely toward deterministic quantities when the spreading factor N and the number of users K converge to ∞ in such a way that $\frac{K}{N} \to \alpha$ where $0 < \alpha < 1$. The forms of these limit SINR become quite explicit, and allow to obtain more insight on the parameters that influence the performance of the detectors.

To our knowledge, the vast majority of these works modeled the code sequences of the various users as independent identically distributed (i.i.d.) sequences, mutually independent. Moreover, most of them assumed the channel as frequency flat fading. Noticeable exceptions are [6] and [7] where frequency selective fading channels are considered.

Assuming the code sequences of the various users mutually independent and i.i.d. is certainly justified in the uplink transmission direction. However, in the downlink, code vectors are usually constrained to be orthonormal instead, thanks to the fact that downlink transmissions are synchronized. Code vectors orthonormality allows intuitively to achieve a better separation of the various users. This fact appears clearly when the channel fading is frequency flat, because in this case the matched filter suppresses the multi-user interference. However, this nice property is no longer verified for frequency selective fading channels. The purpose of this paper is to study and compare the performance of two MMSE like receivers in the context of downlink transmissions corrupted by multi-path Rayleigh fading.

The linear receivers considered in this paper, beginning with the traditional MMSE receiver, are implemented on a mobile station. Implementing the MMSE receiver requires the knowledge or at least the estimation of the covariance matrix of the received vector signal. When code vectors are time varying as it is frequently the case, this practically means that the code vectors and powers associated to all interfering user signals within the cell have to be known to the mobile station of interest. In the existing CDMA systems, this is actually not the case. Partly for this reason, we also study the performance of a sub-optimum MMSE receiver. Roughly speaking, it consists in a Wiener filter which equalizes the chip-rate discrete-time equivalent channel. The despread output of this filter gives an estimate of the transmitted symbols ([8], [9], [10]).

In order to get insights into the compared performance of these two receivers and to evaluate clearly the loss of performance induced by the use of the sub-optimum receiver, it is necessary to obtain interpretable expressions of their associated SINR. For this, we still propose to model the spreading codes as random variables, and to study the behavior of the SINR when the spreading factor N and the number of users K converge to $+\infty$ in such a way that $\frac{K}{N} \to \alpha$, $0 < \alpha < 1$. The point is that code vectors orthonormality is now taken into account. More precisely, it will be assumed here that the K codes associated to the K active users of the cell under consideration coincide with K columns of a $N \times N$ random unitary matrix. Another important assumption concerns the propagation channel between the transmitter and the receiver. Here, we assume that when N converges toward

$+\infty$, the channel parameters (i.e. the time delays of the various paths and their complex amplitudes) remain constant. In particular, the delay spread of the channel tends to be negligible with respect to the symbols duration, an hypothesis which is often met in practical CDMA systems.

This paper is structured as follows. In section II, we introduce our notations and assumptions, and give the discrete-time model of the received signal sampled at the chip rate. In section III, we study the optimum MMSE receiver based on the knowledge of the codes and the powers of the interfering user signals within the cell. We first state an important intermediate result showing that the influence of the inter-symbol interference generated by the channel on the SINR can be considered as negligible. We are thus essentially led back to a model in which the N–dimensional vector obtained by stacking N consecutive values of the sampled received signal is obtained as (a noisy version) of the action of the vector of the symbols transmitted to the various users on a certain $N \times K$ matrix. This matrix is defined as the product of three matrices. The first one is a circulant Toeplitz $N \times N$ matrix built from the coefficients of the discrete equivalent channel, the second matrix is the $N \times K$ matrix obtained by putting the K vectors associated to the codes of the users side by side, and the third matrix is a diagonal $K \times K$ matrix with positive entries modeling a possibly non uniform allocated power distribution. When the powers allocated to the various users are equal, it is possible to use the results of the work [11] to study the asymptotic behavior of the SINR. However, the results of [11] are no longer valid in the non uniform power distribution case. Another approach is therefore proposed here to deal with this problem. As in [11], we compare the SINR obtained with an isometric code matrix, with the results provided by mutually independent i.i.d. codes. The obtained formulas allow us to have a better evaluation of the merits of using isometric code matrices in the context of downlink transmissions. In section IV, we study the performance of a sub-optimum MMSE receiver, which consists in a Wiener equalizer of the chip-rate discrete-time equivalent channel followed by a despreading. We evaluate the limit of the corresponding SINR, and compare the corresponding expressions with the results obtained in the case where all the codes are known to the receiver. This allows us to evaluate the loss of performance caused by the non availability of interfering users codes and powers.

We also present some numerical results showing that our evaluations allow to predict accurately the performance of real life CDMA systems. In particular, we implement the specifications of the downlink of the UMTS wide band CDMA mode, and observe that the empirical results match the theoretical ones.

II. Model and assumptions.

We consider a multi-user communication system based on a direct sequence spread spectrum with spreading factor N. We assume that a certain base station transmits simultaneously K centered unit variance symbol sequences $(s_k(n))_{n\in\mathbb{Z}}$ to K mobile receivers. To transmit each symbol sequence $(s_k(n))_{n\in\mathbb{Z}}$, the base station produces for each n the so-called chip sequence $(x_k(m))_{m\in\mathbb{Z}}$ defined by the fact that

$$\mathbf{x}_k(n) = (x_k(nN), \ldots, x_k(nN+N-1))^T = \mathbf{w}_k \sqrt{p_k} s_k(n) \tag{1}$$

where $\sqrt{p_k}$ is a positive scaling factor representing the amplitude allocated to user k, and where the N-dimensional vector $\mathbf{w}_k = (w_k^{(0)}, \ldots, w_k^{(N-1)})^T$ is the code vector allocated to that user. In order to simplify the notations, we assume that the code vectors $(\mathbf{w}_k)_{k=1,\ldots,K}$ are time-invariant (i.e. independent of n), a condition which is not verified in certain existing CDMA systems. However, most of the following results extend trivially to the case where these vectors are time varying. When the adaptation to the time-varying case is not obvious, some comments will be provided. The transmitter delivers the composite chip sequence $x(m) = \sum_{k=1}^{K} x_k(m)$, which is pulse shaped, transmitted across a multiple path frequency selective fading channel, and received by a mobile station, say the mobile station intended to detect symbol sequence $(s_1(n))$. The received signal is filtered then sampled at chip rate. The resulting discrete-time signal expresses as

$$y(m) = \sum_{k=0}^{M} h_k x(m-k) + v(m) \tag{2}$$

where $h(z) = \sum_{k=0}^{M} h_k z^{-k}$ is the transfer function of the discrete time composite channel with a degree M strictly smaller that N, and $v(m)$ is an AWGN independent of $x(m)$ and having of variance of σ^2. The coefficients channel $(h_0, \ldots, h_M)$ are assumed known to the receiver.

It is more convenient for our purpose to express relation (2) in terms of the vector $\mathbf{y}(n)$ defined as $\mathbf{y}(n) = (y(nN), \ldots, y(nN+N-1))^T$. As $M < N$, vector $\mathbf{y}(n)$ can be written as

$$\mathbf{y}(n) = \mathcal{H}_0 \mathbf{x}(n) + \mathcal{H}_1 \mathbf{x}(n-1) + \mathbf{v}(n) \tag{3}$$

where $\mathbf{x}(n)$ and $\mathbf{v}(n)$ are built in the same manner as $\mathbf{y}(n)$ and where $\mathcal{H}_0$ and $\mathcal{H}_1$ are the $N \times N$ Toeplitz matrices

$$\mathcal{H}_0 = \begin{pmatrix} h_0 & & & & & \\ \vdots & \ddots & & \mathbf{0} & & \\ h_M & & \ddots & & & \\ & \ddots & & h_0 & & \\ & & \ddots & \vdots & \ddots & \\ & \mathbf{0} & & h_M & \ldots & h_0 \end{pmatrix} \tag{4}$$

$$\mathcal{H}_1 = \begin{pmatrix} & & h_M & & h_1 \\ & & & \ddots & \\ & & & & h_M \\ & \mathbf{0} & & & \\ & & & & \end{pmatrix}.$$

If we denote by $\mathbf{W}$ the $N \times K$ code matrix defined by

$$\mathbf{W} = (\mathbf{w}_1, \ldots, \mathbf{w}_K)$$

and by $\mathbf{P}$ the diagonal $K \times K$ matrix whose entries are $p_1, \ldots, p_K$, equation (3) can be rewritten as

$$\mathbf{y}(n) = \mathcal{H}_0 \mathbf{W} \sqrt{\mathbf{P}} \mathbf{s}(n) + \mathcal{H}_1 \mathbf{W} \sqrt{\mathbf{P}} \mathbf{s}(n-1) + \mathbf{v}(n) \tag{5}$$

where $\mathbf{s}(n) = (s_1(n), \ldots, s_K(n))^T$. In most cases, code vectors allocated to the various users are orthonormal, in other words matrix $\mathbf{W}$ satisfies the relation

$$\mathbf{W}^H \mathbf{W} = \mathbf{I}_K \tag{6}$$

because this relation intuitively allows to improve the performances of the detection of the symbol sequence $(s_1(n))_{n\in\mathbb{Z}}$. This is in particular the case when the channel is flat fading and the shaping filter is a Nyquist filter. In this context, all the coefficients $(h_k)_{k=1,M}$ are zero, and h_0 is reduced to the complex amplitude of the path. Therefore, (5) reduces to

$$\mathbf{y}(n) = h_0\mathbf{W}\sqrt{\mathbf{P}}\mathbf{s}(n) + \mathbf{v}(n)$$

The matched filter receiver $\mathbf{w}_1^H\mathbf{y}(n)$ coincides in this case with the maximum likelihood detector, a property which is no longer true if matrix $\mathbf{W}$ is not isometric. This simple observation suggests that the orthogonality of the code vectors may have an important impact on the performance of the most classical linear detectors.

In order to assess the respective merits of the detectors that will be introduced in the next two sections, we propose to study their output SINR in the case where the code matrix $\mathbf{W}$ is a random isometric matrix and where N and K converge to $+\infty$ in such a way that the ratio $\frac{K}{N}$ converges to a constant $\alpha < 1$. Before stating more precisely the technical hypotheses we are going to formulate, we notice that as N grows to infinity, the discrete time equivalent channel $h(z) = \sum_{k=0}^{M} h_k z^{-k}$ is supposed to be kept constant. In particular, the degree M of $h(z)$ becomes negligible with respect to the spreading factor. In practice, this means that our results are applicable if the delay spread of the channel is much smaller than the symbol duration and if the complex amplitudes of the paths do not vary within the duration of one symbol. Our numerical evaluations show that our analysis predicts quite well the performances if the ratio of the symbol duration over the delay spread is equal to 20. As it will appear in the next section, the condition that $M/N \to 0$ is important on a technical point of view because it will allow to replace the inter-symbol interference term

$$\mathcal{H}_1\mathbf{W}\sqrt{\mathbf{P}}\mathbf{s}(n-1)$$

by

$$\mathcal{H}_1\mathbf{W}\sqrt{\mathbf{P}}\mathbf{s}(n)$$

in the expression (5) of $\mathbf{y}(n)$, a replacement which will lead to more tractable expressions.

We are now in a position to say how the random isometric matrix $\mathbf{W}$ is modeled. For that purpose, some notations and definitions need to be introduced. Denote by $\mathcal{U}$ the multiplicative group of $N \times N$ unitary matrices, and by $\mathbf{\Theta}$ a random $N \times N$ unitary matrix. $\mathbf{\Theta}$ is said to be Haar distributed if the probability distribution of $\mathbf{\Theta}$ is invariant by left multiplication by constant unitary matrices [1]. Since the group $\mathcal{U}$ is compact, this condition is known to be equivalent to the invariance of the probability distribution of $\mathbf{\Theta}$ by right multiplication by constant unitary matrices. In order to generate Haar distributed unitary random matrices, let $\mathbf{X} = [x_{i,j}]_{1 \le i,j \le N}$ be a $N \times N$ random matrix with independent complex Gaussian centered unit variance entries. Then (see e.g. [11]), the unitary matrix $\mathbf{X}(\mathbf{X}^H\mathbf{X})^{-1/2}$ is Haar distributed. There is another way for building Haar distributed unitary matrices that will be useful to our purpose. Instead of multiplying $\mathbf{X}$ by the inverse of the Hermitian square root of $\mathbf{X}^H\mathbf{X}$, one can introduce the uniquely defined upper triangular matrix with positive diagonal elements $\mathbf{T}(\mathbf{X})$ defined by

$$\mathbf{X}^H\mathbf{X} = \mathbf{T}(\mathbf{X})^H\mathbf{T}(\mathbf{X}) \ .$$

The unitary matrix $\mathbf{V}(\mathbf{X})$ defined by

$$\mathbf{V}(\mathbf{X}) = \mathbf{X}\mathbf{T}(\mathbf{X})^{-1} \tag{7}$$

is also Haar distributed ([11]). Finally, we state an interesting property of Haar distributed unitary random matrices. $\mathbf{\Theta}$ being one such matrix, its probability distribution is also invariant under right multiplication by unitary matrices, hence this distribution coincides with the distribution of $\mathbf{\Theta}\mathcal{P}$ for any permutation matrix $\mathcal{P}$. This shows that the $N \times K$ isometric matrices obtained by extracting *any* subset of K columns from $\mathbf{\Theta}$ have the same probability distribution.

In the following, it will be assumed that

A1 : matrix $\mathbf{W}$ is generated by extracting K columns from a $N \times N$ Haar unitary random matrix $\mathbf{\Theta}$.

Code matrices commonly used in CDMA systems are of course not obtained as realizations of Haar distributed unitary random matrices. They are often deterministic orthogonal

[1] In other words, the probability distribution of $\mathbf{\Theta}$ coincides with the so-called Haar measure on $\mathcal{U}$.

sequences (e.g. Walsh Hadamard sequences) multiplied by a scrambling code. It is possible to show that the results related to the sub-optimum receivers (section IV) are still valid in this context (the proofs are however quite different). It seems more difficult to check that the results of section III remain valid in this context, but simulations results tend to indicate that it is the case.

III. Performance study of the optimum MMSE receiver.

In this section, we study the SINR of the MMSE receiver when N and K converge toward $+\infty$. We assume that the user of interest is user 1 corresponding to symbol sequence $(s_1(n))_{n\in\mathbb{Z}}$. In order to simplify the notations, we denote from now on the code vector $\mathbf{w}_1$ by $\mathbf{w}$, and by $\mathbf{U}$ the $N \times (K-1)$ isometric matrix obtained by deleting the first column of $\mathbf{W}$. In other words,

$$\mathbf{W} = (\mathbf{w}, \mathbf{U}) \ .$$

The $(K-1) \times (K-1)$ diagonal matrix obtained by deleting the first row and the first column of $\mathbf{P}$ is denoted $\mathbf{Q}$.

The MMSE receiver we consider in this section consists in estimating symbol $s_1(n)$ by a linear combination $\tilde{s}_1(n) = \mathbf{g}^T\mathbf{y}(n)$ of the components of $\mathbf{y}(n)$ chosen in such a way that $E|s_1(n) - \tilde{s}_1(n)|^2$ be minimum. It is clear that

$$\begin{aligned} \tilde{s}_1(n) &= E\left[s_1(n)\mathbf{y}^H(n)\right] \left(E\left[\mathbf{y}(n)\mathbf{y}^H(n)\right]\right)^{-1} \mathbf{y}(n) \\ &= \sqrt{p_1}\mathbf{w}^H\mathcal{H}_0^H\left(\mathcal{H}_0\mathbf{W}\mathbf{P}\mathbf{W}^H\mathcal{H}_0^H + \mathcal{H}_1\mathbf{W}\mathbf{P}\mathbf{W}^H\mathcal{H}_1^H + \sigma^2\mathbf{I}_N\right)^{-1}\mathbf{y}(n) \ . \end{aligned} \tag{8}$$

where it should be understood that the mathematical expectation represents the conditional expectation given $\mathbf{W}$.

In existing CDMA systems, mobile station 1 is not aware of the codes allocated to the interfering users nor of the transmitted powers $(p_k)_{k\geq 2}$. In this context, the use of the classical MMSE receiver may be somewhat unrealistic because the use of formula (8) defining the estimate of $s_1(n)$ requires the knowledge of matrices $\mathbf{W}$ and $\mathbf{P}$. However, one should note that if the code vectors $(\mathbf{w}_k)_{k=1,\ldots,K}$ are time-invariant as assumed here, and if the channel coefficients vary slowly, then the covariance matrix of vector $\mathbf{y}(n)$ can be estimated consistently from the received signal. In such a context, the use of the MMSE receiver does not require the explicit knowledge of matrices $\mathbf{W}$ and $\mathbf{P}$. In more general cases, we nevertheless believe that it is important to study the performance of the MMSE receiver, first for the purpose of comparison, and second because one may imagine that if needed, base stations of future CDMA systems would transmit to all mobile stations the relevant information. We thus begin by studying the case where matrices $\mathbf{W}$ and $\mathbf{P}$ are available. In the next section, we will consider the case where the codes allocated to users 2 to K are unknown, and will study a sub-optimal MMSE receiver.

The output $\tilde{s}_1(n)$ of the MMSE receiver is corrupted by both the thermal noise and the multi-user interference due to the contributions of $\{s_k(n)\}_{k\neq 1}$. Poor and Verdú showed ([3]) that the multi-user interference can be considered as Gaussian when N and K are large enough if the code matrix $\mathbf{W}$ is considered as deterministic. Zhang *et al.* ([4]) extended this result to the case where $\mathbf{W}$ is a random matrix with i.i.d. entries. This justifies the use of the signal to interference plus noise ratio as a performance figure, although the situation considered in this paper ($\mathbf{W}$ is a random matrix obtained from a Haar distributed matrix) is not covered by [3] nor by [4].

As is well known, the SINR, denoted $\tilde{\beta}_N$, can be written as

$$\tilde{\beta}_N = p_1 \mathbf{w}^H \mathcal{H}_0^H \left(\mathcal{H}_0 \mathbf{U}\mathbf{Q}\mathbf{U}^H \mathcal{H}_0^H + \mathcal{H}_1 \mathbf{W}\mathbf{P}\mathbf{W}^H \mathcal{H}_1^H + \sigma^2 \mathbf{I}_N \right)^{-1} \mathcal{H}_0 \mathbf{w} \; . \tag{9}$$

The idea at this point is to remark that as $\frac{M}{N}$ converges to zero when $N \to +\infty$, then the SINR $\tilde{\beta}_N$ behaves asymptotically like the SINR corresponding to the following modified observation model :

$$\mathbf{y}(n) = \mathcal{H}_0 \mathbf{W}\sqrt{\mathbf{P}}\mathbf{s}(n) + \mathcal{H}_1 \mathbf{W}\sqrt{\mathbf{P}}\mathbf{s}(n) + \mathbf{v}(n) \tag{10}$$

which can also be rewritten as

$$\mathbf{y}(n) = \mathbf{H}\mathbf{W}\sqrt{\mathbf{P}}\mathbf{s}(n) + \mathbf{v}(n) \tag{11}$$

where $\mathbf{H}$ is the $N \times N$ circulant Toeplitz matrix $\mathcal{H}_0 + \mathcal{H}_1$. We note that equation (11) represents the signal model in the case where the chip sequence corresponding to the transmission of each symbol $s_k(n)$ is augmented by a cyclic prefix. In other words, if the

N–dimensional chip sequence $\left(w_k^{(0)} s_k(n), \ldots, w_k^{(N-1)} s_k(n)\right)$, was replaced by the $N+M$–dimensional sequence defined by $\left(w_k^{(N-M)} s_k(n), \ldots, w_k^{(N-1)} s_k(n), w_k^{(0)} s_k(n), \ldots, w_k^{(N-1)} s_k(n)\right)$, then equation (11) would represent the received signal after cancellation of the cyclic prefix. The output $\hat{s}_1(n)$ of the MMSE receiver corresponding to model (11) is given by

$$\hat{s}_1(n) = \sqrt{p_1}\mathbf{w}^H\mathbf{H}^H \left(\mathbf{HWPW}^H\mathbf{H}^H + \sigma^2\mathbf{I}_N\right)^{-1} \mathbf{y}(n) \tag{12}$$

and the corresponding signal to interference plus noise ratio, denoted β_N in the following, is given by

$$\beta_N = p_1\mathbf{w}^H\mathbf{H}^H \left(\mathbf{HUQU}^H\mathbf{H}^H + \sigma^2\mathbf{I}_N\right)^{-1} \mathbf{Hw}\ . \tag{13}$$

We now state a result showing that $\tilde{\beta}_N$ and β_N have the same behavior if N and K converge to $+\infty$ in such a way that $\frac{K}{N} \to \alpha < 1$. For this, we first formulate the following assumption:

A2 : It exists two strictly positive constant b and B independent of K such that $0 < b \leq p_k \leq B$ for $k = 1, \ldots, K$.

It is also useful to remark that $\|\mathbf{H}\| \leq \sup_{f\in[-1/2,1/2]} |h(e^{2i\pi f})|$, where $\|.\|$ is the spectral norm, due to the fact that $\mathbf{H}$ is circulant. This matrix verifies then

$$\sup_N \|\mathbf{H}\| < \infty\ . \tag{14}$$

Because $\sup_K \|\mathbf{P}\| < \infty$, $\sup_K \|\mathbf{P}^{-1}\| < \infty$, and $\sup_N \|\mathbf{H}\| < \infty$, matrices $\mathbf{P}$, $\mathbf{P}^{-1}$, and $\mathbf{H}$ are said uniformly bounded.

Proposition 1: Assume that conditions **A1** and **A2** hold. Then, $\tilde{\beta}_N - \beta_N$ converges to 0 almost surely when N and K converge to $+\infty$ in such a way that $\frac{K}{N} \to \alpha < 1$

The proof is deferred to appendix A. This proposition shows that instead of studying model (5), it is possible to consider (11) and the corresponding SINR β_N. From now on, we thus replace (5) by (11) and study the behavior of β_N.

In order to introduce the main results of this paper in a comprehensive way, we formulate the following assumption :

A3 : The distribution $(p_k)_{k=1,\ldots,K}$ of the powers of the K users converges when $K \to \infty$ to the distribution $\nu_p = \sum_{l=1}^{L} \rho_l \delta(p - P_l)$ where the coefficients $(\rho_l)_{l=1,\ldots,L}$ are positive weights such that $\sum_{l=1}^{L} \rho_l = 1$. More precisely, for each bounded continuous function ϕ,

$$\lim_{K\to\infty} \frac{\sum_{k=1}^{K} \phi(p_k)}{K} = \int_0^\infty \phi(p)\, d\nu_p(p) = \sum_{l=1}^{L} \rho_l \phi(P_l)$$

In other words, the limit distribution of the $(p_k)_{k=1,\dots,K}$ contains L classes of users. The users of class l have the same power P_l, and coefficients $(\rho_l)_{l=1,\dots,L}$ represent the percentages of users in each class. We note that our results extend immediately to situations where the distribution of powers $(p_k)_{k=1,\dots,K}$ converge to a more general compactly supported distribution. We only consider the case of a discrete distribution in order to simplify the presentation of our results.

We are now in a position to state the two main results of this section.

Theorem 1: Assume **A1** to **A3**, and that N and K converge to $+\infty$ and $\frac{K}{N} \to \alpha < 1$. Then, the normalized SINR $\frac{\beta_N}{p_1}$ converges almost surely toward the deterministic constant $\overline{\beta}$ defined as the unique solution of the equation

$$\overline{\beta} = \int_{-1/2}^{1/2} \frac{|h(e^{2i\pi f})|^2}{\alpha |h(e^{2i\pi f})|^2 \sum_{l=1}^{L} \rho_l \frac{P_l}{P_l\overline{\beta}+1} + \sigma^2 \left(1 - \alpha \sum_{l=1}^{L} \rho_l \frac{P_l\overline{\beta}}{P_l\overline{\beta}+1}\right)} df \tag{15}$$

Notice that the limit $\overline{\beta}$ of the normalized SINR does not depend on the user. For the purpose of comparison, we also give the performance of the MMSE receiver in the case where the entries of $\mathbf{W}$ are i.i.d. random variables.

Theorem 2: Assume that the entries of $\mathbf{W}$ are centered i.i.d. random variables of variance $\frac{1}{N}$ and finite eighth moment. Then, under **A2** and **A3**, the normalized SINR $\frac{\beta_N}{p_1}$ converges almost surely when N and K converges to $+\infty$ and $\frac{K}{N} \to \alpha < 1$ toward the deterministic constant $\overline{\beta}_{\text{iid}}$ defined as the unique solution of the equation

$$\overline{\beta}_{\text{iid}} = \int_{-1/2}^{1/2} \frac{|h(e^{2i\pi f})|^2}{\alpha |h(e^{2i\pi f})|^2 \sum_{l=1}^{L} \rho_l \frac{P_l}{P_l\overline{\beta}_{\text{iid}}+1} + \sigma^2} \tag{16}$$

Proof. The proof of theorem 2 is similar to the proof of theorem 2 of [11]. We therefore only recall the result on which it is based. As $\mathbf{W}$ is i.i.d., then vector $\mathbf{w}$ and matrix

$$\mathbf{B} = \mathbf{H}^H \left(\mathbf{HUQU}^H\mathbf{H}^H + \sigma^2 \mathbf{I}_N\right)^{-1} \mathbf{H}$$

are independent. Therefore (see [12] and [6]), $\frac{\beta_N}{p_1} = \mathbf{w}^H\mathbf{B}\mathbf{w}$ and $\frac{1}{N}\text{tr}(\mathbf{B})$ have the same behavior. The limit of $\frac{1}{N}\text{tr}(\mathbf{B})$ is eventually evaluated by using the classical results of [13].

We now present informally the main steps of the proof of theorem 1. The detailed proof is given in the appendix.

First step. The first difficulty encountered in the isometric case follows from the observation that $\mathbf{w}$ and matrix $\mathbf{B}$ are of course no longer independent. Hence, $\mathbf{w}^H\mathbf{B}\mathbf{w}$ and $\frac{1}{N}\mathrm{tr}(\mathbf{B})$ do not share the same limit. Actually, it is shown in the appendix that if we let

$$\mathbf{A} = \mathbf{\Pi}\mathbf{H}^H \left(\mathbf{H}\mathbf{U}\mathbf{Q}\mathbf{U}^H\mathbf{H}^H + \sigma^2\mathbf{I}_N\right)^{-1}\mathbf{H} \tag{17}$$

where $\mathbf{\Pi} = \mathbf{I}_N - \mathbf{U}\mathbf{U}^H$ is the orthogonal projection matrix onto the subspace orthogonal to the columns of $\mathbf{U}$, then

$$\frac{\beta_N}{p_1} - \frac{1}{N-K}\mathrm{tr}(\mathbf{A}) \to 0 \text{ a.s.} \tag{18}$$

when N and K converge to $+\infty$ in such a way that $\frac{K}{N}$ converges to α. In order to give an intuitive idea of the reasons for which this result holds, assume that vector $\mathbf{w}$ is given by

$$\mathbf{w} = \frac{\mathbf{\Pi}\mathbf{x}}{\|\mathbf{\Pi}\mathbf{x}\|} \tag{19}$$

where $\mathbf{x}$ is a complex Gaussian N-dimensional centered random vector independent of $\mathbf{U}$, and satisfying $E[\mathbf{x}\mathbf{x}^H] = \mathbf{I}_N$. In this case, $\frac{\beta_N}{p_1}$ can be written as

$$\frac{\beta_N}{p_1} = \frac{1}{N}\mathbf{x}^H\mathbf{A}\mathbf{\Pi}\mathbf{x} \times \frac{N}{\|\mathbf{\Pi}\mathbf{x}\|^2}$$

In order to justify (18), we first remark that as $\frac{\mathbf{x}}{\sqrt{N}}$ is independent of $\mathbf{A}$, then $\frac{1}{N}\mathbf{x}^H\mathbf{A}\mathbf{\Pi}\mathbf{x} - \frac{1}{N}\mathrm{tr}(\mathbf{A}) \to 0$. Just notice that $\mathrm{tr}(\mathbf{A}\mathbf{\Pi}) = \mathrm{tr}(\mathbf{\Pi}\mathbf{A}) = \mathrm{tr}(\mathbf{A})$.

We now study the term $\frac{N}{\|\mathbf{\Pi}\mathbf{x}\|^2}$. As $\mathbf{x}$ is Gaussian, $\|\mathbf{\Pi}\mathbf{x}\|^2$ coincides with the sum of the squares of $(N-K+1)$ independent centered unit variance complex Gaussian random variables. Therefore,

$$\frac{\|\mathbf{\Pi}\mathbf{x}\|^2}{N-K} \to 1$$

almost surely, and

$$\frac{N}{\|\mathbf{\Pi}\mathbf{x}\|^2} - \frac{N}{N-K} \to 0$$

This gives (18). The rigorous proof of (18) is presented in the appendix, where it is in particular shown that there is no restriction to assume that $\mathbf{w}$ is given by (19).

We have now to show that $\frac{\mathrm{tr}(\mathbf{A})}{N-K}$ converges almost surely. As matrices $\mathbf{H}^H\mathbf{H}$ and $\mathbf{UQU}^H$ are asymptotically free almost everywhere (see [14] or [11] for a short introduction and [15] for a detailed presentation), it is not hard to show that $\frac{\mathrm{tr}(\mathbf{A})}{N-K}$ converges almost surely toward a certain deterministic constant (see appendix).

Second step. The second difficulty of the proof of theorem 1 is to evaluate this limit. For this to be done, one could use a result of Biane (see [14]) indicating how to evaluate the limit of expressions of the form

$$\frac{1}{N}\mathrm{tr}(f(\mathbf{R}+\mathbf{S})g(\mathbf{R}))$$

where $\mathbf{R}$ and $\mathbf{S}$ are two almost everywhere asymptotically free random matrices and where f and g are smooth enough functions. However, this direct approach needs the introduction of some perhaps complicated tools. We therefore use an alternative method, which, we hope, is easier to follow.

In order to present the present approach, we need to introduce the SINR, denoted $\beta_{N,k}$, that corresponds to the MMSE estimate of each component s_k for $k = 1, \dots, K$. In particular, the SINR β_N under study coincides with $\beta_{N,1}$. We also denote $\mathbf{U}^{(k)}$ and $\mathbf{Q}^{(k)}$ the matrices obtained by deleting the k-th column of $\mathbf{W}$ and the entry (k,k) of $\mathbf{P}$ respectively ($\mathbf{U}^{(1)}$ and $\mathbf{Q}^{(1)}$ thus coincide with $\mathbf{U}$ and $\mathbf{Q}$). Finally, $\mathbf{A}_k$ represents the matrix defined by

$$\mathbf{A}_k = (\mathbf{I} - \mathbf{U}^{(k)}\mathbf{U}^{(k)H})\mathbf{H}^H \left(\mathbf{H}\mathbf{U}^{(k)}\mathbf{Q}^{(k)}\mathbf{U}^{(k)H}\mathbf{H}^H + \sigma^2\mathbf{I}_N\right)^{-1}\mathbf{H}$$

which means in particular that matrix $\mathbf{A}$ defined by (17) coincides with $\mathbf{A}_1$. For each $k = 1, \dots, K$, we define $\overline{\beta}_{N,k}$ by

$$\overline{\beta}_{N,k} = \frac{\beta_{N,k}}{p_k} = \mathbf{w}_k^H\mathbf{H}^H \left(\mathbf{H}\mathbf{U}^{(k)}\mathbf{Q}^{(k)}\mathbf{U}^{(k)H}\mathbf{H}^H + \sigma^2\mathbf{I}_N\right)^{-1}\mathbf{H}\mathbf{w}_k \tag{20}$$

and put

$$\eta_{N,k} = p_k\mathbf{w}_k^H\mathbf{H}^H \left(\mathbf{HWPW}^H\mathbf{H}^H + \sigma^2\mathbf{I}_N\right)^{-1}\mathbf{H}\mathbf{w}_k = \frac{p_k\overline{\beta}_{N,k}}{1+p_k\overline{\beta}_{N,k}}\ . \tag{21}$$

The results of step 1 show that for each sequence $(k(N))_{N\geq 1}$ of integers satisfying $1 \leq k(N) \leq K$, $\overline{\beta}_{N,k(N)}$ and $\frac{\mathrm{tr}(\mathbf{A}_{k(N)})}{N-K}$ have the same asymptotic behavior. Moreover, it is not hard to check that the limit of $\frac{\mathrm{tr}(\mathbf{A}_{k(N)})}{N-K}$ does not depend on the choice of the sequence $(k(N))_{N\geq 1}$ (see appendix). This common limit is precisely the limit normalized SINR

$\overline{\beta}$. It is clear that for each sequence of integers $(k(N))_{N\geq 1}$ satisfying $1 \leq k(N) \leq K$, $\eta_{N,k(N)} - \frac{p_{k(N)}\overline{\beta}}{1+p_{k(N)}\overline{\beta}}$ converges to 0. Using this fact, it is shown in the appendix that when $N \to \infty, K \to \infty$, and $\frac{K}{N} \to \alpha$, then,

$$\frac{\sum_{k=1}^{K} \eta_{N,k}}{K} \longrightarrow \int \frac{\lambda\overline{\beta}}{1+\lambda\overline{\beta}}\, d\nu_p(\lambda) = \sum_{l=1}^{L} \rho_l \frac{P_l\overline{\beta}}{1+P_l\overline{\beta}} \tag{22}$$

Here, we recall that if θ is a certain probability measure, then the ψ transform of θ is the function ψ_θ defined by

$$\psi_\theta(z) = \int \frac{z\lambda}{1-z\lambda}\, d\theta(\lambda) \tag{23}$$

The right hand side of (22) thus coincides with -$\psi_{\nu_p}(-\overline{\beta})$. In order to calculate $\overline{\beta}$, we now observe that the limit of $\frac{\sum_{k=1}^{K}\eta_{N,k}}{K}$ can be evaluated independently. Indeed, $\frac{\sum_{k=1}^{K}\eta_{N,k}}{K}$ can be written as

$$\frac{\sum_{k=1}^{K} \eta_{N,k}}{K} = \frac{1}{K}\mathrm{tr}\left((\mathbf{HWPW}^H\mathbf{H}^H + \sigma^2\mathbf{I}_N)^{-1}\mathbf{HWPW}^H\mathbf{H}^H\right) \tag{24}$$

As $\mathbf{WPW}^H$ and $\mathbf{H}^H\mathbf{H}$ are asymptotically free almost everywhere, the right hand side of (24) converges almost surely toward

$$-\frac{1}{\alpha}\psi_{\mu\boxtimes\nu}(-\frac{1}{\sigma^2}) = \frac{1}{\alpha}\int \frac{\lambda}{\lambda+\sigma^2}\, d\mu\boxtimes\nu(\lambda)$$

where μ and ν represent the limit eigenvalue distributions of $\mathbf{H}^H\mathbf{H}$ and $\mathbf{WPW}^H$ respectively, and $\mu\boxtimes\nu$ denotes the free multiplicative convolution product of μ and ν. Measures μ and ν can be characterized easily (the eigenvalues of $\mathbf{H}^H\mathbf{H}$ are the $(|h(e^{2i\pi k/N}|^2)_{k=1,\dots,N}$ and $\nu = \alpha\nu_p + (1-\alpha)\delta(\lambda)$). Therefore, it is possible to give the expression of $\psi_{\mu\boxtimes\nu}(-\frac{1}{\sigma^2})$. Equating -$\psi_{\nu_p}(-\overline{\beta})$ with $-\frac{1}{\alpha}\psi_{\mu\boxtimes\nu}(-\frac{1}{\sigma^2})$ results in equation (15).

The two above results deserve some comments.

Discussion of Theorem 2. We first compare (16) to the results presented in [5]. [5] considered the case of a flat fading channel, which amounts to $h(e^{2i\pi f}) = h_0$ for each f. Formula (16) then coincides with what is found in [5], *i.e.*,

$$\overline{\beta}_{\mathrm{iid}} = \frac{|h_0|^2}{\alpha|h_0|^2\sum_{l=1}^{L}\rho_l\frac{P_l}{P_l\overline{\beta}_{\mathrm{iid}}+1} + \sigma^2}\ . \tag{25}$$

Recall that $\alpha\sum_{l=1}^{L}\rho_l\frac{P_l}{P_l\overline{\beta}_{\mathrm{iid}}+1} \simeq \frac{1}{N}\sum_{k=2}^{K}\frac{p_k}{p_k\overline{\beta}_{\mathrm{iid}}+1}$ for K and N large enough is interpreted in [5] as the effective interference produced by users 2 to K on the signal of user 1 at a target normalized SINR of $\overline{\beta}_{\rm iid}$. One of the main conclusions of [5] was then that the total multi-user interference can be decoupled into a sum of interference terms from each of the interfering users. By inspecting the result of theorem 2, it turns out that this interpretation can be generalized to the frequency selective channel case. To fix our ideas, let us introduce the model

$$\mathbf{r}(n) = \mathbf{H}\left(\sqrt{p_1}s_1(n)\mathbf{w} + \sum_{k=2}^{K}\mathbf{u}_k(n)\right) + \mathbf{v}(n) \tag{26}$$

where the $K-1$ uncorrelated vectors $\{\mathbf{u}_k\}$ are also uncorrelated with $s_1(n)$ and $\mathbf{v}(n)$, and have $E\left[\mathbf{u}_k\mathbf{u}_k^H\right] = \frac{1}{N}\zeta_k\mathbf{I}_N$ as covariance matrices. In this case, it is not difficult to show that the SINR at the output of the MMSE receiver for detecting $s_1(n)$ is

$$p_1\mathbf{w}^H\mathbf{H}^H\left(\left(\frac{1}{N}\sum_{k=2}^{K}\zeta_k\right)\mathbf{H}\mathbf{H}^H + \sigma^2\mathbf{I}_N\right)^{-1}\mathbf{H}\mathbf{w}$$

For large values of N, this SINR can be approximated by

$$\frac{p_1}{N}\mathrm{tr}(\mathbf{H}^H\left(\left(\frac{1}{N}\sum_{k=2}^{K}\zeta_k\right)\mathbf{H}\mathbf{H}^H + \sigma^2\mathbf{I}_N\right)^{-1}\mathbf{H}$$

Matrix $\mathbf{H}$ is circulant, and can thus be written as $\mathbf{H} = \mathbf{F}\mathbf{D}\mathbf{F}^H$. Here, $\mathbf{F}$ is the $N\times N$ Fourier matrix which (p,q) entry is $\mathbf{F}_{p,q} = \frac{1}{\sqrt{N}}e^{2i\pi pq/N}$ for $(p,q)\in\{0,\ldots,N-1\}$, and $\mathbf{D}$ is the diagonal matrix with entries $(h(e^{2i\pi l/N})_{l=0,\ldots,N-1}$. Therefore, the above SINR can be written as

$$p_1\frac{1}{N}\sum_{n=0}^{N-1}\frac{|h(e^{\frac{2i\pi n}{N}})|^2}{|h(e^{\frac{2i\pi n}{N}})|^2\frac{1}{N}\sum_{k=2}^{K}\zeta_k + \sigma^2}\ .$$

This expression approximates the right hand member of (16) if we choose $\zeta_k = \frac{p_k}{p_k\overline{\beta}_{\rm iid}+1}$. In conclusion, the MMSE receiver operates as if the covariance matrix $\mathbf{U}\mathbf{Q}\mathbf{U}^H$ of the interference term in the transmitted signal was replaced by $\frac{1}{N}\sum_{k=2}^{K}\frac{p_k}{p_k\overline{\beta}_{\rm iid}+1}\mathbf{I}_N$. This last matrix might then be seen as the effective interference covariance matrix for the given SINR target, whether the channel is frequency flat or selective.

As mentioned in the introduction, all our derivations and results are devoted to downlink transmissions over a frequency selective channel. It is interesting to compare theorem

2 with the result of [6] specialized to uplink transmissions over different frequency selective channels. The signal model (11) becomes in the uplink direction

$$\mathbf{y}(n) = \sum_{k=1}^{K} \mathbf{H}_k \mathbf{w}_k s_k(n) \ + \ \mathbf{v}(n) \tag{27}$$

where $\{\mathbf{H}_k\}_{k=1,\ldots,K}$ are the circulant matrices associated to the K channels $h_k(z) = \sum_{m=0}^{M} h_{k,m} z^{-m}$ as described above. Because the channels are now different, we drop the parameters $(p_k)_{k=1,\ldots,K}$ and include the power differences in the channels transfer functions in order to simplify the notations. For each k, vector $\mathbf{H}_k \mathbf{w}_k$ can be seen as a linear combination of shifted versions of the code vector $\mathbf{w}_k$. Under the simplifying assumption that these shifted versions are independent, [6] obtained an expression for the asymptotic SINR $\beta_{1,\text{uplink}}$ of user 1. Specifically, it appeared that the asymptotic normalized SINR

$$\overline{\beta_{\text{uplink}}} = \frac{\beta_{1,\text{uplink}}}{\sum_{m=0}^{M} |h_{1,m}|^2}$$

is identical for all users and is nearly given as the unique solution of the equation

$$\overline{\beta_{\text{uplink}}} = \frac{1}{\frac{1}{N} \sum_{k=2}^{K} \frac{\sum_{m=0}^{M} |h_{k,m}|^2}{1 + \overline{\beta_{\text{uplink}}} \sum_{m=0}^{M} |h_{k,m}|^2} + \sigma^2} \ . \tag{28}$$

In order to compare this formula with theorem 2, we remark that after (28), the limit SINR for user 1 can be written in the frequency domain

$$\beta_{1,\text{uplink}} = \int_{-1/2}^{1/2} \frac{|h_1(e^{2i\pi f})|^2}{\frac{1}{N} \sum_{k=2}^{K} \frac{\sum_{m=0}^{M} |h_{k,m}|^2}{1 + \overline{\beta_{\text{uplink}}} \sum_{m=0}^{M} |h_{k,m}|^2} + \sigma^2} \, df \ .$$

It is clear that if $h_k(z) = h_1(z)$ for each $k \geq 2$, then this formula is not in accordance with theorem 2. This is because the simplifying assumption regarding the independence of shifted versions of the code vectors is not justified in the downlink. A contrario, (28) is certainly correct if one assumes that coefficients $(h_{k,m})_{k=1,\ldots,K, m=0,\ldots,M}$ coincide with the realizations of independent (but not necessarily identically distributed) random variables, a rather common assumption. In the following, we shall give some arguments to motivate this claim. For this, we first give an equation (eq. 29) which should be nearly satisfied by the limit SINRs $(\beta_{k,\text{uplink}})_{k=1,\ldots,K}$ provided by the MMSE receivers of the K users. We mention here that a rigorous proof of this equation requires some work, and is outside the scope of this paper. A informal justification is provided in appendix C when the code vectors $(\mathbf{w_k})_{k=1,\dots,K}$ are Gaussian.

$$\beta_{1,\text{uplink}} \simeq \frac{1}{N}\sum_{l=0}^{N-1}\frac{|h_1(e^{2i\pi l/N})|^2}{\sigma^2+\frac{1}{N}\sum_{k=2}^{K}\frac{|h_k(e^{2i\pi l/N})|^2}{1+\beta_{k,\text{uplink}}}} \ . \tag{29}$$

We note that this equation corresponds to the SINR associated to the model

$$\mathbf{r}(n) = s_1(n)\mathbf{H}_1\mathbf{w}_1 + \sum_{k=2}^{K}\mathbf{H}_k\mathbf{u}_k(n) + \mathbf{v}(n) \tag{30}$$

where the $K-1$ vectors $\{\mathbf{u}_k\}$ are such that $E\left[\mathbf{u}_k\mathbf{u}_k^H\right] = \frac{1}{N}\frac{1}{1+\beta_{k,\text{uplink}}}\mathbf{I}_N$. The MMSE receiver operates as if each interfering term $\mathbf{w}_k s_k$ was a white noise of variance equal to $\frac{1}{N}$ times the effective interference $\frac{1}{1+\beta_{k,\text{uplink}}}$. Note that the effective interference produced by user k on user 1 depends on the SINR $\beta_{k,\text{uplink}}$.

Under the hypothesis that the coefficients $\{h_{k,m}\}_{k=1,\dots,K,m=0,\dots,M}$ are independent random variables, we now infer that

$$\frac{1}{N}\sum_{k=2}^{K}\frac{|h_k(e^{2i\pi l/N})|^2}{1+\beta_{k,\text{uplink}}} \simeq \frac{1}{N}\sum_{k=2}^{K}\frac{E(|h_k(e^{2i\pi l/N})|^2)}{1+\beta_{k,\text{uplink}}} \ . \tag{31}$$

Writing $|h(e^{2i\pi f})|^2$ as $|h(e^{2i\pi f})|^2 = \sum_{m=-M}^{M} r_m e^{-2i\pi m f}$, where $r_m = \sum_l h_{l+|m|}{h_l}^*$, we get immediately that $E(|h(e^{2i\pi f})|^2)$ does not depend on f, and reduces to $E(r_0) = E(\sum_{m=0}^{M}|h_{k,m}|^2)$. Hence,

$$\frac{1}{N}\sum_{k=2}^{K}\frac{|h_k(e^{2i\pi l/N})|^2}{1+\beta_{k,\text{uplink}}} \simeq \frac{1}{N}\sum_{k=2}^{K}\frac{E\left(\sum_{m=0}^{M}|h_{k,m}|^2\right)}{1+\beta_{k,\text{uplink}}} \tag{32}$$

which is itself nearly equal to

$$\frac{1}{N}\sum_{k=2}^{K}\frac{(\sum_{m=0}^{M}|h_{k,m}|^2)}{1+\beta_{k,\text{uplink}}} \ . \tag{33}$$

By plugging (32 and 33) into (29), this last equation can be rewritten

$$\beta_{1,\text{uplink}} \simeq \frac{1}{\sigma^2+\frac{1}{N}\frac{\sum_{k=2}^{K}|h_{k,m}|^2}{1+\beta_{k,\text{uplink}}}}\left(\sum_{m=0}^{M}|h_{1,m}|^2\right)$$

Notice that this equation remains true if user 1 is replaced by any other user and that for every user k, $\beta_{k,\text{uplink}}/\left(\sum_{m=0}^{M}|h_{k,m}|^2\right)$ does not depend on this user in the asymptotic regime. In conclusion, in this regime $\beta_{k,\text{uplink}} = \overline{\beta_{\text{uplink}}}\sum_{m=0}^{M}|h_{k,m}|^2$ where $\overline{\beta_{\text{uplink}}}$ is given by equation (28).

Discussion of Theorem 1. We first compare the two SINR given by (15) and (16). It is clear that for each $\overline{\beta}$,

$$\int_{-1/2}^{1/2} \frac{df}{\alpha \sum_{l=1}^{L} \rho_l \frac{P_l}{P_l\overline{\beta}+1} + \frac{\sigma^2}{|h(e^{2i\pi f})|^2}\left(1-\alpha \sum_{l=1}^{L} \rho_l \frac{P_l\overline{\beta}}{P_l\overline{\beta}+1}\right)} \geq \int_{-1/2}^{1/2} \frac{df}{\alpha \sum_{l=1}^{L} \rho_l \frac{P_l}{P_l\overline{\beta}+1} + \frac{\sigma^2}{|h(e^{2i\pi f})|^2}}$$

Therefore, $\overline{\beta}$ is greater than $\overline{\beta}_{\text{iid}}$, in confirmation of the fact that the use of an isometric code matrix improves the performance of the MMSE detector on a frequency selective channel. Moreover, for a given SINR target $\overline{\beta}$, (15) and (16) show that a system corrupted by a background noise of variance σ^2 using an isometric code matrix provides the same performance as an i.i.d. one corrupted by a noise of variance

$$\sigma^2 \left(1-\alpha \sum_{l=1}^{L} \rho_l \frac{P_l\overline{\beta}}{P_l\overline{\beta}+1}\right) .$$

In other words, the effective interference term introduced in the discussion of Theorem 2 is not modified by the use of an isometric code matrix. However, an isometric code matrix reduces in some sense the noise variance to the "effective noise" variance given above. This expression provides directly the gain on the signal to noise ratio $\frac{E_b}{N_0}$ resulting from the use of isometric codes instead of i.i.d. ones :

$$\gamma_{dB} = -10\log_{10}\left(1-\alpha \sum_{l=1}^{L} \rho_l \frac{P_l\overline{\beta}}{P_l\overline{\beta}+1}\right)$$

We note in particular that the attenuation factor is all the more favorable that α is close to 1.

Numerical illustration. Here, the theoretical performance of the optimum MMSE receiver with isometric codes spreading is compared to the performance of the same receiver with i.i.d. codes spreading, based on the results of theorems 1 and 2 respectively. The figure of merit will be the theoretical Bit Error Rate (BER), given by $Q(\sqrt{SINR})$ when a QPSK constellation is used, as it will be the case. $Q(x)$ is of course defined by

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-\frac{t^2}{2}}\, dt$$

In figure 1, we assume that all the users have the same power, and compare the performance obtained with an isometric spreading matrix to the one obtained with an i.i.d. spreading matrix. The load α has been chosen equal to 1/4, 1/2, and 1. The plots confirm the fact that the difference between the two cases is all the more significant that α is close from 1. In contrast, the two BERs for $\alpha = \frac{1}{4}$ are rather close. In this experiment, the propagation channel is the so-called vehicular A channel. This is a multiple path Rayleigh fading channel whose time delays are $0, 1.2T_c, 2.8T_c, 4.2T_c, 6.6T_c$ and $9.6T_c$. The variances of the corresponding complex amplitudes are equal to 0 dB, -1 dB, -9 dB, -10 dB, -15 dB, and -20 dB, and the BER represented in figure 1 are obtained by averaging $Q(\sqrt{SINR})$ on 10 realizations of the complex amplitudes. Finally, the shaping filter is a square root raised cosine with a roll-off factor of 0.22.

Adaptation of the results to the context of time-varying codes. Let us give now some remarks on these results in the case where code matrices $\mathbf{W}$ are time-varying. In order to mention explicitly this time-dependence, they are denoted $\mathbf{W}(n)$ in this paragraph, and we assume that the sequence of $N \times K$ matrices $(\mathbf{W}(n))_{n\in\mathbb{Z}}$ is (Haar) identically distributed and ergodic. In this context, the various SINRs of course also depend on n. The reader may check that the proof of Proposition 1 can be extended to this context, and that $\tilde{\beta}_N(n)$ shows the same asymptotic behavior as the SINR $\beta_N(n)$ defined by

$$\beta_N(n) = p_1 \mathbf{w}^H(n)\mathbf{H}^H \left(\mathbf{H}\mathbf{U}(n)\mathbf{Q}\mathbf{U}(n)^H\mathbf{H}^H + \sigma^2\mathbf{I}_N\right)^{-1} \mathbf{H}\mathbf{w}(n) \tag{34}$$

and associated to the output MMSE receiver $\hat{s}_1(n)$ of model

$$\mathbf{y}(n) = \mathbf{H}\mathbf{W}(n)\sqrt{\mathbf{P}}\mathbf{s}(n) + \mathbf{v}(n) \tag{35}$$

Under the hypothesis that the channel coefficients $(h_k)_{k=0,\ldots,M}$ remain constant over a duration large enough, the performance of the MMSE receiver is of course characterized by $E_{\mathbf{W}}(\frac{1}{\beta_N(n)})$, where $E_{\mathbf{W}}(.)$ is the mathematical expectation with respect to the Haar distribution. Theorem 1 shows that for each n, $\beta_N(n)$ converges almost surely toward β defined by (15). If we assume that $|h(e^{2i\pi f})| > 0$ for each f, then it is easy to check that $\frac{1}{\beta_N(n)}$ is upper-bounded by a deterministic constant term. The Lebesgue dominated convergence theorem thus implies that $E_{\mathbf{W}}(\frac{1}{\beta_N(n)})$ converges toward $\frac{1}{\beta}$. It turns out that, like in the time-invariant case, the performances of the MMSE receiver are completely characterized by β as $N \to +\infty$ and $\frac{K}{N} \to \alpha < 1$.

We now tackle the question related to the values of N for which our asymptotic analysis is likely to provide reliable performance evaluations. In the time invariant case, the asymptotic analysis is relevant if N is chosen in such a way that $\frac{1}{\beta_N} \simeq \frac{1}{\beta}$. One can intuitively suppose that in the time varying case, the condition $E_{\mathbf{W}}(\frac{1}{\beta_N(n)}) \simeq \frac{1}{\beta}$ will be satisfied for smaller values of N, thanks to the averaging effect w.r.t. the values of the code matrices. This claim will be justified by the simulations presented in section IV in which it is shown that, in the time-varying code vectors case, our asymptotic results are quite reliable for $N = 256$.

IV. The sub-optimum MMSE receiver.

In this section, we address the case where matrices $\mathbf{U}$ and $\mathbf{Q}$ are not available. In this context, it is often difficult to obtain reliable estimates of the covariance matrix $\mathbf{HWPW}^H\mathbf{H}^H + \sigma^2\mathbf{I}_N$ directly from the observation. It is natural in these conditions to study the receiver consisting in a chip rate filter that equalizes the transfer function $h(z)$ shown in (2), followed by despreading. Denote by $\overline{p}^{(K)} = \frac{1}{K}\sum_{k=1}^{K} p_k$ the mean of the power distribution allocated to the various users of the cell. The power of the received signal, defined as $\lim_{T\to+\infty} \frac{1}{T}(\sum_{m=0}^{T-1} |y(m)|^2)$, is given by $\frac{K}{N}(\sum_{k=0}^{M} |h_k|^2)\overline{p}^{(K)} + \sigma^2$. If σ^2 is known, it is therefore straightforward to estimate consistently $\frac{K}{N}\overline{p}^{(K)}$. This justifies that it is relevant to assume that $\frac{K}{N}\overline{p}^{(K)}$ is known at the mobile station side although the $(p_k)_{k\geq 2}$ and $\frac{K}{N}$ are not. The sub-optimal Wiener filter we consider here is derived under the assumption that the chip sequences $(x_k)_{k=1,\ldots,K}$ generated by the various users are uncorrelated white sequences with variance $\frac{1}{N}\overline{p}^{(K)}$. Based on this assumption, the $N \times N$ matrix $\overline{\mathbf{G}}$ minimizing

$$E\|\mathbf{x}(n) - \overline{\mathbf{G}}\mathbf{y}(n)\|^2$$

is equal to (see formula (3))

$$\overline{\mathbf{G}} = \mathcal{H}_0^H(\mathcal{H}_0\mathcal{H}_0^H + \mathcal{H}_1\mathcal{H}_1^H + \frac{\sigma^2}{\frac{K}{N}\overline{p}^{(K)}}\mathbf{I})^{-1} .$$

Vector $\overline{\mathbf{G}}\mathbf{y}(n)$ thus estimates vector $\mathbf{x}(n)$, so that the action of $\overline{\mathbf{G}}$ on $\mathbf{y}(n)$ is equivalent to the action of a chip-rate equalizer $g(z)$ on signal $y(m)$. The sub-optimum MMSE receiver consists in despreading $\overline{\mathbf{G}}\mathbf{y}(n)$, i.e. $s_1(n)$ is estimated from

$$\mathbf{w}^H\overline{\mathbf{G}}\mathbf{y}(n) \ . \tag{36}$$

It is possible to show that as $N \to \infty$ and $\frac{K}{N} \to \alpha < 1$, the SINR corresponding to this sub-optimal receiver has the same behavior as the SINR associated to the action of the receiver $\mathbf{w}^H\mathbf{H}^H(\mathbf{H}\mathbf{H}^H + \frac{\sigma^2}{\frac{K}{N}\overline{p}^{(K)}}\mathbf{I}_N)^{-1}$ on vector $\mathbf{y}(n)$ defined by (11). The proof of this statement is similar to the proof of Proposition 1, and is thus omitted.

We therefore propose to evaluate the performance of the sub-optimum MMSE receiver described by the equation

$$\overline{s}_1(n) = \mathbf{w}^H\mathbf{H}^H(\mathbf{H}\mathbf{H}^H + \frac{\sigma^2}{\frac{K}{N}\overline{p}^{(K)}}\mathbf{I})^{-1}(\mathbf{H}\mathbf{W}\sqrt{\mathbf{P}}\mathbf{s}(n) + \mathbf{v}(n)) \tag{37}$$

We first evaluate the SINR, denoted $\beta_{\text{chip},N}$, associated to this receiver. Using (37), we immediately obtain

$$\beta_{\text{chip},N} = p_1 \frac{(\overline{\eta}_N)^2}{\overline{\gamma}_N} \tag{38}$$

where

$$\overline{\eta}_N = \mathbf{w}^H\mathbf{H}^H\left(\mathbf{H}\mathbf{H}^H + \frac{\sigma^2}{\frac{K}{N}\overline{p}^{(K)}}\mathbf{I}\right)^{-1}\mathbf{H}\mathbf{w} \tag{39}$$

$$\overline{\gamma}_N = \mathbf{w}^H\mathbf{H}^H\left(\mathbf{H}\mathbf{H}^H + \frac{\sigma^2}{\frac{K}{N}\overline{p}^{(K)}}\mathbf{I}\right)^{-1}\left(\mathbf{H}\mathbf{U}\mathbf{Q}\mathbf{U}^H\mathbf{H}^H + \sigma^2\mathbf{I}\right)\left(\mathbf{H}\mathbf{H}^H + \frac{\sigma^2}{\frac{K}{N}\overline{p}^{(K)}}\mathbf{I}\right)^{-1}\mathbf{H}\mathbf{w} \tag{40}$$

In order to study the behavior of $\beta_{\text{chip},N}$ when N and K converge to $+\infty$ and $\frac{K}{N} \to \alpha$, we shall consider $\overline{\eta}_N$ and $\overline{\gamma}_N$ separately, and begin by investigating the behavior of $\overline{\eta}_N$. In the following, we denote by $\overline{p} = \sum_{l=1}^{L} \rho_l P_l = \lim_{K\to+\infty} \overline{p}^{(K)}$ the average of the limit power distribution.

Proposition 2: Under assumptions **A1** to **A3**, $\overline{\eta}_N$ converges almost surely to $\overline{\eta}$ defined by

$$\overline{\eta} = \int_{-1/2}^{1/2} \frac{|h(e^{2i\pi f})|^2}{|h(e^{2i\pi f})|^2 + \frac{\sigma^2}{\overline{p}\alpha}}\, df \tag{41}$$

Proof. The result follows immediately from Lemma 1 : just use this result and use the fact that the eigenvalues of matrix $\mathbf{H}^H(\mathbf{H}\mathbf{H}^H+\frac{\sigma^2}{K/N\overline{p}^{(K)}}\mathbf{I})^{-1}\mathbf{H}$ are the $(\frac{|h(e^{2i\pi l/N})|^2}{|h(e^{2i\pi l/N})|^2+\frac{\sigma^2}{K/N\overline{p}^{(K)}}})_{l=0,\ldots,\frac{N-1}{N}}$.

$\overline{\gamma}_N$ can also be written as $\mathbf{w}^H\mathbf{B}\mathbf{w}$ for a certain uniformly bounded matrix $\mathbf{B}$. However, $\mathbf{B}$ is not independent of $\mathbf{w}$ in the isometric case.

Proposition 3: Under assumptions **A1** to **A3**, $\overline{\gamma}_N$ converges almost surely toward a deterministic constant $\overline{\gamma}$. Moreover,

$$\overline{\gamma}-\alpha\overline{p}\mathbf{w}^H\mathbf{H}^H(\mathbf{H}\mathbf{H}^H+\frac{\sigma^2}{\alpha\overline{p}}\mathbf{I})^{-1}(\mathbf{H}(I-\mathbf{w}\mathbf{w}^H)\mathbf{H}^H+\frac{\sigma^2}{\alpha\overline{p}}\mathbf{I})(\mathbf{H}\mathbf{H}^H+\frac{\sigma^2}{\alpha\overline{p}}\mathbf{I})^{-1}\mathbf{H}\mathbf{w}\rightarrow 0 \quad (42)$$

in the least mean-square sense when N and K converge to ∞ and $\frac{K}{N}\rightarrow\alpha<1$.

Proof. See appendix.

The second term of the left hand side of (42) clearly converges almost surely toward $\alpha\overline{p}\,\overline{\eta}(1-\overline{\eta})$. This in turn shows that $\overline{\gamma}_N$ converges almost surely toward $\alpha\overline{p}\,\overline{\eta}(1-\overline{\eta})$. Therefore, we have proved the following result :

Theorem 3: Under assumptions **A1** to **A3**, the SINR $\beta_{\text{chip},N}$ converges almost surely to the quantity β_{chip} defined by

$$\beta_{\text{chip}}=\frac{p_1}{\alpha\overline{p}}\frac{\overline{\eta}}{1-\overline{\eta}} \quad (43)$$

Let us comment this result. If we consider the classical Single Input Single Output (SISO) signal model where the channel impulse response is $h(z)$, the useful signal power is $\alpha\overline{p}$ and the noise power is σ^2, then it is known that the MMSE receiver output SINR is $\overline{\eta}/(1-\overline{\eta})$. The extra factor $\frac{p_1}{\alpha\overline{p}}$ in (43) can thus be interpreted as the SINR gain provided by the despreading.

We furthermore remark that $\overline{\eta}$ depends on the power limit distribution through its average $\overline{p}$ only. This shows that the effect of a non uniform power distribution on β_{chip} only depends on the ratio $\frac{p_1}{\overline{p}}$, and not on the particular form of the limit distribution.

We now compare the performances of the optimum and sub-optimum receivers. We first consider the case where the power distribution is uniform, all powers being equal to P. In this case, we trivially have $\overline{p}=P$, and we denote the corresponding value of β_{chip} by $\underline{\beta}_{upd}$ (the subscript "upd" stands for uniform power distribution). Accordingly, let

$\beta_{upd} = \overline{\beta} P$, where $\overline{\beta}$ is given by equation (15), be the limit SINR provided by the optimum Wiener filter of subsection III. (43) leads immediately to

$$\underline{\beta}_{upd} = \frac{1}{\alpha} \frac{\overline{\eta}}{1 - \overline{\eta}} \ . \tag{44}$$

In order to compare the performance of the optimum to that of the sub-optimum MMSE receiver, we remark that

$$\frac{\underline{\beta}_{upd}}{\alpha \underline{\beta}_{upd} + 1} = \frac{\overline{\eta}}{\alpha} = \int_{-1/2}^{1/2} \frac{P|h(e^{2i\pi f})|^2}{\alpha P |h(e^{2i\pi f})|^2 + \sigma^2} \, df$$

It turns out that $\underline{\beta}_{upd}$ may thus be interpreted as the unique solution of the equation

$$\underline{\beta}_{upd} = \int_{-1/2}^{1/2} \frac{P|h(e^{2i\pi f})|^2}{\alpha |h(e^{2i\pi f})|^2 \frac{P}{\alpha \underline{\beta}_{upd} + 1} + \frac{\sigma^2}{(\alpha \underline{\beta}_{upd} + 1)}} \, df \ . \tag{45}$$

On the other hand, the optimum MMSE limit SINR β_{upd} is defined as the solution of

$$\beta_{upd} = \int_{-1/2}^{1/2} \frac{P|h(e^{2i\pi f})|^2}{\alpha |h(e^{2i\pi f})|^2 \frac{P}{\beta_{upd} + 1} + \sigma^2 \left(1 - \frac{\alpha \beta_{upd}}{\beta_{upd} + 1}\right)} \, df \ . \tag{46}$$

We first notice that the two expressions coincide when $\alpha \to 1$. When $\alpha < 1$, we remark that for a given target SINR $\underline{\beta}_{upd}$, the effective interference term $\frac{P}{\alpha \underline{\beta}_{upd} + 1}$ is less favorable than in formula (46) because $\alpha < 1$. However, the term $\frac{1}{\alpha \underline{\beta}_{upd} + 1} < 1$ attenuating the variance σ^2 in (45) is more favorable than the corresponding term $(1 - \frac{\alpha \beta_{upd}}{\beta_{upd} + 1})$ in (46). Yet, $\underline{\beta}_{upd}$ is of course smaller than β_{upd}. As the corresponding formulas are difficult to interpret, we only resort to numerical simulations to compare $\overline{\beta} p_1$ with β_{chip} in the non uniform power distribution case.

One should also remark that the sub-optimum Wiener filter is in principle still near far resistant, i.e. it is able to cancel perfectly the multi-user interference if $\sigma^2 \to 0$. To see this, notice that $\underline{\beta}_{upd}$ satisfies

$$\frac{\underline{\beta}_{upd}}{\alpha \underline{\beta}_{upd} + 1} = \int_{-1/2}^{1/2} \frac{P|h(e^{2i\pi f})|^2}{\alpha P |h(e^{2i\pi f})|^2 + \sigma^2} \, df \tag{47}$$

When $\sigma^2 \to 0$, then the right hand side of (47) converges to $\frac{1}{\alpha}$, a condition which implies that $\underline{\beta}_{upd}$ converges to $+\infty$. We note that this property is not surprising because the knowledge of the codes of the other users is not necessary to design near far resistant receivers in the downlink if the code matrix is isometric : a simple chip rate zero forcing equalizer (equivalent to the inversion of matrix $\mathbf{H}$), followed by a correlation with the code of the desired user, is near far resistant. However, as shown in the simulations below, the performance of the optimum MMSE receiver is nearly independent of the power p_1 allocated to the user of interest, while that of the sub-optimum MMSE receiver depends strongly on the ratio $\frac{p_1}{\overline{p}}$.

Adaptation of the results to the context of time-varying codes. We just mention that the above results are still valid when code vectors are time-varying. As in section III, the performance of the sub-optimum MMSE receiver is characterized by $E_{\mathbf{W}}(\frac{1}{\beta_{\text{chip},N}})$, which converges toward $\frac{1}{\beta_{\text{chip}}}$.

Numerical simulations. We first check that our theoretical results (based in particular on the assumption that the code matrix is obtained from a Haar distributed random unitary matrix) allow to predict accurately the performance of a real-life CDMA system. For this, we simulated the downlink of the wide-band CDMA mode of the UMTS (see the specifications [16] for more details). In this context, the codes allocated to the users are obtained by multiplying time-invariant Walsh-Hadamard codes with a cell specific time-varying QPSK scrambling code whose period is equal to 150 symbols. The corresponding matrices $\mathbf{W}(n)$ are thus time-varying. Note also that the symbols transmitted by the base station are QPSK symbols.

We consider the case where $N = 256$, $K = 64$ ($\alpha = \frac{1}{4}$) or $K = 128$ ($\alpha = \frac{1}{2}$), and assume that there exists $K_c = 5$ classes of users with a limit power distribution given by $(\alpha_1, \ldots, \alpha_5) = (5/16, 1/4, 13/64, 11/64, 1/16)$ and $(p_1, \ldots, p_5) = (P, 2P, 4P, 8P, 16P)$. Like in section III, the propagation channel is the vehicular A channel, and the shaping filter is a square root raised cosine. In this case, one can check that the degree M of filter $h(z)$ is nearly equal to 10. The complex amplitudes of the channels remain constant on each frame of 150 QPSK symbols, and differ from one frame to another frame. The empirical BERs are averaged on 350 frames. In order to check the validity of our theoretical SINRs formulas we have evaluated, for each user of class 2, the empirical BERs provided by the optimum and sub-optimum Wiener filters defined by formula (8) and (36). These empirical results are then compared to the BERs given by formulas $Q(\sqrt{\beta p_k})$ (optimum Wiener filter) and $Q(\sqrt{\beta_{\text{chip}}})$ (sub-optimum Wiener filter). Figures 2 ($K = 64$, i.e. $\alpha = \frac{1}{4}$) and 3 ($K = 128$, i.e. $\alpha = \frac{1}{2}$) allow to compare the theoretical performances with the empirical results. These results show that our formulas allow to predict quite well the performances of the downlink of the wide band CDMA mode of the UMTS, but that a certain dispersion occurs when $\frac{E_b}{N_0}$ increases. This observation is in accordance with [17] in which it is shown, in a simple case, that the variance of the difference between the SINR and its asymptotic expression increases when $\frac{E_b}{N_0}$ increases.

We now compare the theoretical performances of the optimum with that of the sub-optimum Wiener filter. For this to be done, we first consider the uniform power distribution case, and represent in figure 4 the theoretical BERs provided by the two receivers for $\alpha = 1, \alpha = \frac{1}{2}$, and $\alpha = \frac{1}{4}$. The channel is still a realization of the vehicular A. As mentioned above, the optimum and sub-optimum Wiener detectors provide the same performance when $\alpha = 1$. However, the differences between the two receivers tend to increase if $\alpha = \frac{1}{2}$ and $\alpha = \frac{1}{4}$. We finally compare the effect of a non uniform power distribution on the behavior of the optimum and sub-optimum Wiener filter. Here we assume that $\alpha = \frac{1}{4}$, and that the distribution of the powers is the same as in the previous experiments. We represent in figure 5 the theoretical BER provided by the two receivers for each class. Figure 5 shows that the performance of the optimum Wiener filter is nearly independent of the user's class. However, this is not at all the case when the sub-optimum receiver is implemented. In particular, for a target BER of 10^{-2}, we observe a loss of performance of 4 dB between class 1 and class 5.

V. Conclusion

We studied the performance of two linear receivers acting on CDMA signals transmitted in the downlink direction over a frequency selective channel with unequal power allocation. The first receiver is the optimum MMSE receiver and the second one is a chip rate equalizer followed by despreading. Spreading codes are modeled as random variables and the analysis is made in the asymptotic regime where the spreading factor and the number of users grow toward infinity at the same rate. Both isometric code matrices and code matrices with i.i.d. elements are considered. Asymptotic expressions for the SINRs at the outputs of these receivers are derived. It appears that these quantities depend only on the channel transfer function, the power distribution and the asymptotic ratio α between the number of active users and the spreading factor.

In particular, the following phenomena are quantified in a precise manner: when the optimum MMSE receiver is used, the performance gain which results from the use of isometric spreading codes rather than i.i.d. ones grows as $\alpha \to 1$.

While the two receivers provide the same performance when an isometric code matrix is used and $\alpha = 1$, the performance loss induced by the use of the sub-optimum receiver increases as $\alpha \to 0$.

Finally, even though the sub-optimum receiver is in principle near-far resistant when the codes are isometric, the low-power users are much more penalized by this receiver than by the optimum MMSE receiver.

Acknowledgment

The authors would like to thank Prof. David Tse for his insightful remarks concerning the interpretation of theorems 1 and 2.

Appendix

I. Proof of Proposition 1

We shall prove in this appendix that under **A1** and **A2**, the difference between the SINR $\tilde{\beta}_N$ given by (9) and the more tractable expression β_N given by (13) converges to zero almost surely. Writing for conciseness

$$\begin{aligned} \mathbf{R} &= \mathbf{HUQU}^H\mathbf{H}^H + \sigma^2\mathbf{I}_N, && (48)\\ \mathbf{T} &= -\mathcal{H}_0\mathbf{UQU}^H\mathcal{H}_1^H - \mathcal{H}_1\mathbf{UQU}^H\mathcal{H}_0^H, \text{ and} && (49)\\ \mathbf{z} &= \sqrt{p_1}\mathcal{H}_1\mathbf{w}\ , && (50) \end{aligned}$$

we clearly have

$$\tilde{\beta}_N = p_1\mathbf{w}^H(\mathbf{H} - \mathcal{H}_1)^H\left(\mathbf{R} + \mathbf{T} + \mathbf{zz}^H\right)^{-1}(\mathbf{H} - \mathcal{H}_1)\mathbf{w} \qquad (51)$$

and $\beta_N = p_1 \mathbf{w}^H \mathbf{H}^H \mathbf{R}^{-1} \mathbf{H} \mathbf{w}$. Before going into the proof, some remarks are in order. First, $\mathcal{H}_0$ and $\mathcal{H}_1$ are both uniformly bounded. To see this, let $\mathcal{H} = [\mathcal{H}_0^T \ \mathcal{H}_1^T]^T$ and notice that $\mathcal{H}^H\mathcal{H}$ is a Toeplitz matrix associated to the spectral density $|h(e^{2i\pi f})|^2$. As a consequence, $\|\mathcal{H}\| \leq \max_f |h(e^{2i\pi f})|$. Because $\|\mathcal{H}_k\| \leq \|\mathcal{H}\|$ for $k = 0$ and 1, these matrices are uniformly bounded. Second, $\mathbf{R} + \mathbf{T} = \mathcal{H}_0 \mathbf{U}\mathbf{Q}\mathbf{U}^H \mathcal{H}_0^H + \mathcal{H}_1 \mathbf{U}\mathbf{Q}\mathbf{U}^H \mathcal{H}_1^H + \sigma^2 \mathbf{I}_N$ is also uniformly bounded: use the inequalities $\|\mathbf{X} + \mathbf{Y}\| \leq \|\mathbf{X}\| + \|\mathbf{Y}\|$ and $\|\mathbf{X}\mathbf{Y}\| \leq \|\mathbf{X}\|\|\mathbf{Y}\|$ repeatedly, and notice that $\|\mathbf{Q}\| < \infty$ thanks to **A2**. Moreover, $\|(\mathbf{R} + \mathbf{T})^{-1}\|$ is upper bounded by $1/\sigma^2$. Remember that $\mathbf{H}$ is also uniformly bounded (see (14)).

As a first step, we shall prove that the difference between $\tilde{\beta}_N$ and

$$\tilde{\beta}_N^{(1)} = p_1 \mathbf{w}^H (\mathbf{H} - \mathcal{H}_1)^H \left(\mathbf{R} + \mathbf{T}\right)^{-1} (\mathbf{H} - \mathcal{H}_1)\mathbf{w}$$

converges almost surely to 0, in other words, the term $\mathbf{z}\mathbf{z}^H$ in (51) can be neglected. By the matrix inversion lemma,

$$(\mathbf{R} + \mathbf{T} + \mathbf{z}\mathbf{z}^H)^{-1} - (\mathbf{R} + \mathbf{T})^{-1} = -(1 + \mathbf{z}^H(\mathbf{R} + \mathbf{T})^{-1}\mathbf{z})^{-1} \ (\mathbf{R} + \mathbf{T})^{-1}\mathbf{z}\mathbf{z}^H(\mathbf{R} + \mathbf{T})^{-1} \ .$$

The scalar term $(1 + \mathbf{z}^H(\mathbf{R} + \mathbf{T})^{-1}\mathbf{z})^{-1}$ of the right hand member is upper bounded by 1, so

$$\|(\mathbf{R} + \mathbf{T} + \mathbf{z}\mathbf{z}^H)^{-1} - (\mathbf{R} + \mathbf{T})^{-1}\| \leq \|(\mathbf{R} + \mathbf{T})^{-1}\|^2 \|\mathbf{z}\|^2 \leq (1/\sigma^4)\|\mathbf{z}\|^2$$

and we need to prove that $\|\mathbf{z}\|^2 \to 0$ almost surely. The code vector being $\mathbf{w} = (w_1^{(0)}, \ldots, w_1^{(N-1)})^T$, $\mathbf{z}$ has the form $\mathbf{z} = (z_N^{(1)}, \ldots, z_N^{(M)}, 0, \ldots, 0)^T$, where a non zero element is written $z_N^{(m)} = \sqrt{p_1} \sum_{i=0}^{M-m} h_{M-i} w_1^{(N-M+i+m-1)}$. Every one of these elements is thus a finite weighted sum of elements of a Haar distributed unitary matrix. By Minkowski's inequality,

$$\sqrt[4]{E[|z_N^{(m)}|^4]} \leq \sum_{i=0}^{M-m} \sqrt[4]{{p_1}^2 |h_{M-i}|^4 \, E[|w_1^{(N-M+i+m-1)}|^4]} \ .$$

As an element of a $N \times N$ Haar unitary random matrix, w_1^k satisfies $E[|w_1^{(k)}|^4] = 2/(N(N+1))$ (see [15, chap.4]), and moreover, $p_1 < \infty$ by **A2**. It results that $E[|z_N^{(m)}|^4] = O(1/N^2)$. Now, Markov's inequality implies that

$$\forall \, \epsilon > 0, \ \mathbb{P}\left(|z_N^{(m)}|^2 > \epsilon\right) \leq \frac{E\left[\left|z_N^{(m)}\right|^4\right]}{\epsilon^2} = O(N^{-2}) \ .$$

Therefore, $\sum_{N=1}^{\infty} \mathbb{P}\left(|z_N^{(m)}|^2 > \epsilon\right) < \infty$ and the fact that $|z_N^{(m)}|^2 \to 0$ a.s. follows from the Borel-Cantelli lemma. So, $\|\mathbf{z}\|^2 = \sum_{m=1}^{M} |z_N^{(m)}|^2 \to 0$ almost surely. Writing

$$\begin{aligned} |\tilde{\beta}_N - \tilde{\beta}_N^{(1)}| &\leq p_1 \|(\mathbf{R}+\mathbf{T}+\mathbf{z}\mathbf{z}^H)^{-1} - (\mathbf{R}+\mathbf{T})^{-1}\| \, \|(\mathbf{H}-\mathcal{H}_1)\mathbf{w}\|^2 \\ &\leq p_1 \|(\mathbf{R}+\mathbf{T}+\mathbf{z}\mathbf{z}^H)^{-1} - (\mathbf{R}+\mathbf{T})^{-1}\| \, (\|\mathbf{H}\| + \|\mathcal{H}_1\|)^2 , \end{aligned}$$

the term $\|\mathbf{H}\| + \|\mathcal{H}_1\|$ is bounded, so $\tilde{\beta}_N - \tilde{\beta}_N^{(1)} \to 0$ almost surely.

The second step consists in proving that $\tilde{\beta}_N^{(1)} - \beta_N \to 0$ almost surely. Writing $\mathbf{\Pi} = (\mathbf{I} - \mathbf{U}\mathbf{U}^H)$ and

$$\tilde{\mathbf{A}} = \mathbf{\Pi}(\mathbf{H} - \mathcal{H}_1)^H \left(\mathbf{R}+\mathbf{T}\right)^{-1} \left(\mathbf{H} - \mathcal{H}_1\right) ,$$

and adapting proposition 4 below, we get that $\tilde{\beta}_N^{(1)} - p_1 \mathrm{tr}(\tilde{\mathbf{A}})/(N-K) \to 0$ almost surely. By the same argument, $\beta_N - p_1\mathrm{tr}(\mathbf{A})/(N-K) \to 0$ a.s., where $\mathbf{A} = \mathbf{\Pi}\mathbf{H}^H\mathbf{R}^{-1}\mathbf{H}$ is given by (17). Let us develop the expression of $\tilde{\mathbf{A}}$. Looking at the expression (49) of $\mathbf{T}$, one notices that the rank of this matrix does not exceed $2M$. This is because the rank of both terms is upper bounded by the rank of $\mathcal{H}_1$ which is M. It is therefore possible to factor $\mathbf{T}$ as $\mathbf{T} = \mathbf{C}_N\mathbf{D}_N^H$, where $\mathbf{C}_N$ and $\mathbf{D}_N$ are $N \times 2M$ matrices. Applying the matrix inversion lemma, we have

$$\left(\mathbf{R}+\mathbf{T}\right)^{-1} = \mathbf{R}^{-1} - \mathbf{R}^{-1}\mathbf{C}\left(\mathbf{I}_{2M} + \mathbf{D}^H\mathbf{R}^{-1}\mathbf{C}\right)^{-1}\mathbf{D}^H\mathbf{R}^{-1}$$

hence we can write $\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{X} + \mathbf{Y}$, where

$$\mathbf{X} = -\mathbf{\Pi}(\mathbf{H}-\mathcal{H}_1)^H\mathbf{R}^{-1}\mathbf{C}\left(\mathbf{I} + \mathbf{D}^H\mathbf{R}^{-1}\mathbf{C}\right)^{-1}\mathbf{D}^H\mathbf{R}^{-1}(\mathbf{H}-\mathcal{H}_1)$$

and

$$\mathbf{Y} = \mathbf{\Pi}\left(\mathcal{H}_1^H\mathbf{R}^{-1}\mathcal{H}_1 - \mathcal{H}_1^H\mathbf{R}^{-1}\mathbf{H} - \mathbf{H}^H\mathbf{R}^{-1}\mathcal{H}_1\right) .$$

$\mathbf{X}$ is a.s. uniformly bounded because each of its factors is, and furthermore, its rank is upper bounded by $2M$. It follows that $|\mathrm{tr}(\mathbf{X})|/(N-K) \leq 2M\|\mathbf{X}\|/(N-K) = O(1/N)$, thus $|\mathrm{tr}(\mathbf{X})|/(N-K) \to 0$ a.s. One can easily show by a similar argument that $|\mathrm{tr}(\mathbf{Y})|/(N-K) \to 0$ a.s. It results that $\mathrm{tr}(\tilde{\mathbf{A}})/(N-K) - \mathrm{tr}(\mathbf{A})/(N-K) \to 0$, thus $\tilde{\beta}_N^{(1)} - \beta_N \to 0$ a.s., which ends the proof.

II. Proof of Theorem 1.

A. Proof of the claims of the first step.

In order to prove theorem 1, we first establish in proposition 4 that $\frac{\beta_N}{p_1} - \frac{tr(\mathbf{A})}{N-K}$, where $\mathbf{A}$ is defined in (17), converges to zero almost surely. The result of proposition 4 can be seen as a generalization to the context of isometric random matrices of corollary 1 in [6]. Next, we use asymptotic freeness results to show that $\frac{tr(\mathbf{A})}{N-K}$ converges almost surely toward a deterministic constant.

We first recall the following useful result of [12] and [6] :

Lemma 1: Let $\mathbf{z}$ be a $N \times 1$ random vector and $\mathbf{B}$ a $N \times N$ random matrix independent of $\mathbf{z}$. Assume that the elements of $\mathbf{z}$ are i.i.d. and have a unit variance and a finite eighth order moment and that $\sup_{N \in \mathbb{N}} \|\mathbf{B}\| < +\infty$. Denote by ξ_N the random variable defined by

$$\xi_N = \frac{1}{N} \left(\mathbf{z}^H \mathbf{B} \mathbf{z} \ - \ \mathrm{tr}(\mathbf{B}) \right)$$

Then,

$$E\left(|\xi_N|^4 \right) \leq C/N^2 \tag{52}$$

where C is independent of N.

We now show the following result.

Proposition 4: Assume that $K/N \to \alpha$ when $N \to +\infty$. Then,

$$\lim_{N \to +\infty} \left(\frac{\beta_N}{p_1} - \frac{\mathrm{tr}(\mathbf{A})}{N-K} \right) = 0 \ \text{ a.s.} \tag{53}$$

Proof: Put $e_N = \frac{\beta_N}{p_1} - \frac{\mathrm{tr}(\mathbf{A})}{N-K}$. In order to establish (53),it is sufficient to show that if $K/N \to \alpha$, then,

$$E(|e_N|^4) = O(N^{-2}) \tag{54}$$

Indeed, if (54) holds, Markov's inequality implies that

$$\forall \ \epsilon > 0, \ \mathbb{P}\left(|e_N| > \epsilon \right) \leq \frac{E\left(|e_N|^4 \right)}{\epsilon^4} = O(N^{-2}) \ .$$

Therefore,

$$\sum_{N=1}^{\infty} \mathbb{P}\left(|e_N| > \epsilon \right) < \infty$$

and (53) follows from the Borel-Cantelli lemma .

We now establish (54). Recall that we have assumed that $\mathbf{w}$ is the column 1 of $\mathbf{W}$. $e_N = \frac{\beta_N}{p_1} - \frac{\mathrm{tr}(\mathbf{A})}{N-K}$ is a function $f_1(\mathbf{W})$ of the random matrix $\mathbf{W}$. Let $\mathbf{S}$ be the permutation matrix exchanging column 1 with column K, and denote by $\tilde{\mathbf{W}}$ the matrix $\tilde{\mathbf{W}} = \mathbf{WS}$. It is obvious that $f_1(\mathbf{W}) = f_K(\tilde{\mathbf{W}})$. Moreover, $\tilde{\mathbf{W}}$ and $\mathbf{W}$ have the same distribution. Random variables $f_K(\tilde{\mathbf{W}})$ and $f_K(\mathbf{W})$ are thus identically distributed, which implies that

$$E|f_1(\mathbf{W})|^4 = E|f_K(\mathbf{W})|^4$$

Therefore, there is no restriction to assume in the proof of identity (54) that $\mathbf{w}$ is the column K of $\mathbf{W}$, a condition which is supposed to hold all along the proof. Finally, $E|f_K(\mathbf{W})|^4$ does not of course depend of the particular way $\mathbf{W}$ is generated, provided it is Haar distributed. We can therefore assume that $\mathbf{W}$ consists of the first K columns of the Haar distributed random matrix $\mathbf{V}(\mathbf{X})$ obtained from a complex Gaussian i.i.d. random matrix $\mathbf{X}$ through formula (7). As $\mathbf{V}(\mathbf{X})$ is obtained by a Gram-Schmidt orthogonalization of $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_K, \ldots, \mathbf{x}_N]$ and vector $\mathbf{w}$ is assumed to be the column K of $\mathbf{V}(\mathbf{X})$, $\mathbf{w}$ can be written as

$$\mathbf{w} = \frac{\mathbf{\Pi}\mathbf{x}_K}{\|\mathbf{\Pi}\mathbf{x}_K\|} \tag{55}$$

It is important to note that $\mathbf{U}$ depends only on vectors $\mathbf{x}_1, \ldots, \mathbf{x}_{K-1}$. Therefore, $\mathbf{x}_K$ and $\mathbf{\Pi}$ are independent. The expression (13) of the SINR β_N becomes

$$\beta_N = p_1 \frac{\mathbf{x}_K^H \mathbf{A} \mathbf{\Pi} \mathbf{x}_K}{\|\mathbf{\Pi}\mathbf{x}_K\|^2} \ . \tag{56}$$

Put $e_N = e_{1,N} + e_{2,N}$ where $e_{1,N}$ and $e_{2,N}$ are defined by

$$e_{1,N} = \frac{\mathbf{x}_K^H \mathbf{A} \mathbf{\Pi} \mathbf{x}_K}{\|\mathbf{\Pi}\mathbf{x}_K\|^2} - \frac{\mathbf{x}_K^H \mathbf{A} \mathbf{\Pi} \mathbf{x}_K}{N-K}$$

and

$$e_{2,N} = \frac{\mathbf{x}_K^H \mathbf{A} \mathbf{\Pi} \mathbf{x}_K}{N-K} - \frac{\mathrm{tr}(\mathbf{A})}{N-K} = \frac{N}{N-K} \frac{1}{N} \left(\mathbf{x}_K^H \mathbf{A} \mathbf{\Pi} \mathbf{x}_K - \mathrm{tr}(\mathbf{A}) \right) \ .$$

Then, $E|e_N|^4 \leq 8(E|e_{1,N}|^4 + E|e_{2,N}|^4)$. We first use lemma 1 to show that $E|e_{2,N}|^4 = O(N^{-2})$. As $\mathbf{U}$ and $\mathbf{x}_K$ are independent, $\mathbf{A}$ and $\mathbf{x}_K$ are also independent. Moreover, the spectral norm of $(\mathbf{HUQU}^H\mathbf{H}^H + \sigma^2 I)^{-1}$ is uniformly bounded by $1/\sigma^2$. As $\|\mathbf{H}\|$ is bounded, the inequality $\|\mathbf{CD}\| \leq \|\mathbf{C}\| \, \|\mathbf{D}\|$ implies that $\mathbf{A}$ is bounded. Furthermore, $\mathbf{x}_K$ meets the conditions of lemma 1 since it has Gaussian independent elements. This lemma and the fact that $N/(N-K) \to 1/(1-\alpha)$ imply that $E|e_{2,N}|^4 = O(N^{-2})$ converges to 0 almost surely (set $\mathbf{z} = \mathbf{x}$ and $\mathbf{B} = \mathbf{A\Pi}$, and notice that $\mathrm{tr}(\mathbf{B}) = \mathrm{tr}(\mathbf{A})$). By the Borel-Cantelli lemma, $E|e_{2,N}|^4 = O(N^{-2})$ implies that $e_{2,N}$ converges to 0 almost surely.

We now study the behavior of $E|e_{1,N}|^4$. $e_{1,N}$ can be written as

$$e_{1,N} = \left(\frac{\mathbf{x}_K^H \mathbf{A\Pi x}_K}{N-K} \right) \left(\frac{N-K}{\|\mathbf{\Pi x}_K\|^2} - 1 \right) .$$

As $e_{2,N}$ converges to 0 almost everywhere,

$$\frac{\mathbf{x}_K^H \mathbf{A\Pi x}_K}{(N-K)} < 2 \frac{\mathrm{tr}(\mathbf{A})}{(N-K)} \leq 2\|\mathbf{A}\| \quad \text{a.s.}$$

for N large enough. The last inequality comes from the facts that for a given matrix $\mathbf{X}$, $\mathrm{tr}(\mathbf{X}) \leq \|\mathbf{X}\| \, \mathrm{rank}(\mathbf{X})$, and that the rank of $\mathbf{A}$ does not exceed $N-K$ as can be seen from the expression of this matrix. As $\sup_{N \in \mathbb{N}} \|\mathbf{A}\| < \infty$, $\left(\mathbf{x}_K^H \mathbf{A x}_K\right) / (N-K)$ is bounded almost everywhere.

We now show that

$$E \left(\frac{N-K}{\|\mathbf{\Pi x}_K\|^2} - 1 \right)^4 = O(N^{-2}).$$

Let $\mathbf{U}'$ be a $N \times (N-K+1)$ isometric matrix such that $\mathbf{\Pi} = \mathbf{U}'\mathbf{U}'^H$. Then, $\|\mathbf{\Pi x}_K\|^2 = \|\mathbf{U}'^H \mathbf{x}_K\|^2$. As $\mathbf{x}_K$ and $\mathbf{U}'$ are independent, $\mathbf{U}'^H \mathbf{x}_K$ is a $(N-K+1)$-dimensional complex Gaussian random vector with covariance matrix $\mathbf{I}_{N-K+1}$. Therefore, $\|\mathbf{U}'^H \mathbf{x}_K\|^2$ is χ^2 distributed with $2(N-K+1)$ degrees of freedom. Its probability density is the function $\frac{t^{(N-K)}}{(N-K)!} e^{-t}$, and a straightforward direct computation shows that

$$E \left(\frac{N-K}{\|\mathbf{\Pi x}_K\|^2} - 1 \right)^4 = O((N-K)^{-2})$$

which coincides with $O(N^{-2})$ if $K/N \to \alpha$. As $\left(\mathbf{x}_K^H \mathbf{A\Pi x}_K\right) / (N-K)$ is bounded almost everywhere, we get that $E|e_{1,N}|^4 = O(N^{-2})$. Putting all the pieces together, this implies that $E|e_N|^4 = O(N^{-2})$, and that e_N converges almost surely to 0. ■

We now establish that $\frac{\mathrm{tr}(\mathbf{A})}{N-K}$ converges almost surely toward a certain deterministic limit. For this, we are going to show that this is the case for $\frac{\mathrm{tr}(\mathbf{A})}{N}$ by using the concept of almost sure asymptotic freeness of random matrices. We put $\mathbf{R} = \mathbf{UQU}^H$, and first justify that $\mathbf{R}$ and $\mathbf{H}^H\mathbf{H}$ are asymptotically free almost everywhere. For this, we have first to justify that the eigenvalue distributions of both matrices converge almost surely toward a deterministic probability distribution. This is of course true for $\mathbf{R}$, the limit distribution of which is the measure $\nu = (1-\alpha)\delta(\lambda) + \alpha\nu_p$ (we recall that ν_p is the users power limit distribution). The eigenvalues of $\mathbf{H}^H\mathbf{H}$ coincide with the $(|h(e^{2i\pi k/N})|^2)_{k=0,\ldots,N-1}$. Therefore, for each bounded continuous function ϕ, $\frac{\mathrm{tr}(\phi(\mathbf{H}^H\mathbf{H}))}{N}$ can be written as

$$\frac{\mathrm{tr}(\phi(\mathbf{H}^H\mathbf{H}))}{N} = \frac{\sum_{k=0}^{N-1}\phi(|h(e^{2i\pi k/N})|^2)}{N}$$

When $N \to \infty$, this of course converges toward $\int_0^1 \phi(|h(e^{2i\pi f})|^2 df$, which can be written as $\int \phi(\lambda)d\mu(\lambda)$ for a certain probability measure μ supported by the interval $[0, \max_f |h(e^{2i\pi f})|^2]$. We note in particular that the ψ-transform ψ_μ of μ is given by

$$\psi_\mu(z) = \int \frac{z\lambda}{1-z\lambda}\, d\mu(\lambda) = \int_0^1 \frac{z|h(e^{2i\pi f})|^2}{1-z|h(e^{2i\pi f})|^2}\, df \tag{57}$$

We finally remark that $\mathbf{R}$ is unitarily invariant in the sense that for each constant $N \times N$ unitary matrix $\mathbf{\Theta}$, the probability distribution of $\mathbf{R}$ coincides with the probability distribution of $\mathbf{\Theta}^H\mathbf{R}\mathbf{\Theta}$. Therefore (see [15, proposition 4.3.9]), $\mathbf{R}$ and $\mathbf{H}^H\mathbf{H}$ are asymptotically free almost everywhere. This in particular implies that the normalized trace of every non commutative monomial in $\mathbf{R}$ and $\mathbf{H}^H\mathbf{H}$ converges almost surely toward a deterministic limit. The last step consists in showing that $\frac{\mathrm{tr}(\mathbf{A})}{N}$ can be approximated by a linear combination of such monomials. For this, we first remark that the eigenvalues of $\mathbf{R}$ belong to $[b, B]$ for each N (see assumption **A2**). Let $f(\lambda)$ be a continuous function satisfying $f(\lambda) = 0$ on $[b, B]$, and $f(\lambda) = 1$ in a neighborhood of 0. Then, it is clear that $I - \mathbf{U}\mathbf{U}^H$ coincides with $f(\mathbf{R})$. By the Stone-Weierstrass theorem, for each $\epsilon > 0$, it exists a polynomial P_1 for which $|f(\lambda) - P_1(\lambda)| < \epsilon$ on $[0, B]$. Therefore, $\|f(\mathbf{R}) - P_1(\mathbf{R})\| < \epsilon$ for each N. The spectrum of $\mathbf{HUQU}^H\mathbf{H}^H + \sigma^2\mathbf{I}$ is included into $[\sigma^2, \sigma^2 + B\max_f |h(e^{2i\pi f})|^2]$. As the function $\lambda \to \frac{1}{\lambda}$ is continuous on $[\sigma^2, \sigma^2 + B\max_f |h(e^{2i\pi f})|^2]$, it also exists a polynomial P_2 such that

$$\|(\mathbf{HUQU}^H\mathbf{H}^H + \sigma^2\mathbf{I})^{-1} - P_2(\mathbf{HUQU}^H\mathbf{H}^H + \sigma^2\mathbf{I})\| < \epsilon$$

for each N. This shows that $\frac{\mathrm{tr}(\mathbf{A})}{N}$ can be approximated with an arbitrary good accuracy by a (finite) linear combination of terms of the form

$$\frac{1}{N}\mathrm{tr}\left(\mathbf{R}^k\mathbf{H}^H(\mathbf{H}\mathbf{R}\mathbf{H}^H+\sigma^2\mathbf{I})^l\mathbf{H}\right)$$

By expanding $(\mathbf{H}\mathbf{R}\mathbf{H}^H+\sigma^2\mathbf{I})^l$, it is easy to check that the above term is the normalized trace of the value taken by a non-commutative polynomial in the indeterminates X_1 and X_2 where X_1 and X_2 are replaced by $\mathbf{H}^H\mathbf{H}$ and $\mathbf{R}$ respectively. As $\mathbf{R}$ and $\mathbf{H}^H\mathbf{H}$ are asymptotically free almost everywhere, the above term converges almost surely toward a deterministic constant. This in turn shows that $\frac{\mathrm{tr}(\mathbf{A})}{N}$, and thus $\frac{\beta_N}{p_1}$ as well as β_N converge almost surely to a deterministic term.

B. Proof of the claims of the second step.

The reader may check that the proof of the first step of theorem 1 implies that for each sequence $(k(N))_{N\geq 1}$ of integers satisfying $1\leq k(N)\leq K$, then $\overline{\beta}_{N,k(N)}$ and $\frac{\mathrm{tr}(\mathbf{A}_{k(N)})}{N-K}$ have the same asymptotic behavior. Moreover, $\frac{\mathrm{tr}(\mathbf{A}_{k(N)})}{N-K}$ converges toward a deterministic constant. In order to justify that this constant does not depend on the sequence $(k(N))_{N\geq 1}$, we remark that the difference between matrix $\mathbf{A}_{k(N)}$ and

$$(\mathbf{I}-\mathbf{W}\mathbf{W}^H)\mathbf{H}^H\left(\mathbf{H}\mathbf{W}\mathbf{P}\mathbf{W}^H\mathbf{H}^H+\sigma^2\mathbf{I}_N\right)^{-1}\mathbf{H}$$

is a uniformly bounded rank 3 matrix (just use the matrix inversion lemma). Hence, $\frac{\mathrm{tr}(\mathbf{A}_{k(N)})}{N-K}$ and

$$\frac{\mathrm{tr}\left((\mathbf{I}-\mathbf{W}\mathbf{W}^H)\mathbf{H}^H\left(\mathbf{H}\mathbf{W}\mathbf{P}\mathbf{W}^H\mathbf{H}^H+\sigma^2\mathbf{I}_N\right)^{-1}\mathbf{H}\right)}{N-K}$$

have the same limit $\overline{\beta}$, which is of course independent of sequence $(k(N))_{N\geq 1}$. We now establish relation (22). For each K, we denote by $g_K(x)$ the random process defined for $x\in[0,1[$ by

$$g_K(x)=\eta_{N,\lfloor xK+1\rfloor}$$

where $\lfloor xK+1\rfloor$ represents the greatest integer less than or equal to $xK+1$. It is obvious that

$$\frac{\sum_{k=1}^K\eta_{N,k}}{K}=\int_0^1 g_K(x)\,dx$$

We denote $\overline{g}_K(x)$ the deterministic function defined by

$$\overline{g}_K(x) = \frac{p_{\lfloor xK+1 \rfloor}\overline{\beta}}{1 + p_{\lfloor xK+1 \rfloor}\overline{\beta}}$$

It is clear that

$$\int_0^1 \overline{g}_K(x)\, dx = \frac{1}{K}\left(\sum_{k=1}^{K} \frac{p_k\overline{\beta}}{1+p_k\overline{\beta}}\right)$$

By the very definition of the concept of limit distribution, it therefore turns out that

$$\lim_{K\to\infty} \int_0^1 \overline{g}_K(x)\, dx = \int \frac{\lambda\overline{\beta}}{1+\lambda\overline{\beta}}\, d\nu_p(\lambda)$$

In order to show (22), it is thus sufficient to establish that

$$\lim_{K\to\infty} \int_0^1 g_K(x)\, dx = \lim_{K\to\infty} \int_0^1 \overline{g}_K(x)\, dx \tag{58}$$

almost surely. As $\overline{\beta}_{N,\lfloor xK+1 \rfloor}$ converges almost surely to $\overline{\beta}$ when $N \to \infty$, it is clear that for each $x \in [0,1[$, $\lim_{K\to\infty}(g_K(x) - \overline{g}_K(x)) = 0$ almost surely. In other words, the probability of the event $G(x) = \{(g_K(x) - \overline{g}_K(x))\text{does not converge to } 0\}$ is equal to 0, i.e.,

$$E\left[\mathbf{1}_{G(x)}\right] = 0$$

where $\mathbf{1}_{G(x)}$ represents the set indicator function of the event $G(x)$. Integrating on $[0,1]$ this identity with respect to x, and using the Fubini theorem yields to

$$E\left[\int_0^1 \mathbf{1}_{G(x)}\, dx\right] = 0$$

The random variable f defined by

$$f = \int_0^1 \mathbf{1}_{G(x)}\, dx$$

is thus equal to 0 almost surely. We now use the following identity :

$$\begin{aligned} \int_0^1 |g_K(x) - \overline{g}_K(x)|\, dx &= \int_0^1 \mathbf{1}_{\{g_K(x)-\overline{g}_K(x)\to 0\}}\, |g_K(x) - \overline{g}_K(x)|\, dx \\ &+ \int_0^1 \mathbf{1}_{G(x)}\, |g_K(x) - \overline{g}_K(x)|\, dx \end{aligned} \tag{59}$$

As $|g_K(x) - \overline{g}_K(x)| \leq 2$, the Lebesgue dominated convergence theorem implies that the first term of the right hand side of (59) converges to 0. As for the second term, it is less that $2f$, and is thus equal to 0. This in turn establishes (22).

III. Justification of eq. (29).

In this section, we justify equation (29) in the case where code vectors $(\mathbf{w}_k)_{k=1,\ldots,K}$ are Gaussian. Notice that it should be possible to release the Gaussian assumption by using the results of [18, Chap. 16]. For each k, matrix $\mathbf{H}_k$ is circulant, and can thus be written as $\mathbf{H}_k = \mathbf{F}\mathbf{D}_k\mathbf{F}^H$. We recall that $\mathbf{F}$ is the $N \times N$ Fourier matrix which (p,q) entry is $\mathbf{F}_{p,q} = \frac{1}{\sqrt{N}} e^{2i\pi pq/N}$ for $(p,q) \in \{0,\ldots,N-1\}$, and $\mathbf{D}_k$ is the diagonal matrix with entries $(h_k(e^{2i\pi l/N})_{l=0,\ldots,N-1}$. Thanks to the Gaussian character of the code vectors, it is possible to replace matrix $\mathbf{H}_k$ by the diagonal matrix $\mathbf{D}_k$ in order to evaluate the asymptotic SINR of user 1. The non asymptotic SINR $\beta_{1,N}$ associated to this user in the uplink is now given by

$$\beta_{1,N} = \mathbf{w}_1^H \mathbf{D}_1^H (\sum_{k=2}^{K} \mathbf{D}_k \mathbf{w}_k \mathbf{w}_k^H \mathbf{D}_k^H + \sigma^2 \mathbf{I})^{-1} \mathbf{D}_1 \mathbf{w}_1$$

For each k, the entries of vector $\mathbf{D}_k\mathbf{w}_k$ are independent, the entry l having a variance equal to $\frac{1}{N}|h_k(e^{2i\pi l/N})|^2$ for $l = 0,\ldots,N-1$. Therefore $\beta_{1,N}$ behaves asymptotically as

$$\frac{1}{N}\sum_{l=0}^{N-1} |h_1(e^{2i\pi l/N})|^2 \left((\sum_{k=2}^{K} \mathbf{D}_k \mathbf{w}_k \mathbf{w}_k^H \mathbf{D}_k^H + \sigma^2 \mathbf{I})^{-1} \right)_{l,l} .$$

It is possible to analyze the asymptotic behavior of the diagonal terms of $(\sum_{k=2}^{K} \mathbf{D}_k \mathbf{w}_k \mathbf{w}_k^H \mathbf{D}_k^H + \sigma^2 \mathbf{I})^{-1}$ by using the results of [18, Chap.7], and thus to evaluate the limit of $\beta_{1,N}$. Let $(\phi_l(z))_{l=0,\ldots,N-1}$ and $(\psi_j(z))_{j=1,\ldots,K}$ be the functions in the class of Stieltjes transforms uniquely defined by the system of equations

$$\begin{aligned} \phi_l(z) &= -z + \frac{1}{N}\sum_{k=1}^{K} |h_k(e^{2i\pi l/N})|^2 \frac{1}{\psi_k(z)} \\ \psi_j(z) &= 1 + \frac{1}{N}\sum_{n=0}^{N-1} |h_j(e^{2i\pi n/N})|^2 \frac{1}{\phi_n(z)} . \end{aligned} \tag{60}$$

Then, the lth diagonal term of $(\sum_{k=2}^{K} \mathbf{D}_k \mathbf{w}_k \mathbf{w}_k^H \mathbf{D}_k^H + \sigma^2 \mathbf{I})^{-1}$ has the same asympptotic behavior than $\frac{1}{\phi_l(-\sigma^2)}$. Therefore, we have

$$\beta_{1,N} - \frac{1}{N}\sum_{l=0}^{N-1} |h_1(e^{2i\pi l/N})|^2 \frac{1}{\phi_l(-\sigma^2)} \to 0$$

in probability. Similarly, the non asymptotic SINR $\beta_{k,N}$ of user k is such that

$$\beta_{k,N} - \frac{1}{N}\sum_{l=0}^{N-1} |h_k(e^{2i\pi l/N})|^2 \frac{1}{\phi_l(-\sigma^2)} \to 0$$

in probability. Therefore, equation (60) leads to $\psi_j(-\sigma^2) \simeq 1 + \beta_{j,N}$, so that

$$\frac{1}{\phi_l(-\sigma^2)} \simeq \frac{1}{\sigma^2 + \frac{1}{N}\sum_{k=1}^{K} \frac{|h_k(e^{2i\pi l/N})|^2}{1+\beta_{k,N}}}$$

and therefore

$$\beta_{1,N} \simeq \frac{1}{N}\sum_{l=0}^{N-1} \frac{|h_1(e^{2i\pi l/N})|^2}{\sigma^2 + \frac{1}{N}\sum_{k=1}^{K} \frac{|h_k(e^{2i\pi l/N})|^2}{1+\beta_{k,N}}} \ . \tag{61}$$

Therefore, the asymptotic SINRs $(\beta_{k,\text{uplink}})_{k=1,\dots,K}$ are such that

$$\beta_{1,\text{uplink}} \simeq \frac{1}{N}\sum_{l=0}^{N-1} \frac{|h_1(e^{2i\pi l/N})|^2}{\sigma^2 + \frac{1}{N}\sum_{k=2}^{K} \frac{|h_k(e^{2i\pi l/N})|^2}{1+\beta_{k,\text{uplink}}}}$$

IV. Proof of Proposition 3

In order to prove Proposition 3, we first justify that $\overline{\gamma}_N$ converges almost surely toward a certain deterministic constant $\overline{\gamma}$. We define matrix $\mathbf{B}$ by

$$\mathbf{B} = (\mathbf{I} - \mathbf{U}\mathbf{U}^H)\mathbf{H}^H \left(\mathbf{H}\mathbf{H}^H + \frac{\sigma^2}{\frac{K}{N}\overline{p}}\mathbf{I}\right)^{-1} \left(\mathbf{H}\mathbf{U}\mathbf{Q}\mathbf{U}^H\mathbf{H}^H + \sigma^2\mathbf{I}\right) \left(\mathbf{H}\mathbf{H}^H + \frac{\sigma^2}{\frac{K}{N}\overline{p}}\mathbf{I}\right)^{-1} \mathbf{H}$$

Using again Proposition 4, it is easily seen that

$$\lim_{N\to+\infty} \overline{\gamma}_N - \frac{\text{tr}(\mathbf{B})}{N-K} = 0$$

almost surely. The fact that $\overline{\gamma}_N$ converges almost surely to a constant $\overline{\gamma}$ is shown as in the proof of the first step of Theorem 1 (use that matrices $\mathbf{H}^H\mathbf{H}$ and $\mathbf{R} = \mathbf{U}\mathbf{Q}\mathbf{U}^H$ are asymptotically free almost everywhere, and that $\overline{\gamma}_N$ can be approximated by the trace of non commutative polynomials of $\mathbf{H}^H\mathbf{H}$ and $\mathbf{R}$).

We now show that the above mentioned limit $\overline{\gamma}$ satisfies (42). For this, we remark that $\overline{\gamma}_N$ and $\overline{\gamma}$ are bounded almost surely. Therefore, the Lebesgue dominated convergence theorem implies that $E|\overline{\gamma}_N - \overline{\gamma}|^2 \to 0$ when $N \to +\infty$. This implies that

$$\lim_{N\to+\infty} E\left|E(\overline{\gamma}_N|\mathbf{w}) - \overline{\gamma}\right|^2 = 0.$$

Therefore, we get that

$$\overline{\gamma} = \lim_{N \to +\infty} E(\overline{\gamma}_N | \mathbf{w})$$

where the limit is understood in the least mean square sense. In order to evaluate $E(\overline{\gamma}_N|\mathbf{w})$, we prove the following lemma.

Lemma 2: The following property holds :

$$E(\mathbf{UQU}^H|\mathbf{w}) = \frac{\sum_{k=2}^K p_k}{N-1} \left(\mathbf{I} - \mathbf{ww}^H\right) \tag{62}$$

Proof. As relation (62) only depends on the statistical properties of the Haar distribution, we can choose the way $\mathbf{W}$ is generated, provided it is obtained by extracting K columns from a Haar distributed unitary matrix. We first generate a Haar distributed $N \times N$ random unitary matrix $\tilde{\mathbf{\Theta}}_{N,N} = (\mathbf{w}(n), \tilde{\mathbf{\Theta}}_{N,N-1})$, i.e. $\mathbf{w}$ is the first column of $\tilde{\mathbf{\Theta}}_{N,N}$. Let $\overline{\mathbf{\Theta}}_{N-1,N-1}$ be a $(N-1) \times (N-1)$ Haar distributed random unitary matrix **independent from** $\tilde{\mathbf{\Theta}}_{N,N}$. Then, it is easily seen that the $N \times N$ matrix $\mathbf{\Theta}_{N,N}$ defined by

$$\mathbf{\Theta}_{N,N} = \left[\mathbf{w}, \tilde{\mathbf{\Theta}}_{N,N-1}\overline{\mathbf{\Theta}}_{N-1,N-1}\right]$$

is Haar distributed. We then define $\mathbf{U}$ as the matrix obtained by extracting columns 2 to K of $\mathbf{\Theta}_{N,N}$. If $\overline{\mathbf{U}}$ represents the first $(K-1)$ columns of $\overline{\mathbf{\Theta}}_{N-1,N-1}$, then,

$$\mathbf{U} = \tilde{\mathbf{\Theta}}_{N,N-1}\overline{\mathbf{U}}$$

It is clear that the conditional expectation of $\mathbf{UQU}^H$ given $\tilde{\mathbf{\Theta}}_{N,N}$ is equal to

$$\tilde{\mathbf{\Theta}}_{N,N-1} E\left(\overline{\mathbf{U}}\mathbf{Q}\overline{\mathbf{U}}^H\right) \tilde{\mathbf{\Theta}}_{N,N-1}^H$$

But, using the properties of the Haar distribution, it is easily seen that two different entries of $\overline{\mathbf{U}}$ are decorrelated, while their second order moments all coincide with $\frac{1}{N-1}$. From this, we get immediately that

$$E(\overline{\mathbf{U}}\mathbf{Q}\overline{\mathbf{U}}^H) = \frac{\sum_{k=2}^K p_k}{N-1}\mathbf{I}_{N-1}$$

Therefore, the conditional expectation of $\mathbf{UQU}^H$ given $\tilde{\mathbf{\Theta}}_{N,N}$ is equal to

$$\frac{\sum_{k=2}^K p_k}{N-1}\tilde{\mathbf{\Theta}}_{N,N-1}\tilde{\mathbf{\Theta}}_{N,N-1}^H \ ,$$

which coincides with $\frac{\sum_{k=2}^K p_k}{N-1}(\mathbf{I}_N - \mathbf{ww}^H)$. This only depends on $\mathbf{w}$, so that this identity implies (62).

This shows (42) and completes the proof of Proposition (3).

References

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

[2] D.R. Brown, M. Motani, V.V. Veeravalli, H.V. Poor, and C.R. Johnson, "On the Performance of Linear Parallel Interference Cancellation," *IEEE Transactions on Information Theory*, vol. 47, no. 5, pp. 1957–1970, July 2001.

[3] H. Poor and S. Verdú, "Probability of Error in MMSE Multiuser Detection," *IEEE Transactions on Information Theory*, vol. 43, no. 3, pp. 858–871, May 1997.

[4] J. Zhang, E.K.P. Chong, and D.N.C. Tse, "Output MAI Distributions of Linear MMSE Multiuser Receivers in DS-CDMA Systems," *IEEE Transactions on Information Theory*, vol. 47, no. 3, pp. 1128–1144, Mar. 2001.

[5] D.N.C Tse and S. Hanly, "Linear Multi-User Receiver: Effective Interference, Effective Bandwidth and User Capacity," *IEEE Transactions on Information Theory*, vol. 45, no. 2, pp. 641–657, Mar. 1999.

[6] J. Evans and D.N.C Tse, "Large System Performance of Linear Multiuser Receivers in Multipath Fading Channels," *IEEE Transactions on Information Theory*, vol. 46, no. 6, pp. 2059–2078, Sept. 2000.

[7] W.G. Phoel and M.L. Honig, "Performance of Coded DS-CDMA With Pilot-Assisted Channel Estimation and Linear Interference Suppression," *IEEE Transactions on Communications*, vol. 50, no. 5, pp. 822–832, May 2002.

[8] C.D. Frank and E. Visotsky, "Adaptive Interference Suppression for CDMA with Long Spreading Codes," in *Proceedings of the Allerton Conference on Communication, Control, and Computing*, Monticello, Illinois, USA, Sept. 1998, pp. 411–420.

[9] M. Lenardi, A. Medles, and D.T.M. Slock, "A SINR Maximizing RAKE Receiver for DS-CDMA Downlinks," in *Proceedings of the 34th Asilomar conference on Signals, Systems and Computers*, Pacific Grove, California, USA, Oct. 2000.

[10] K. Hooli, M. Juntti, M. Heikkilä, P. Komulainen, M. Latva-Aho, and J. Lilleberg, "Chip-Level Channel Equalization in WCDMA Downlink," *EURASIP Journal of Applied Signal Processing*, , no. 8, pp. 757–770, Aug. 2002.

[11] M. Debbah, W. Hachem, P. Loubaton, and M. de Courville, "MMSE Analysis of Certain Large Isometric Random Precoded Systems," *IEEE Transactions on Information Theory*, vol. 49, no. 5, pp. 1293–1311, May 2003.

[12] Z.D. Bai and J.W. Silverstein, "No Eigenvalues Outside the Support of the Limiting Spectral Distribution of Large Dimensional Sample Covariance Matrices," *Annals of Probability*, vol. 26, no. 1, pp. 316–345, 1998.

[13] J.W. Silverstein and Z.D. Bai, "On the Empirical Distribution of Eigenvalues of a Class of Large Dimensional Random Matrices," *J. Multivariate Anal.*, vol. 54, no. 2, pp. 175–192, 1995.

[14] Ph. Biane, "Free Probability for Probabilists," *http://www.dma.ens.fr/~biane/*, 2000, download by choosing "Des notes d'introduction aux probabilités libres".

[15] F. Hiai and D. Petz, *The Semicircle Law, Free Random Variables and Entropy*, vol. 77 of *Mathematical Surveys and Monographs*, AMS, 2000.

[16] ETSI, Ed., *3GPP, Technical Specifications 3G TS 25.211, Physical Channels and Mapping og Transport Channel onto Physical Channels (FDD)*, vol. 1, Release 99, 2000.

[17] D.N.C Tse and O. Zeitouni, "Linear multiusers receivers in random environments," *IEEE Transactions on Information Theory*, vol. 46, no. 1, pp. 641–657, Jan. 2000.

[18] V. L. Girko, *Theory of Stochastic Canonical Equations*, vol. 1, Kluwer, 2001.

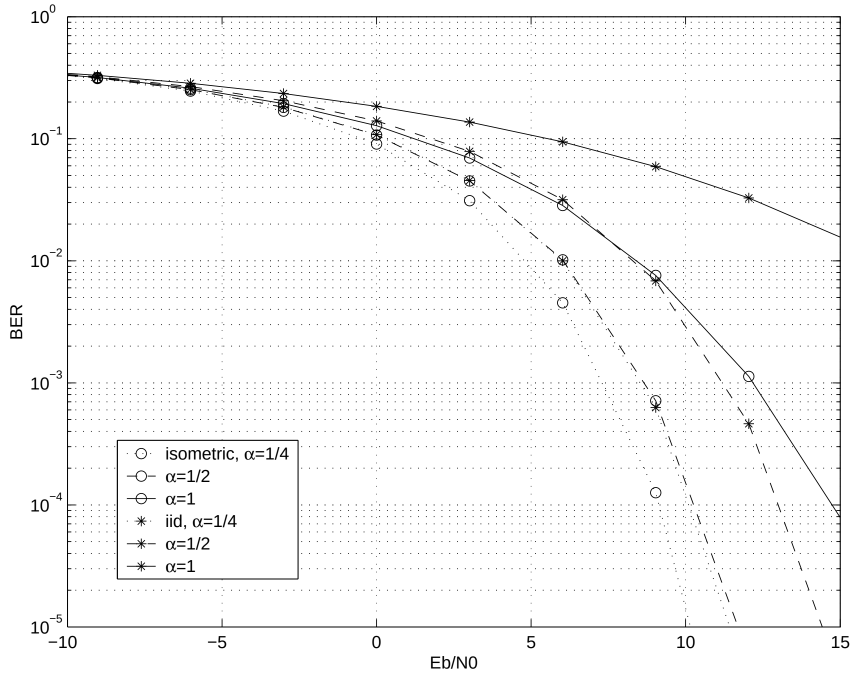

Fig. 1. Optimal Wiener Filter in isometric and i.i.d. cases

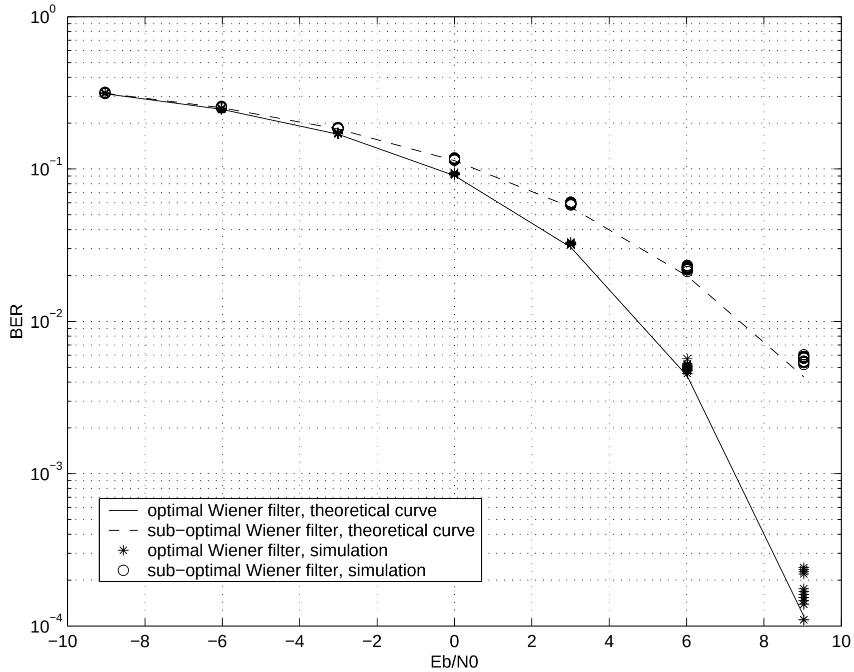

Fig. 2. Theoretical and empirical performance of the optimal and sub-optimal Wiener filters, $\alpha = \frac{1}{4}$

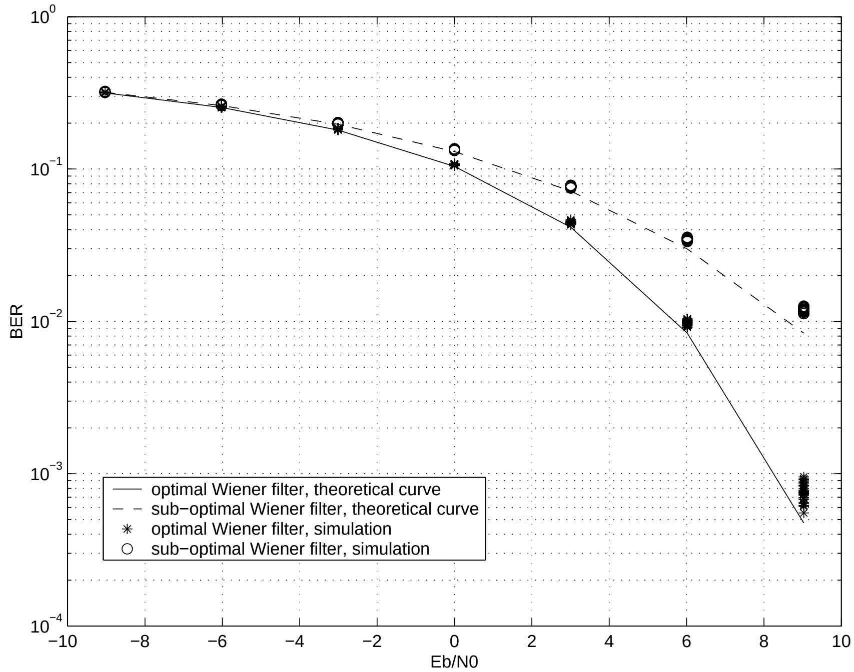

Fig. 3. Theoretical and empirical performance of the optimal and sub-optimal Wiener filters, $\alpha = \frac{1}{2}$

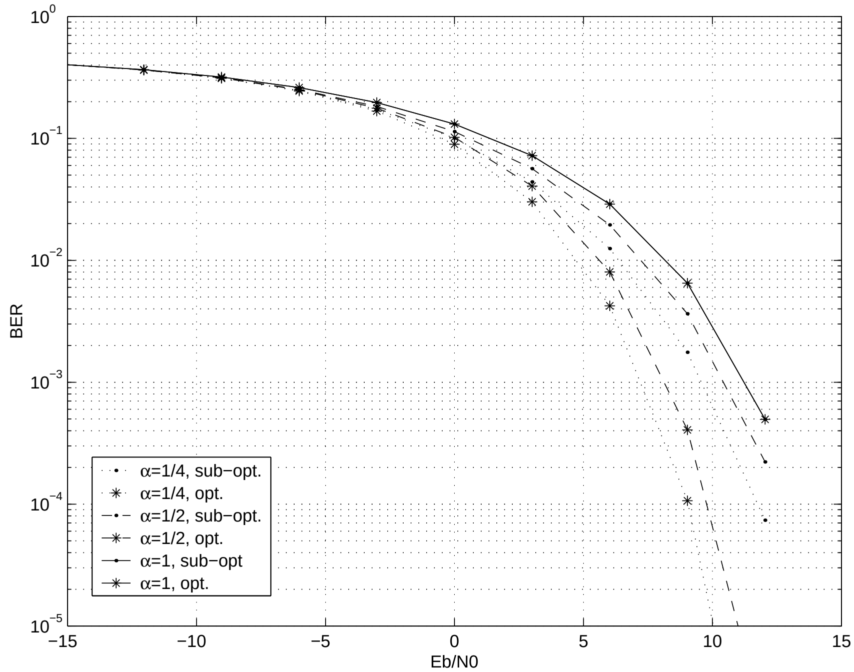

Fig. 4. Comparison of optimal and sub-optimal Wiener filters

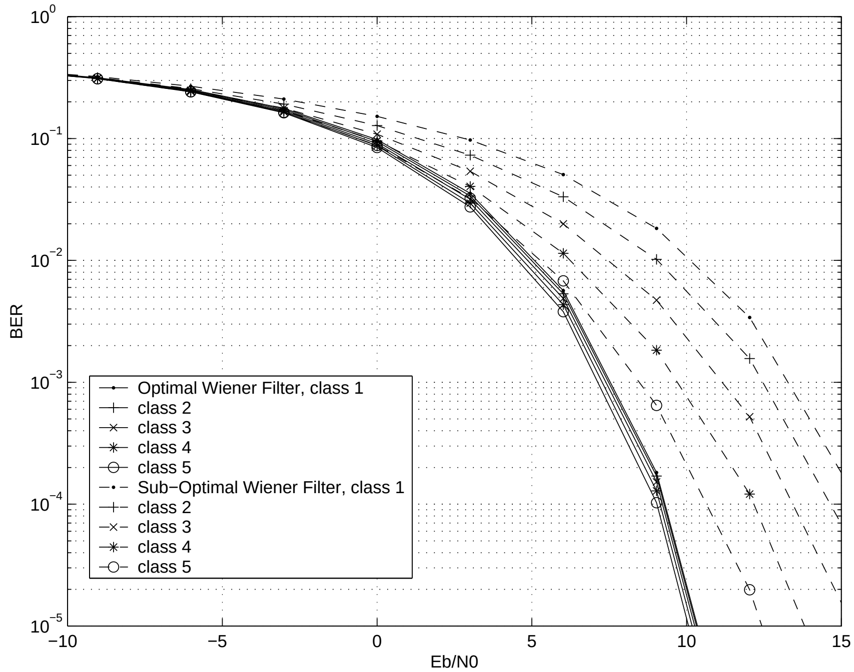

Fig. 5. Performance of the various power classes with optimal and sub-optimal Wiener filters