A Central Limit Theorem for the SINR at the LMMSE Estimator Output for Large Dimensional Signals

Abla Kammoun[(1)], Malika Kharouf[(2)], Walid Hachem[(3)] and Jamal Najim[(3)]

Abstract

This paper is devoted to the performance study of the Linear Minimum Mean Squared Error estimator for multidimensional signals in the large dimension regime. Such an estimator is frequently encountered in wireless communications and in array processing, and the Signal to Interference and Noise Ratio (SINR) at its output is a popular performance index. The SINR can be modeled as a random quadratic form which can be studied with the help of large random matrix theory, if one assumes that the dimension of the received and transmitted signals go to infinity at the same pace. This paper considers the asymptotic behavior of the SINR for a wide class of multidimensional signal models that includes general multi-antenna as well as spread spectrum transmission models.

The expression of the deterministic approximation of the SINR in the large dimension regime is recalled and the SINR fluctuations around this deterministic approximation are studied. These fluctuations are shown to converge in distribution to the Gaussian law in the large dimension regime, and their variance is shown to decrease as the inverse of the signal dimension.

Index Terms

Antenna Arrays, CDMA, Central Limit Theorem, LMMSE, Martingales, MC-CDMA, MIMO, Random Matrix Theory.

(1) Telecom ParisTech (ENST), Paris, France. abla.kammoun@enst.fr
(2) Telecom ParisTech (ENST) and Casablanca University, Morocco. malika.kharouf@enst.fr
(3) CNRS / ENST, Paris, France. walid.hachem, jamal.najim@enst.fr

I. INTRODUCTION

Large Random Matrix Theory (LRMT) is a powerful mathematical tool used to study the performance of multi-user and multi-access communication systems such as Multiple Input Multiple Output (MIMO) digital wireless systems, antenna arrays for source detection and localization, spread spectrum communication systems as Code Division Multiple Access (CDMA) and Multi-Carrier CDMA (MC-CDMA) systems. In most of these communication systems, the N dimensional received random vector $\mathbf{r} \in \mathbb{C}^N$ is described by the model

$$\mathbf{r} = \mathbf{\Sigma s} + \mathbf{n} \tag{1}$$

where $\mathbf{s} = [s_0, s_1, \ldots, s_K]^{\mathrm{T}}$ is the unknown random vector of transmitted symbols with size $K+1$ satisfying $\mathbb{E}\mathbf{ss}^* = \mathbf{I}_{K+1}$, the noise $\mathbf{n}$ is an independent Additive White Gaussian Noise (AWGN) with covariance matrix $\mathbb{E}\mathbf{nn}^* = \rho \mathbf{I}_N$ whose variance $\rho > 0$ is known, and matrix $\mathbf{\Sigma}$ represents the known "channel" in the wide sense whose structure depends on the particular system under study. One typical problem addressed by LRMT concerns the estimation performance by the receiver of a given transmitted symbol, say s_0.

In this paper we focus on one of the most popular estimators, namely the linear Wiener estimator, also called LMMSE for Linear Minimum Mean Squared Error estimator: the LMMSE estimate $\hat{s}_0 = \mathbf{g}^*\mathbf{r}$ of signal s_0 is the one for which the $N \times 1$ vector $\mathbf{g}$ minimizes $\mathbb{E}|\hat{s}_0 - s_0|^2$. If we partition the channel matrix as $\mathbf{\Sigma} = [\mathbf{y}\ \mathbf{Y}]$ where $\mathbf{y}$ is the first column of $\mathbf{\Sigma}$ and where matrix $\mathbf{Y}$ has dimensions $N \times K$, then it is well known that vector $\mathbf{g}$ is given by $\mathbf{g} = (\mathbf{\Sigma\Sigma}^* + \rho \mathbf{I}_N)^{-1}\mathbf{y}$. Usually, the performance of this estimator is evaluated in terms of the Signal to Interference plus Noise Ratio (SINR) at its output. Writing the received vector $\mathbf{r}$ as $\mathbf{r} = s_0\mathbf{y} + \mathbf{r}_{\text{in}}$ where $s_0\mathbf{y}$ is the relevant term and $\mathbf{r}_{\text{in}}$ represents the so-called interference plus noise term, the SINR is given by $\beta_K = |\mathbf{g}^*\mathbf{y}|^2/\mathbb{E}|\mathbf{g}^*\mathbf{r}_{\text{in}}|^2$. Plugging the expression of $\mathbf{g}$ given above into this expression, one can prove that the SINR β_K is given by the well-known expression:

$$\beta_K = \mathbf{y}^* \left(\mathbf{YY}^* + \rho \mathbf{I}_N\right)^{-1} \mathbf{y} \ . \tag{2}$$

In general, this expression does not provide a clear insight on the impact of the channel model parameters (such as the load factor KN^{-1}, the power distribution of the transmission data streams, or the correlation structure of the channel paths in the context of multi-antenna transmissions) on the performance of the LMMSE estimator.

An alternative approach, justified by the fluctuating nature of the channel paths in the context of MIMO communications and by the pseudo-random nature of the spreading sequences in spread spectrum applications consists to model matrix $\mathbf{\Sigma}$ as a random matrix (in this case, β_K becomes a random SINR). The simplest random matrix model for $\mathbf{\Sigma}$, corresponding to the most canonical MIMO or CDMA transmission channels, corresponds to independent and identically distributed (i.i.d.) entries with mean zero and variance N^{-1}. In that case, LRMT shows that when $K \to \infty$ and the load factor KN^{-1} converges to a limiting load factor $\alpha > 0$, the SINR β_K converges almost surely (a.s.) to an explicit deterministic quantity $\overline{\beta}(\alpha, \rho)$ which simply depends on the limiting load factor α and on the noise variance ρ. As a result, the impact of these two parameters on the LMMSE performance can be easily evaluated [1], [2].

The LMMSE SINR large dimensional behavior for more sophisticated random matrix models has also been thoroughly studied (cf. [1], [3]–[9]) and it has been proved that there exists a deterministic sequence $(\overline{\beta}_K)$, generally defined as the solution of an implicit equation, such that $\beta_K - \overline{\beta}_K \to 0$ almost surely as $K \to \infty$ and $\frac{K}{N}$ remains bounded away from zero and from infinity.

Beyond the convergence $\beta_K - \overline{\beta}_K \to 0$, a natural question arises concerning the accuracy of $\overline{\beta}_K$ for finite values of K. A first answer to this question consists in evaluating the Mean Squared Error (MSE) of the SINR $\mathbb{E}|\beta_K - \overline{\beta}_K|^2$ for large K. A further problem is the computation of outage probability, that is the probability for $\beta_K - \overline{\beta}_K$ to be below a certain level. Both problems can be addressed by establishing a Central Limit Theorem (CLT) for $\beta_K - \overline{\beta}_K$. In this paper, we establish such a CLT (Theorem 3 below) for a large class of random matrices $\mathbf{\Sigma}$. We prove that there exists a sequence $\Theta_K^2 = \mathcal{O}(1)$ such that $\frac{\sqrt{K}}{\Theta_K}(\beta_K - \overline{\beta}_K)$ converges in distribution to the standard normal law $\mathcal{N}(0,1)$ in the asymptotic regime. One can therefore infer that the MSE asymptotically behaves like $\frac{\Theta_K^2}{K}$ and that the outage probability can be simply approximated by a Gaussian tail function.

The class of random matrices $\mathbf{\Sigma}$ we consider in this paper is described by the following statistical model: Assume that

$$\mathbf{\Sigma} = \left(\Sigma_{nk}\right)_{n=1,k=0}^{N,K} = \left(\frac{\sigma_{nk}}{\sqrt{K}} W_{nk}\right)_{n=1,k=0}^{N,K} \tag{3}$$

where the complex random variables W_{nk} are i.i.d. with $\mathbb{E}W_{nk} = 0$, $\mathbb{E}W_{nk}^2 = 0$ and $\mathbb{E}|W_{nk}|^2 = 1$ and where $(\sigma_{nk}^2;\ 1 \le n \le N;\ 0 \le k \le K)$ is an array of real numbers. Due to the fact that $\mathbb{E}|\Sigma_{nk}|^2 = \frac{\sigma_{nk}^2}{K}$, the array (σ_{nk}^2) is referred to as a variance profile. An important particular case is when σ_{nk}^2 is *separable*, that is, writes:

$$\sigma_{nk}^2 = d_n \tilde{d}_k \ , \tag{4}$$

where $(d_1, \ldots, d_N)$ and $(\tilde{d}_0, \ldots, \tilde{d}_K)$ are two vectors of real positive numbers.

Applicative contexts.

Among the applicative contexts where the channel is described appropriately by model (3) or by its particular case (4), let us mention:

- Multiple antenna transmissions with $K+1$ distant sources sending their signals toward an array of N antennas. The corresponding transmission model is $\mathbf{r} = \mathbf{\Xi s} + \mathbf{n}$ where $\mathbf{\Xi} = \frac{1}{\sqrt{K}}\mathbf{H}\mathbf{P}^{1/2}$, matrix $\mathbf{H}$ is a $N \times (K+1)$ random matrix with complex Gaussian elements representing the radio channel, $\mathbf{P} = \mathrm{diag}(p_0, \ldots, p_K)$ is the (deterministic) matrix of the powers given to the different sources, and $\mathbf{n}$ is the usual AWGN satisfying $\mathbb{E}\mathbf{n}\mathbf{n}^* = \rho \mathbf{I}_N$. Write $\mathbf{H} = [\mathbf{h}_0 \ \cdots \ \mathbf{h}_K]$, and assume that the columns $\mathbf{h}_k$ are independent, which is realistic when the sources are distant one from another. Let $\mathbf{C}_k$ be the covariance matrix $\mathbf{C}_k = \mathbb{E}\mathbf{h}_k\mathbf{h}_k^*$ and let $\mathbf{C}_k = \mathbf{U}_k\mathbf{\Lambda}_k\mathbf{U}_k$ be a spectral decomposition of $\mathbf{C}_k$ where $\mathbf{\Lambda}_k = \mathrm{diag}(\lambda_{nk};\ 1 \le n \le N)$ is the matrix of eigenvalues. Assume now that the eigenvector matrices $\mathbf{U}_0, \ldots, \mathbf{U}_K$ are all equal (to some matrix $\mathbf{U}$, for instance), a case considered in e.g. [10] (note that sometimes they are all identified with the Fourier $N \times N$ matrix [11]). Let $\mathbf{\Sigma} = \mathbf{U}^*\mathbf{\Xi}$. Then matrix $\mathbf{\Sigma}$ is described by the statistical model (3) where the W_{nk} are standard Gaussian i.i.d., and $\sigma_{nk}^2 = \lambda_{nk} p_k$. If we partition $\mathbf{\Xi}$ as $\mathbf{\Xi} = [\mathbf{x}\ \mathbf{X}]$ similarly to the partition $\mathbf{\Sigma} = [\mathbf{y}\ \mathbf{Y}]$ above, then the SINR β at the output of the LMMSE estimator for the first element of vector $\mathbf{s}$ in the transmission model $\mathbf{r} = \mathbf{\Xi s} + \mathbf{n}$ is

$$\beta = \mathbf{x}^* \left(\mathbf{X}\mathbf{X}^* + \rho\mathbf{I}_N\right)^{-1} \mathbf{x} = \mathbf{y}^* \left(\mathbf{Y}\mathbf{Y}^* + \rho\mathbf{I}_N\right)^{-1} \mathbf{y}$$

  due to the fact that $\mathbf{U}$ is a unitary matrix. Therefore, the problem of LMMSE SINR convergence for this MIMO model is a particular case of the general problem of convergence of the right-hand member of (2) for model (3).

  It is also worth to say a few words about the particular case (4) in this context. If we assume that $\mathbf{\Lambda}_0 = \cdots = \mathbf{\Lambda}_K$ and these matrices are equal to $\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_N)$, then the model for $\mathbf{H}$ is the well-known Kronecker model with correlations at reception [12]. In this case,

$$\mathbf{\Sigma} = \mathbf{U}^*\mathbf{\Xi} = \frac{1}{\sqrt{K}}\mathbf{U}^*\mathbf{H}\mathbf{P}^{1/2} = \frac{1}{\sqrt{K}}\mathbf{\Lambda}^{1/2}\mathbf{W}\mathbf{P}^{1/2} \tag{5}$$

  where $\mathbf{W}$ is a random matrix with iid standard Gaussian elements. This model coincides with the separable variance profile model (4) with $d_n = \lambda_n$ and $\tilde{d}_k = p_k$.

- CDMA transmissions on flat fading channels. Here N is the spreading factor, $K+1$ is the number of users, and

$$\mathbf{\Sigma} = \mathbf{V}\mathbf{P}^{1/2} \tag{6}$$

where $\mathbf{V}$ is the $N \times (K+1)$ signature matrix assumed here to have random i.i.d. elements with mean zero and variance N^{-1}, and where $\mathbf{P} = \mathrm{diag}(p_0, \ldots, p_K)$ is the users powers matrix. In this case, the variance profile is separable with $d_n = 1$ and $\tilde{d}_k = \frac{K}{N} p_k$. Note that elements of $\mathbf{V}$ are not Gaussian in general.

- Cellular MC-CDMA transmissions on frequency selective channels. In the uplink direction, the matrix $\mathbf{\Sigma}$ is written as:

$$\mathbf{\Sigma} = [\mathbf{H}_0 \mathbf{v}_0 \ \cdots \ \mathbf{H}_{K+1} \mathbf{v}_{K+1}] \ , \tag{7}$$

where $\mathbf{H}_k = \mathrm{diag}(h_k(\exp(2\imath\pi(n-1)/N); \ 1 \leq n \leq N)$ is the radio channel matrix of user k ($\imath = \sqrt{-1}$) in the discrete Fourier domain (here N is the number of frequency bins) and $\mathbf{V} = [\mathbf{v}_0, \cdots, \mathbf{v}_K]$ is the $N \times (K+1)$ signature matrix with i.i.d. elements as in the CDMA case above. Modeling this time the channel transfer functions as deterministic functions, we have $\sigma^2_{nk} = \frac{K}{N} |h_k(\exp(2\imath\pi(n-1)/N))|^2$.

In the downlink direction, we have

$$\mathbf{\Sigma} = \mathbf{H}\mathbf{V}\mathbf{P}^{1/2} \tag{8}$$

where $\mathbf{H} = \mathrm{diag}(h(\exp(2\imath\pi(n-1)/N); \ 1 \leq n \leq N)$ is the radio channel matrix in the discrete Fourier domain, the $N \times (K+1)$ signature matrix $\mathbf{V}$ is as above, and $\mathbf{P} = \mathrm{diag}(p_0, \ldots, p_K)$ is the matrix of the powers given to the different users. Model (8) coincides with the separable variance profile model (4) with $d_n = \frac{K}{N} |h(\exp(2\imath\pi(n-1)/N))|^2$ and $d_k = p_k$.

About the literature.

The asymptotic approximation $\overline{\beta}_K$ (first order result) is connected with the asymptotic eigenvalue distribution of Gram matrices $\mathbf{Y}\mathbf{Y}^*$ where elements of $\mathbf{Y}$ are described by the model (3), and can be found in the mathematical LRMT literature in the work of Girko [13] (see also [14] and [15]). Applications in the field of wireless communications can be found in e.g. [6] in the separable case and in [8] in the general variance profile case.

Concerning the CLT for $\beta_K - \overline{\beta}_K$ (second order result), only some particular cases of the general model (3) have been considered in the literature among which the i.i.d. case ($\sigma^2_{nk} = 1$) is studied in [16] (and based on a result of [17] pertaining to the asymptotic behavior of the eigenvectors of $\mathbf{Y}\mathbf{Y}^*$). The more general CDMA model (6) has been considered in [18], using a result of [19]. The model used in this paper includes the models of [16] and [18] as particular cases.

Fluctuations of other performance indexes such as Shannon's mutual information $\mathbb{E}\log\det\left(\frac{\Sigma\Sigma^*}{\rho}+\mathbf{I}_N\right)$ have also been studied at length. Let us cite [20] where the CLT is established in the separable case and [21] for a CLT in the general variance profile case. Similar results concerning the mutual information are found in [22] and in [23].

Limiting expressions vs K-dependent expressions.

As one may check in Theorems 2 and 3 below, we deliberately chose to provide deterministic expressions $\overline{\beta}_K$ and Θ_K^2 which remain bounded but do not necessarily converge as $K\to\infty$. For instance, Theorem 2 only states that $\beta_K-\overline{\beta}_K\to 0$ almost surely. No conditions which would guarantee the convergence of β_K are added. This approach has two advantages: $1)$ such expressions for $\overline{\beta}_K$ and Θ_K^2 exist for very general variance profiles (σ_{nk}^2) while limiting expressions may not, and $2)$ they provide a natural discretization which can easily be implemented.

The statements about these deterministic approximations are valid within the following asymptotic regime:

$$K\to\infty,\quad \liminf\frac{K}{N}>0\quad\text{and}\quad\limsup\frac{K}{N}<\infty\ . \tag{9}$$

Note that $\frac{K}{N}$ is not required to converge. In the remainder of the paper, the notation "$K\to\infty$" will refer to (9).

We note that in the particular case where $\frac{K}{N}\to\alpha>0$ and the variance profile is obtained by a regular sampling of a continuous function f i.e. $\sigma_{nk}^2=f\left(\frac{n}{N},\frac{k}{K+1}\right)$, it is possible to prove that $\overline{\beta}_K$ and Θ_K^2 converge towards limits that can be characterized by integral equations.

Principle of the approach.

The approach used here is simple and powerful. It is based on the approximation of β_K by the sum of a martingale difference sequence and on the use of the CLT for martingales [24]. We note that apart from the LRMT context, such a technique has been used recently in [25] to establish a CLT on general quadratic forms of the type $\mathbf{z}^*\mathbf{A}\mathbf{z}$ where $\mathbf{A}$ is a deterministic matrix and $\mathbf{z}$ is a random vector with i.i.d. elements.

Paper organization.

In Section II, first-order results, whose presentation and understanding is compulsory to state the CLT, are recalled. The CLT, which is the main contribution of this paper, is provided in Section III. In Section

IV, simulations and numerical illustrations are provided. The proof of the main theorem (Theorem 3) in given in Section V while the Appendix gathers proofs of intermediate results.

Notations.

Given a complex $N \times N$ matrix $\mathbf{X} = [x_{ij}]_{i,j=1}^{N}$, denote by $\|\mathbf{X}\|$ its spectral norm, and by $|\!|\!|\mathbf{X}|\!|\!|_\infty$ its maximum row sum norm, i.e., $|\!|\!|\mathbf{X}|\!|\!|_\infty = \max_{1\le i\le N}\sum_{j=1}^{N}|x_{ij}|$. Denote by $\|\cdot\|$ the Euclidean norm of a vector and by $\|\cdot\|_\infty$ its $\max$ (or ℓ_∞) norm.

II. First Order Results: The SINR Deterministic Approximation

In the sequel, we shall often show explicitly the dependence on K in the notations. Consider the quadratic form (2):

$$\beta_K = \mathbf{y}^*\left(\mathbf{Y}\mathbf{Y}^* + \rho\mathbf{I}_N\right)^{-1}\mathbf{y}\ ,$$

where the sequence of matrices $\mathbf{\Sigma}(K) = [\mathbf{y}(K)\ \mathbf{Y}(K)]$ is given by

$$\mathbf{\Sigma}(K) \;=\; \left(\Sigma_{nk}(K)\right)_{n=1,k=0}^{N,K} \;=\; \left(\frac{\sigma_{nk}(K)}{\sqrt{K}}W_{nk}\right)_{n=1,k=0}^{N,K}\ .$$

Let us state the main assumptions:

A1*:* The complex random variables $(W_{nk};\ n\ge 1,\ k\ge 0)$ are i.i.d. with $\mathbb{E}W_{10}=0$, $\mathbb{E}W_{10}^2=0$, $\mathbb{E}|W_{10}|^2=1$ and $\mathbb{E}|W_{10}|^8<\infty$.

A2*:* There exists a real number $\sigma_{\max}<\infty$ such that

$$\sup_{K\ge 1}\ \max_{\substack{1\le n\le N\\ 0\le k\le K}} |\sigma_{nk}(K)| \le \sigma_{\max}\ .$$

Let $(a_m; 1\le m\le M)$ be complex numbers, then $\operatorname{diag}(a_m; 1\le m\le M)$ refers to the $M\times M$ diagonal matrix whose diagonal elements are the a_m's. If $\mathbf{A}=(a_{ij})$ is a square matrix, then $\operatorname{diag}(\mathbf{A})$ refers to the matrix $\operatorname{diag}(a_{ii})$. Consider the following diagonal matrices based on the variance profile along the columns and the rows of $\mathbf{\Sigma}$:

$$\begin{aligned}\mathbf{D}_k(K) &= \operatorname{diag}(\sigma_{1k}^2(K),\cdots,\sigma_{Nk}^2(K)), \quad 0\le k\le K\\ \widetilde{\mathbf{D}}_n(K) &= \operatorname{diag}(\sigma_{n1}^2(K),\cdots,\sigma_{nK}^2(K)), \quad 1\le n\le N.\end{aligned} \tag{10}$$

A3*:* The variance profile satisfies

$$\liminf_{K\ge 1}\ \min_{0\le k\le K}\frac{1}{K}\operatorname{tr}\mathbf{D}_k(K) > 0\ .$$

Since $\mathbb{E}|W_{10}|^2=1$, one has $\mathbb{E}|W_{10}|^4\ge 1$. The following is needed:

A4*:* At least one of the following conditions is satisfied:

$$\mathbb{E}|W_{10}|^4 > 1 \quad \text{or} \quad \liminf_K \frac{1}{K^2}\mathrm{tr}\left(\mathbf{D}_0(K)\sum_{k=1}^K \mathbf{D}_k(K)\right) > 0 \ .$$

Remark 1: If needed, one can attenuate the assumption on the eighth moment in **A1**. For instance, one can adapt without difficulty the proofs in this paper to the case where $\mathbb{E}|W_{10}|^{4+\epsilon} < \infty$ for $\varepsilon > 0$. We assumed $\mathbb{E}|W_{10}|^8 < \infty$ because at some places we rely on results of [21] which are stated with the assumption on the eighth moment.

Assumption **A3** is technical. It has already appeared in [26].

Assumption **A4** is necessary to get a non-vanishing variance Θ_K^2 in Theorem 3.

The following definitions will be of help in the sequel. A complex function $t(z)$ belongs to class $\mathcal{S}$ if $t(z)$ is analytical in the upper half plane $\mathbb{C}_+ = \{z \in \mathbb{C} \ ; \ \mathrm{im}(z) > 0\}$, if $t(z) \in \mathbb{C}_+$ for all $z \in \mathbb{C}_+$ and if $\mathrm{im}(z)|t(z)|$ is bounded over the upper half plane $\mathbb{C}_+$.

Denote by $\mathbf{Q}_K(z)$ and $\widetilde{\mathbf{Q}}_K(z)$ the resolvents of $\mathbf{Y}(K)\mathbf{Y}(K)^*$ and $\mathbf{Y}(K)^*\mathbf{Y}(K)$ respectively, that is the $N \times N$ and $K \times K$ matrices defined by:

$$\mathbf{Q}_K(z) = (\mathbf{Y}(K)\mathbf{Y}(K)^* - z\mathbf{I}_N)^{-1} \qquad \text{and} \qquad \widetilde{\mathbf{Q}}_K(z) = (\mathbf{Y}(K)^*\mathbf{Y}(K) - z\mathbf{I}_K)^{-1} \ .$$

A. The SINR Deterministic approximation

It is known [13], [26] that there exists a deterministic diagonal $N \times N$ matrix function $\mathbf{T}(z)$ that approximates the resolvent $\mathbf{Q}(z)$ in the following sense: Given a test matrix $\mathbf{S}$ with bounded spectral norm, the quantity $\frac{1}{K}\mathrm{tr}\ \mathbf{S}(\mathbf{Q}(z) - \mathbf{T}(z))$ converges a.s. to zero as $K \to \infty$. It is also known that the approximation $\overline{\beta}_K$ of the SINR β_K is simply related to $\mathbf{T}(z)$ (cf. Theorem 2). As we shall see, matrix $\mathbf{T}(z)$ also plays a fundamental role in the second order result (Theorem 3).

In the following theorem, we recall the definition and some of the main properties of $\mathbf{T}(z)$.

Theorem 1: The following hold true:

1) [26, Theorem 2.4] Let $(\sigma_{nk}^2(K);\ 1 \le n \le N;\ 1 \le k \le K)$ be a sequence of arrays of real numbers and consider the matrices $\mathbf{D}_k(K)$ and $\widetilde{\mathbf{D}}_n(K)$ defined in (10). The system of $N + K$ functional equations

$$\left\{\begin{array}{rcll} t_{n,K}(z) & = & \dfrac{-1}{z\left(1 + \frac{1}{K}\mathrm{tr}(\widetilde{\mathbf{D}}_n(K)\widetilde{\mathbf{T}}_K(z))\right)}, & 1 \le n \le N \\ \tilde{t}_{k,K}(z) & = & \dfrac{-1}{z\left(1 + \frac{1}{K}\mathrm{tr}(\mathbf{D}_k(K)\mathbf{T}_K(z))\right)}, & 1 \le k \le K \end{array}\right. \tag{11}$$

where

$$\mathbf{T}_K(z) = \mathrm{diag}(t_{1,K}(z), \ldots, t_{N,K}(z)), \quad \widetilde{\mathbf{T}}_K(z) = \mathrm{diag}(\tilde{t}_{1,K}(z), \ldots, \tilde{t}_{K,K}(z))$$

admits a unique solution $(\mathbf{T}, \widetilde{\mathbf{T}})$ among the diagonal matrices for which the $t_{n,K}$'s and the $\tilde{t}_{k,K}$'s belong to class $\mathcal{S}$. Moreover, functions $t_{n,K}(z)$ and $\tilde{t}_{k,K}(z)$ admit an analytical continuation over $\mathbb{C} - \mathbb{R}_+$ which is real and positive for $z \in (-\infty, 0)$.

2) [26, Theorem 2.5] Assume that Assumptions **A1** and **A2** hold true. Consider the sequence of random matrices $\mathbf{Y}(K)\mathbf{Y}(K)^*$ where $\mathbf{Y}$ has dimensions $N \times K$ and whose entries are given by $\mathbf{Y}_{nk} = \frac{\sigma_{nk}}{\sqrt{K}} W_{nk}$. For every sequence $\mathbf{S}_K$ of $N \times N$ diagonal matrices and every sequence $\widetilde{\mathbf{S}}_K$ of $K \times K$ diagonal matrices with

$$\sup_K \max\left(\|\mathbf{S}_K\|, \|\widetilde{\mathbf{S}}_K\|\right) < \infty \,,$$

the following limits hold true almost surely:

$$\begin{aligned}
\lim_{K\to\infty} \frac{1}{K}\mathrm{tr}\ \mathbf{S}_K \left(\mathbf{Q}_K(z) - \mathbf{T}_K(z)\right) &= 0, \qquad \forall z \in \mathbb{C} - \mathbb{R}_+, \\
\lim_{K\to\infty} \frac{1}{K}\mathrm{tr}\ \widetilde{\mathbf{S}}_K \left(\widetilde{\mathbf{Q}}_K(z) - \widetilde{\mathbf{T}}_K(z)\right) &= 0, \qquad \forall z \in \mathbb{C} - \mathbb{R}_+ \ .
\end{aligned}$$

The following lemma which reproduces [27, Lemma 2.7] will be used throughout the paper. It characterizes the asymptotic behavior of an important class of quadratic forms:

Lemma 1: Let $\mathbf{x} = [X_1, \ldots, X_N]^{\mathrm{T}}$ be a $N \times 1$ vector where the X_n are centered i.i.d. complex random variables with unit variance. Let $\mathbf{A}$ be a deterministic $N \times N$ complex matrix. Then, for any $p \geq 2$, there exists a constant C_p depending on p only such that

$$\mathbb{E}\left|\frac{1}{N}\mathbf{x}^*\mathbf{A}\mathbf{x} - \frac{1}{N}\mathrm{tr}(\mathbf{A})\right|^p \leq \frac{C_p}{N^p}\left(\left(\mathbb{E}|X_1|^4\mathrm{tr}(\mathbf{A}\mathbf{A}^*)\right)^{p/2} + \mathbb{E}|X_1|^{2p}\mathrm{tr}\left((\mathbf{A}\mathbf{A}^*)^{p/2}\right)\right) \ . \tag{12}$$

Noticing that $\mathrm{tr}(\mathbf{A}\mathbf{A}^*) \leq N\|\mathbf{A}\|^2$ and that $\mathrm{tr}\left((\mathbf{A}\mathbf{A}^*)^{p/2}\right) \leq N\|\mathbf{A}\|^p$, we obtain the simpler inequality

$$\mathbb{E}\left|\frac{1}{N}\mathbf{x}^*\mathbf{A}\mathbf{x} - \frac{1}{N}\mathrm{tr}(\mathbf{A})\right|^p \leq \frac{C_p}{N^{p/2}}\|\mathbf{A}\|^p\left(\left(\mathbb{E}|X_1|^4\right)^{p/2} + \mathbb{E}|X_1|^{2p}\right) \tag{13}$$

which is useful in case one has bounds on $\|\mathbf{A}\|$.

Using Theorem 1 and Lemma 1, we are in position to characterize the asymptotic behavior of the quadratic form β_K given by (2). We begin by rewriting β_K as

$$\beta_K = \frac{1}{K}\mathbf{w}_0^*\mathbf{D}_0^{1/2}\left(\mathbf{Y}\mathbf{Y}^* + \rho\mathbf{I}_N\right)^{-1}\mathbf{D}_0^{1/2}\mathbf{w}_0 = \frac{1}{K}\mathbf{w}_0^*\mathbf{D}_0^{1/2}\mathbf{Q}(-\rho)\mathbf{D}_0^{1/2}\mathbf{w}_0 \tag{14}$$

where the $N \times 1$ vector $\mathbf{w}_0$ is given by $\mathbf{w}_0 = [W_{10}, \ldots, W_{N0}]^{\mathrm{T}}$ and the diagonal matrix $\mathbf{D}_0$ is given by (10). Recall that $\mathbf{w}_0$ and $\mathbf{Q}$ are independent and that $\|\mathbf{D}_0\| \le \sigma^2_{\max}$ by **A2**. Furthermore, one can easily notice that $\|\mathbf{Q}(-\rho)\| = \|(\mathbf{YY}^* + \rho\mathbf{I})^{-1}\| \le 1/\rho$.

Denote by $\mathbb{E}_{\mathbf{Q}}$ the conditional expectation with respect to $\mathbf{Q}$, i.e. $\mathbb{E}_{\mathbf{Q}} = \mathbb{E}(\,\cdot\,\|\mathbf{Q})$. From Inequality (13), there exists a constant $C > 0$ for which

$$\begin{aligned}
\mathbb{E}\mathbb{E}_{\mathbf{Q}} \left| \beta_K - \frac{1}{K}\mathrm{tr}\mathbf{D}_0\mathbf{Q}(-\rho) \right|^4 &\le \frac{C}{K^2}\left(\frac{N}{K}\right)^2 \mathbb{E}\|\mathbf{D}_0\mathbf{Q}\|^4 \left((\mathbb{E}|W_{10}|^4)^2 + \mathbb{E}|W_{10}|^8\right) \\
&\le \frac{C}{K^2}\left(\frac{N}{K}\right)^2 \left(\frac{\sigma^2_{\max}}{\rho}\right)^4 \left((\mathbb{E}|W_{10}|^4)^2 + \mathbb{E}|W_{10}|^8\right) \\
&= \mathcal{O}\left(\frac{1}{K^2}\right) .
\end{aligned}$$

By the Borel-Cantelli Lemma, we therefore have

$$\beta_K - \frac{1}{K}\mathrm{tr}(\mathbf{D}_0\mathbf{Q}(-\rho)) \xrightarrow[K\to\infty]{} 0 \quad \text{a.s.}$$

Using this result, simply apply Theorem 1–(2) with $\mathbf{S} = \mathbf{D}_0$ (recall that $\|\mathbf{D}_0\| \le \sigma^2_{\max}$) to obtain:

Theorem 2: Let $\overline{\beta}_K = \dfrac{1}{K}\mathrm{tr}(\mathbf{D}_0(K)\mathbf{T}_K(-\rho))$ where $\mathbf{T}_K$ is given by Theorem 1–(1). Assume **A1** and **A2**. Then

$$\beta_K - \overline{\beta}_K \xrightarrow[K\to\infty]{} 0 \quad \text{a.s.}$$

B. The deterministic approximation in the separable case

In the separable case $\sigma_{nk}(K) = d_n(K)\tilde{d}_k(K)$, matrices $\mathbf{D}_k(K)$ and $\widetilde{\mathbf{D}}_n(K)$ are written as $\mathbf{D}_k(K) = \tilde{d}_k(K)\mathbf{D}(K)$ and $\widetilde{\mathbf{D}}_n(K) = d_n(K)\widetilde{\mathbf{D}}(K)$ where $\mathbf{D}(K)$ and $\widetilde{\mathbf{D}}(K)$ are the diagonal matrices

$$\mathbf{D}(K) = \mathrm{diag}(d_1(K), \ldots, d_N(K)), \quad \widetilde{\mathbf{D}}(K) = \mathrm{diag}(\tilde{d}_1(K), \ldots, \tilde{d}_K(K)) \ . \tag{15}$$

and one can check that the system of $N + K$ equations leading to $\mathbf{T}_K$ and $\widetilde{\mathbf{T}}_K$ simplifies into a system of two equations, and Theorem 1 takes the following form:

Proposition 1: [26, Sec. 3.2]

1) Assume $\sigma^2_{nk}(K) = d_n(K)\tilde{d}_k(K)$. Given $\rho > 0$, the system of two equations

$$\left\{ \begin{array}{rcl} \delta_K(\rho) &=& \frac{1}{K}\mathrm{tr}\left(\mathbf{D}\left(\rho(\mathbf{I}_N + \tilde{\delta}_K(\rho)\mathbf{D})\right)^{-1}\right) \\ \tilde{\delta}_K(\rho) &=& \frac{1}{K}\mathrm{tr}\left(\widetilde{\mathbf{D}}\left(\rho(\mathbf{I}_K + \delta_K(\rho)\widetilde{\mathbf{D}})\right)^{-1}\right) \end{array} \right. \tag{16}$$

where $\mathbf{D}$ and $\widetilde{\mathbf{D}}$ are given by (15) admits a unique solution $(\delta_K(\rho), \tilde{\delta}_K(\rho))$. Moreover, in this case matrices $\mathbf{T}(-\rho)$ and $\widetilde{\mathbf{T}}(-\rho)$ provided by Theorem 1–(1) coincide with

$$\mathbf{T}(-\rho) = \frac{1}{\rho}(\mathbf{I} + \tilde{\delta}(\rho)\mathbf{D})^{-1} \quad \text{and} \quad \widetilde{\mathbf{T}}(-\rho) = \frac{1}{\rho}(\mathbf{I} + \delta(\rho)\widetilde{\mathbf{D}})^{-1} \ . \tag{17}$$

2) Assume that **A1** and **A2** hold true. Let matrices $\mathbf{S}_K$ and $\widetilde{\mathbf{S}}_K$ be as in Theorem 1–(2). Then, almost surely $\frac{1}{K}\mathrm{tr}\left(\mathbf{S}_K\left(\mathbf{Q}_K(-\rho) - \mathbf{T}_K(-\rho)\right)\right) \to 0$ and $\frac{1}{K}\mathrm{tr}\left(\widetilde{\mathbf{S}}_K\left(\widetilde{\mathbf{Q}}_K(-\rho) - \widetilde{\mathbf{T}}_K(-\rho)\right)\right) \to 0$ as $K \to \infty$.

With these equations we can adapt the result of Theorem 2 to the separable case. Notice that $\mathbf{D}_0 = \tilde{d}_0\mathbf{D}$ and that $\delta(\rho)$ given by the system (16) coincides with $\frac{1}{K}\mathrm{tr}(\mathbf{DT})$, hence

Proposition 2: Assume that $\sigma^2_{nk}(K) = d_n(K)\tilde{d}_k(K)$, and that **A1** and **A2** hold true. Then

$$\frac{\beta_K}{\tilde{d}_0} - \delta_K(\rho) \xrightarrow[K\to\infty]{} 0 \quad \text{a.s.}$$

where $\delta_K(\rho)$ is given by Proposition 1–(1).

Let us provide a more explicit expression of δ_K which will be used in Section IV to illustrate the SINR behavior for the MIMO Model (5) and for MC-CDMA downlink Model (8). By combining the two equations in System (16), it turns out that $\delta = \delta_K(\rho)$ is the unique solution of the implicit equation

$$\delta = \frac{1}{K}\sum_{n=0}^{N-1} \frac{d_n}{\rho + \frac{1}{K}d_n\sum_{k=1}^K \frac{p_k}{1+p_k\delta}} \ . \tag{18}$$

Recall that in the case of the MIMO model (5), $d_n = \lambda_n$ and $\tilde{d}_k = p_k$, while in the case of the MC-CDMA downlink model (8), $d_n = \frac{K}{N}|h(\exp(2\imath\pi(n-1)/N)|^2$ and $\tilde{d}_k = p_k$ again. Here $\tilde{d}_0 = p_0$ is the power of the user of interest (user 0), and therefore $\beta_K/\tilde{d}_0$ is the normalized SINR of this user. Notice that $\delta_K(\rho)$ is almost the same for all users, hence the normalized SINRs for all users are close to each other for large K. Their common deterministic approximation is given by (18) which is the discrete analogue of the integral equation (16) in [6].

This example will be continued in Section III.

III. Second order results: The Central Limit Theorem

The following theorem is the main result of this paper. Its proof is postponed to Section V.

Theorem 3: 1) Assume that **A2**, **A3** and **A4** hold true. Let $\mathbf{A}_K$ and $\mathbf{\Delta}_K$ be the $K \times K$ matrices

$$\mathbf{A}_K = \left[\frac{1}{K} \frac{\frac{1}{K}\mathrm{tr}\mathbf{D}_\ell \mathbf{D}_m \mathbf{T}(-\rho)^2}{\left(1 + \frac{1}{K}\mathrm{tr}\mathbf{D}_\ell \mathbf{T}(-\rho)\right)^2} \right]_{\ell,m=1}^{K} \quad \text{and} \tag{19}$$

$$\mathbf{\Delta}_K = \mathrm{diag}\left(\left(1 + \frac{1}{K}\mathrm{tr}\mathbf{D}_\ell \mathbf{T}(-\rho)\right)^2 ; 1 \le \ell \le K \right) ,$$

where $\mathbf{T}$ is defined in Theorem 1–(1). Let $\mathbf{g}_K$ be the $K \times 1$ vector

$$\mathbf{g}_K = \left[\frac{1}{K}\mathrm{tr}\mathbf{D}_0 \mathbf{D}_1 \mathbf{T}(-\rho)^2, \cdots, \frac{1}{K}\mathrm{tr}\mathbf{D}_0 \mathbf{D}_K \mathbf{T}(-\rho)^2 \right]^{\mathrm{T}} .$$

Then the sequence of real numbers

$$\Theta_K^2 = \frac{1}{K}\mathbf{g}^{\mathrm{T}}(\mathbf{I}_K - \mathbf{A})^{-1}\mathbf{\Delta}^{-1}\mathbf{g} + (\mathbb{E}|W_{10}|^4 - 1)\frac{1}{K}\mathrm{tr}\mathbf{D}_0^2\mathbf{T}(-\rho)^2 \tag{20}$$

is well defined and furthermore

$$0 < \liminf_K \Theta_K^2 \le \limsup_K \Theta_K^2 < \infty .$$

2) Assume in addition **A1**. Then the sequence $\beta_K = \mathbf{y}^*(\mathbf{Y}\mathbf{Y}^* + \rho\mathbf{I})^{-1}\mathbf{y}$ satisfies

$$\frac{\sqrt{K}}{\Theta_K}\left(\beta_K - \overline{\beta}_K\right) \xrightarrow[K\to\infty]{} \mathcal{N}(0,1)$$

in distribution where $\overline{\beta}_K = \frac{1}{K}\mathrm{tr}\,\mathbf{D}_0\mathbf{T}_K$ is defined in the statement of Theorem 2.

Remark 2: (Comparison with other performance indexes) It is interesting to compare the "Mean Squared Error" (MSE) related to the SINR β_K: $MSE(\beta_K) = \mathbb{E}(\beta_K - \overline{\beta}_K)^2$, with the MSE related to Shannon's mutual information per transmit dimension $I = \frac{1}{K}\log\det(\rho\mathbf{\Sigma}\mathbf{\Sigma}^* + \mathbf{I})$ (studied in [21], [22] for instance):

$$MSE(\beta_K) \propto \mathcal{O}\left(\frac{1}{K}\right) \quad \text{while} \quad MSE(I) \propto \mathcal{O}\left(\frac{1}{K^2}\right) .$$

Remark 3: (On the achievability of the minimum of the variance) Recall that the variance writes

$$\Theta_K^2 = \frac{1}{K}\mathbf{g}^{\mathrm{T}}(\mathbf{I}_K - \mathbf{A})^{-1}\mathbf{\Delta}^{-1}\mathbf{g} + (\mathbb{E}|W_{10}|^4 - 1)\frac{1}{K}\mathrm{tr}\,\mathbf{D}_0^2\mathbf{T}^2 .$$

As $\mathbb{E}|W_{10}|^2 = 1$, one clearly has $\mathbb{E}|W_{10}|^4 - 1 \ge 0$ with equality if and only if $|W_{10}| = 1$ with probability one. Moreover, we shall prove in the sequel (Section V-B) that $\liminf_K \frac{1}{K}\mathbf{D}_0(K)\mathbf{T}_K^2 > 0$. Therefore $(\mathbb{E}|W_{10}|^4 - 1)\frac{1}{K}\mathrm{tr}\,\mathbf{D}_0^2\mathbf{T}^2$ is nonnegative, and is zero if and only if $|W_{10}| = 1$ with probability one. As a consequence, Θ_K^2 is minimum with respect to the distribution of the W_{nk} if and only if these random variables have their values on the unit circle. In the context of CDMA and MC-CDMA, this is the case when the signature matrix elements are elements of a PSK constellation. In multi-antenna systems, the

W_{nk}'s are frequently considered as Gaussian which induces a penalty on the SINR asymptotic MSE with respect to the unit norm case.

In the separable case, $\Theta_K^2 = \tilde{d}_0^2 \Omega_K^2$ where Ω_K^2 is given by the following corollary.

Corollary 1: Assume that **A2** is satisfied and that $\sigma_{nk}^2 = d_n \tilde{d}_k$. Assume moreover that

$$\min\left(\liminf_K \frac{1}{K}\mathrm{tr}(\mathbf{D}(K)), \liminf_K \frac{1}{K}\mathrm{tr}(\widetilde{\mathbf{D}}(K))\right) > 0 \tag{21}$$

where $\mathbf{D}$ and $\widetilde{\mathbf{D}}$ are given by (15). Let $\gamma = \frac{1}{K}\mathrm{tr}\mathbf{D}^2\mathbf{T}^2$ and $\tilde{\gamma} = \frac{1}{K}\mathrm{tr}\widetilde{\mathbf{D}}^2\widetilde{\mathbf{T}}^2$. Then the sequence

$$\Omega_K^2 = \gamma\left(\frac{\rho^2\gamma\tilde{\gamma}}{1-\rho^2\gamma\tilde{\gamma}} + \left(\mathbb{E}|W_{10}|^4 - 1\right)\right) \tag{22}$$

satisfies $0 < \liminf_K \Omega_K^2 \le \limsup_K \Omega_K^2 < \infty$. If, in addition, **A1** holds true, then:

$$\frac{\sqrt{K}}{\Omega_K}\left(\frac{\beta_K}{\tilde{d}_0} - \delta_K\right) \xrightarrow[K\to\infty]{} \mathcal{N}(0,1)$$

in distribution.

Remark 4: Condition (21) is the counterpart of Assumption **A3** in the case of a separable variance profile and suffices to establish $0 < \liminf_K(1-\rho^2\gamma\tilde{\gamma}) \le \limsup_K(1-\rho^2\gamma\tilde{\gamma}) < 1$ (see for instance [20]), hence the fact that $0 < \liminf_K \Omega_K^2 \le \limsup_K \Omega_K^2 < \infty$. The remainder of the proof of Corollary 1 is postponed to Appendix B.

Remark 5: As a direct application of Corollary 1 (to be used in Section IV below), let us provide the expressions of γ and $\tilde{\gamma}$ for the MIMO Model (5) or MC-CDMA downlink Model (8). From (15)–(17), we get

$$\begin{aligned}
\gamma &= \frac{1}{K}\sum_{n=0}^{N-1}\left(\frac{d_n}{\rho+\rho d_n\tilde{\delta}}\right)^2 = \frac{1}{K}\sum_{n=0}^{N-1}\left(\frac{d_n}{\rho+\frac{1}{K}d_n\sum_{k=1}^{K}\frac{p_k}{1+p_k\delta}}\right)^2 \\
\tilde{\gamma} &= \frac{1}{K}\sum_{k=1}^{K}\left(\frac{p_k}{\rho+\rho p_k\delta}\right)^2
\end{aligned}$$

where we recall that $d_n = \lambda_n$ for Model (5), $d_n = \frac{K}{N}|h(\exp(2\imath\pi(n-1)/N)|^2$ for Model (8), and δ is the solution of (18).

IV. Simulations

A. The general (non necessarily separable) case

In this section, the accuracy of the Gaussian approximation is verified by simulation. In order to validate the results of Theorems 2 and 3 for practical values of K, we consider the example of a MC-CDMA transmission in the uplink direction. We recall that K is the number of interfering users in this context. In the simulation, the discrete time channel impulse response of user k is represented by the vector with $L = 5$ coefficients $\mathbf{g}_k = [g_{k,0}, \ldots, g_{k,L-1}]^{\mathrm{T}}$. In the simulations, these vectors are generated pseudo-randomly according to the complex multivariate Gaussian law $\mathcal{CN}(0, 1/L\mathbf{I}_L)$. Setting the number of frequency bins to N, the channel matrix $\mathbf{H}_k$ for user k in the frequency domain (see Eq. (7)) is $\mathbf{H}_k = \mathrm{diag}(h_k(\exp(2\imath\pi(n-1)/N); 1 \le n \le N)$ where $h_k(z) = \frac{\sqrt{P_k}}{\|\mathbf{g}_k\|} \sum_{l=0}^{L-1} g_{k,l} z^{-l}$, the norm $\|\mathbf{g}_k\|$ is the Euclidean norm of $\mathbf{g}_k$ and P_k is the power received from user k. Concerning the distribution of the user powers P_k, we assume that these are arranged into five power classes with powers $P, 2P, 4P, 8P$ and $16P$ with relative frequencies given by Table IV-A. The user of interest (User 0) is assumed to belong

TABLE I

POWER CLASSES AND RELATIVE FREQUENCIES

Class	1	2	3	4	5
Power	P	$2P$	$4P$	$8P$	$16P$
Relative frequency	$1/8$	$1/4$	$1/4$	$1/8$	$1/4$

to Class 1. Finally, we assume that the number K of interfering users is set to $K = N/2$.

In Figure 1, the Signal over Noise Ratio (SNR) P/ρ for the user of interest is fixed to 10 dB. The evolution of $K\mathbb{E}(\beta_K - \bar{\beta}_K)^2/\Theta_K^2$ for this user (where $\mathbb{E}(\beta_K - \bar{\beta}_K)^2$ is measured numerically) is shown with respect to K. We note that this quantity is close to one for values of K as small as $K = 8$.

In Figure 2, K is set to $K = 64$, and the SINR normalized MSE $K\mathbb{E}(\beta_K - \bar{\beta}_K)^2/\Theta_K^2$ is plotted with respect to the input SNR P/ρ. This figure also confirms the fact that the MSE asymptotic approximation is highly accurate.

Figure 3 shows the histogram of $\sqrt{K}(\beta_K - \bar{\beta}_K)/\Theta_K$ for $N = 16$ and $N = 64$. This figure gives an idea of the similarity between the distribution of $\sqrt{K}(\beta_K - \bar{\beta}_K)/\Theta_K$ and $\mathcal{N}(0,1)$.

More precisely, Figure 4 quantifies this similarity through a Quantile-Quantile plot.

B. The separable case

In order to test the results of Proposition 2 and Corollary 1, we consider the following multiple antenna (MIMO) model with exponentially decaying correlation at reception:

$$\mathbf{\Sigma} = \frac{1}{\sqrt{K}} \mathbf{\Psi}^{1/2} \mathbf{W} \mathbf{P}^{1/2}$$

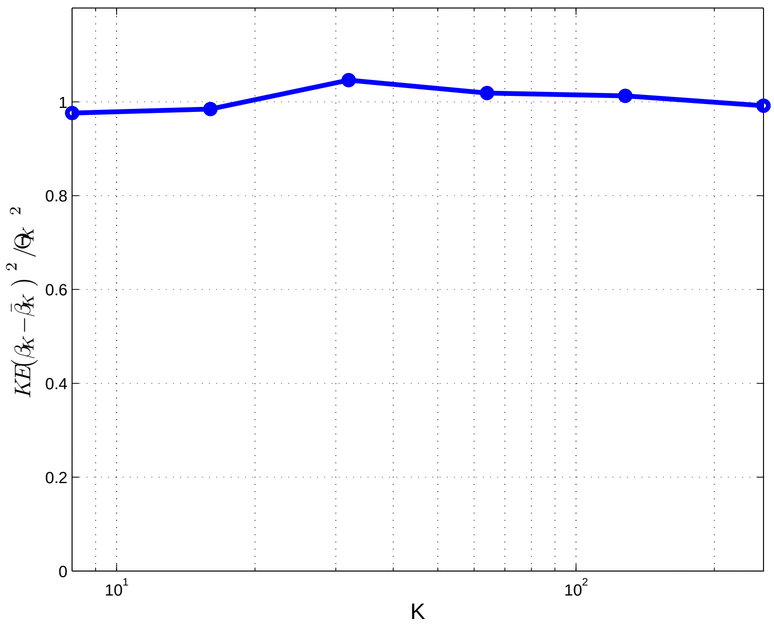

Fig. 1. SINR normalized MSE vs K

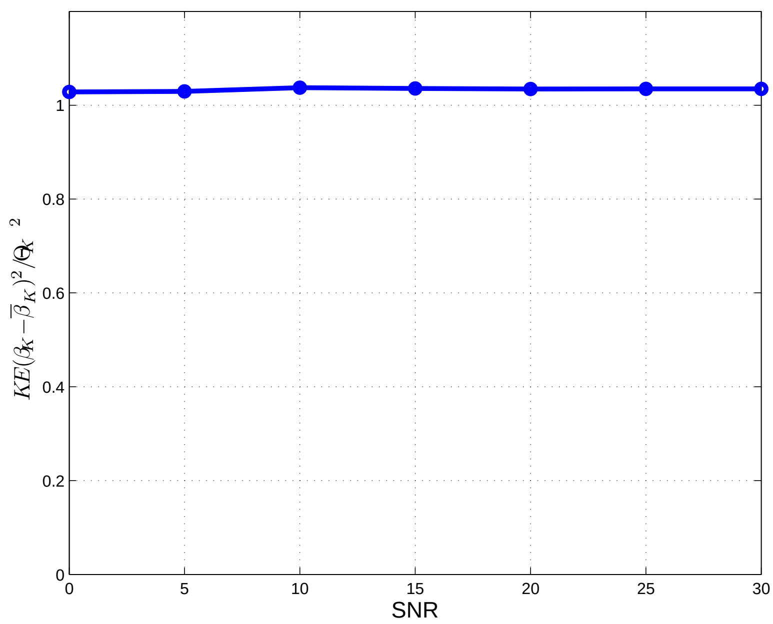

Fig. 2. SINR normalized MSE vs SNR

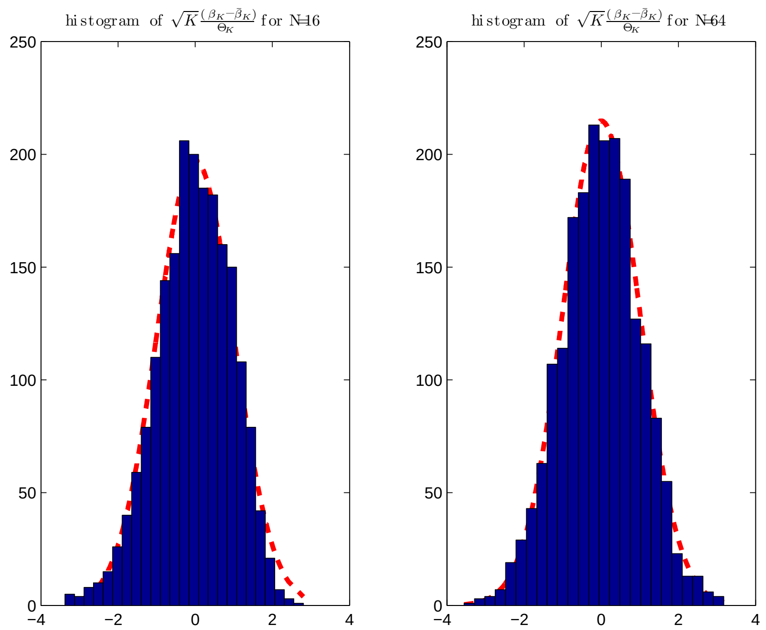

Fig. 3. Histogram of $\sqrt{K}(\beta_K - \bar{\beta}_K)$ for $N = 16$ and $N = 64$.

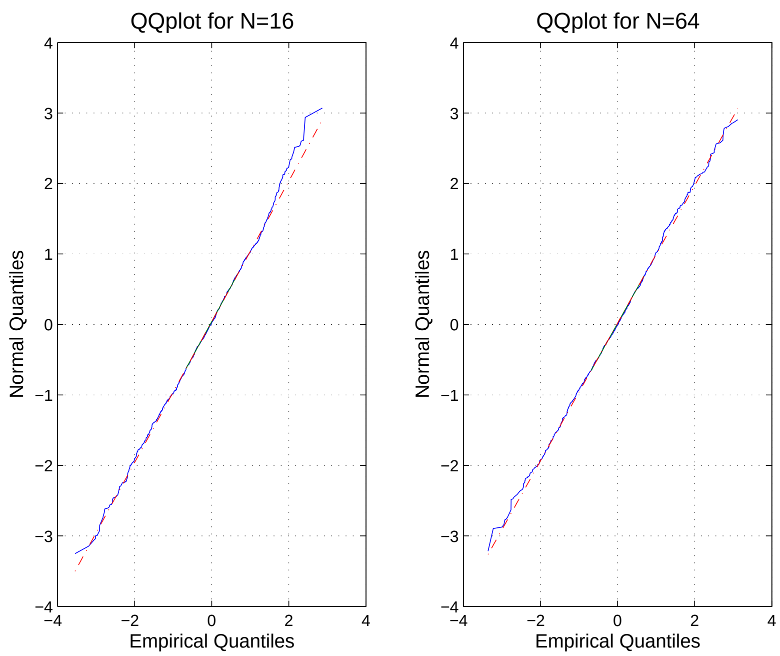

Fig. 4. Q-Q plot for $\sqrt{K}(\beta_K - \bar{\beta}_K)$, $N = 16$ and $N = 64$; dash doted line is the 45 degree line.

where $\mathbf{\Psi} = [a^{m-n}]_{m,n=0}^{N-1}$ with $0 < a < 1$ is the covariance matrix that accounts for the correlations at the receiver side, $P = \text{diag}\,(p_0, \cdots, p_K)$ is the matrix of the powers given to the different sources and W is a $N \times (K+1)$ matrix with Gaussian standard iid elements. Let $\mathbf{P_u}$ denote the vector containing the powers of the interfering sources. We set $\mathbf{P_u}$ (up to a permutation of its elements) to:

$$\mathbf{P_u} = \begin{cases} [4P \quad 5P] \quad \text{if} \quad K = 2 \\ [P \quad P \quad 2P \quad 4P] \quad \text{if} \quad K = 4 \\ [P \quad P \quad 2P \quad 2P \quad 2P \quad 4P \quad 4P \quad 4P \quad 8P \quad 16P \quad 16P \quad 16P\ ] \text{ if } K = 12\ . \end{cases}$$

For $K = 2^p$ with $3 \leq p \leq 7$, we assume that the powers of the interfering sources are arranged into 5 classes as in Table IV-A. We set the SNR P/ρ to 10 dB and a to 0.1. We investigate in this section the accuracy of the Gaussian approximation in terms of the outage probability. In Fig.5, we compare the empirical 1% outage SINR with the one predicted by the Central Limit Theorem. We note that the Gaussian approximation tends to under estimate the 1% outage SINR. We also note that it has a good accuracy for small values of α and for enough large values of N ($N \geq 64$).

Observe that all these simulations confirm a fact announced in Remark 2 above: compared with functionals of the channel singular values such as Shannon's mutual information, larger signal dimensions are needed to attain the asymptotic regime for quadratic forms such as the SINR (see for instance outage probability approximations for mutual information in [22] and in [23]). This observation holds for first order as well as for second order results.

V. Proof of Theorem 3

This section is devoted to the proof of Theorem 3. We begin with mathematical preliminaries.

A. Preliminaries

The following lemma gathers useful matrix results, whose proofs can be found in [28]:

Lemma 2: Assume $\mathbf{X} = [x_{ij}]_{i,j=1}^{N}$ and $\mathbf{Y}$ are complex $N \times N$ matrices. Then

1) For every $i, j \leq N$, $|x_{ij}| \leq \|\mathbf{X}\|$. In particular, $\|\text{diag}(\mathbf{X})\| \leq \|\mathbf{X}\|$.
2) $\|\mathbf{XY}\| \leq \|\mathbf{X}\|\,\|\mathbf{Y}\|$.
3) For $\rho > 0$, the resolvent $(\mathbf{XX}^* + \rho\mathbf{I})^{-1}$ satisfies $\|(\mathbf{XX}^* + \rho\mathbf{I})^{-1}\| \leq \rho^{-1}$.
4) If $\mathbf{Y}$ is Hermitian nonnegative, then $|\text{tr}(\mathbf{XY})| \leq \|\mathbf{X}\|\text{tr}(\mathbf{Y})$.

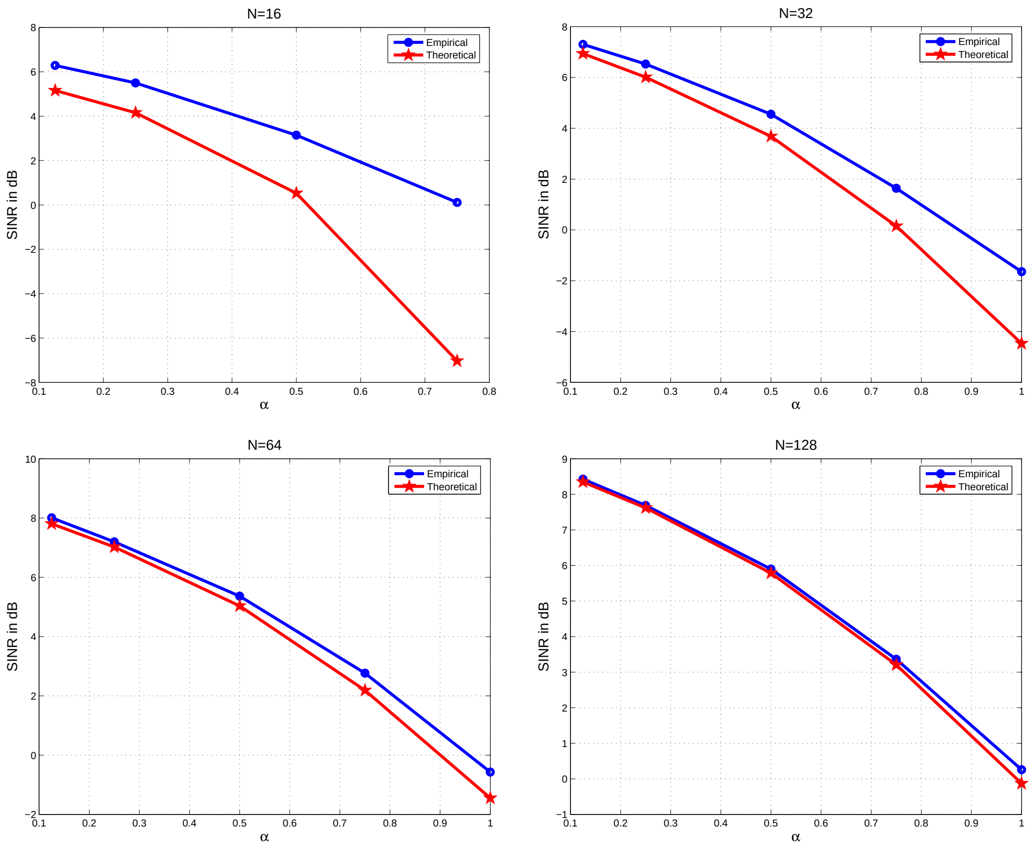

Fig. 5. Theoretical and empirical 1% outage SINR

Let $\mathbf{X} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^*$ be a spectral decomposition of $\mathbf{X}$ where $\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ is the matrix of singular values of $\mathbf{X}$. For a real $p \geq 1$, the Schatten ℓ_p-norm of $\mathbf{X}$ is defined as $\|\mathbf{X}\|_p = \left(\sum \lambda_i^p\right)^{1/p}$. The following bound over the Schatten ℓ_p-norm of a triangular matrix will be of help (for a proof, see [25], [29, page 278]):

Lemma 3: Let $\mathbf{X} = [x_{ij}]_{i,j=1}^N$ be a $N \times N$ complex matrix and let $\widetilde{\mathbf{X}} = [x_{ij}\mathbf{1}_{i>j}]_{i,j=1}^N$ be the strictly lower triangular matrix extracted from $\mathbf{X}$. Then for every $p \geq 1$, there exists a constant C_p depending on p only such that

$$\|\widetilde{\mathbf{X}}\|_p \leq C_p \|\mathbf{X}\|_p \ .$$

The following lemma lists some properties of the resolvent $\mathbf{Q}$ and the deterministic approximation matrix $\mathbf{T}$. Its proof is postponed to Appendix A.

Lemma 4: The following facts hold true:

1) Assume **A2**. Consider matrices $\mathbf{T}_K(-\rho) = \mathrm{diag}(t_1(-\rho), \ldots, t_N(-\rho))$ defined by Theorem 1–(1). Then for every $1 \leq n \leq N$,
$$\frac{1}{\rho + \sigma^2_{\max}} \leq t_n(-\rho) \leq \frac{1}{\rho} \ . \tag{23}$$
2) Assume in addition **A1** and **A3**. Let $\mathbf{Q}_K(-\rho) = (\mathbf{Y}\mathbf{Y}^* + \rho \mathbf{I})^{-1}$ and let matrices $\mathbf{S}_K$ be as in the statement of Theorem 1–(2). Then
$$\sup_K \mathbb{E} \left| \mathrm{tr}\ \mathbf{S}_K(\mathbf{Q}_K - \mathbf{T}_K) \right|^2 < \infty \ . \tag{24}$$

B. Proof of Theorem 3–(1)

We introduce the following notations. Assume that $\mathbf{X}$ is a real matrix, by $\mathbf{X} \succcurlyeq \mathbf{0}$ we mean $X_{ij} \geq 0$ for every element X_{ij}. For a vector $\mathbf{x}$, $\mathbf{x} \succcurlyeq \mathbf{0}$ is defined similarly. In the remainder of the paper, $C = C(\rho, \sigma^2_{\max}, \liminf \frac{N}{K}, \sup \frac{N}{K}) < \infty$ denotes a positive constant whose value may change from line to line.

The following lemma, which directly follows from [21, Lemma 5.2 and Proposition 5.5], states some important properties of the matrices $\mathbf{A}_K$ defined in the statement of Theorem 3.

Lemma 5: Assume **A2** and **A3**. Consider matrices $\mathbf{A}_K$ defined by (19). Then the following facts hold true:

1) Matrix $\mathbf{I}_K - \mathbf{A}_K$ is invertible, and $(\mathbf{I}_K - \mathbf{A}_K)^{-1} \succcurlyeq \mathbf{0}$.
2) Element (k,k) of the inverse satisfies $\left[(\mathbf{I}_K - \mathbf{A}_K)^{-1} \right]_{k,k} \geq 1$ for every $1 \leq k \leq K$.
3) The maximum row sum norm of the inverse satisfies $\limsup_K \left|\!\left|\!\left| (\mathbf{I}_K - \mathbf{A}_K)^{-1} \right|\!\right|\!\right|_\infty < \infty$.

Due to Lemma 5–(1), Θ^2_K is well defined. Let us prove that $\limsup_K \Theta^2_K < \infty$. The first term of the right-hand side of (20) satisfies

$$\begin{aligned} \frac{1}{K}\mathbf{g}^{\mathrm{T}}(\mathbf{I}_K - \mathbf{A}_K)^{-1}\boldsymbol{\Delta}^{-1}\mathbf{g} &\leq \|\mathbf{g}\|_\infty \|(\mathbf{I}_K - \mathbf{A}_K)^{-1}\boldsymbol{\Delta}^{-1}\mathbf{g}\|_\infty \\ &\leq \|\mathbf{g}\|_\infty \left|\!\left|\!\left| (\mathbf{I}_K - \mathbf{A}_K)^{-1} \right|\!\right|\!\right|_\infty \|\boldsymbol{\Delta}^{-1}\mathbf{g}\|_\infty \leq \|\mathbf{g}\|^2_\infty \left|\!\left|\!\left| (\mathbf{I}_K - \mathbf{A}_K)^{-1} \right|\!\right|\!\right|_\infty \end{aligned} \tag{25}$$

due to $\left\|\left\|\boldsymbol{\Delta}^{-1}\right\|\right\|_\infty \leq 1$. Recall that $\|\mathbf{T}\| \leq \rho^{-1}$ by Lemma 4–(1). Therefore, any element of $\mathbf{g}$ satisfies

$$\frac{1}{K}\mathrm{tr}\mathbf{D}_0\mathbf{D}_k\mathbf{T}^2 \leq \frac{N}{K}\|\mathbf{D}_0\|\|\mathbf{D}_k\|\|\mathbf{T}\|^2 \leq \frac{N}{K}\frac{\sigma^4_{\max}}{\rho^2} \tag{26}$$

by **A2**, hence $\sup_K \|\mathbf{g}\| \leq C$. From Lemma 5–(3) and (25), we then obtain

$$\limsup_K \frac{1}{K}\mathbf{g}^{\mathrm{T}}(\mathbf{I}_K - \mathbf{A}_K)^{-1}\boldsymbol{\Delta}^{-1}\mathbf{g} \leq C. \tag{27}$$

We can prove similarly that the second term in the right-hand side of (20) satisfies $\sup_K((\mathbb{E}|W_{10}|^4 - 1)\frac{1}{K}\mathrm{tr}\mathbf{D}_0^2\mathbf{T}(-\rho)^2) \leq C$. Hence $\limsup_K \Theta_K^2 < \infty$.

Let us prove that $\liminf_K \Theta_K^2 > 0$. We have

$$\begin{aligned}
\frac{1}{K}\mathbf{g}^{\mathrm{T}}(\mathbf{I}_K - \mathbf{A}_K)^{-1}\boldsymbol{\Delta}^{-1}\mathbf{g} &\overset{(a)}{\geq} \frac{1}{K}\mathbf{g}^{\mathrm{T}}\mathrm{diag}\left((\mathbf{I}_K - \mathbf{A}_K)^{-1}\right)\boldsymbol{\Delta}^{-1}\mathbf{g} \\
&\overset{(b)}{\geq} \frac{1}{\left(1+\frac{N}{K}\frac{\sigma^2_{\max}}{\rho}\right)^2}\frac{1}{K}\sum_{k=1}^K\left(\frac{1}{K}\mathrm{tr}\mathbf{D}_0\mathbf{D}_k\mathbf{T}^2\right)^2 \\
&\overset{(c)}{\geq} \frac{1}{\left(1+\frac{N}{K}\frac{\sigma^2_{\max}}{\rho}\right)^2}\left(\frac{1}{K^2}\mathrm{tr}\ \mathbf{D}_0\left(\sum_{k=1}^K\mathbf{D}_k\right)\mathbf{T}^2\right)^2 \\
&\overset{(d)}{\geq} \frac{1}{\left(1+\frac{N}{K}\frac{\sigma^2_{\max}}{\rho}\right)^2(\rho+\sigma^2_{\max})^4}\left(\frac{1}{K^2}\mathrm{tr}\ \mathbf{D}_0\sum_{k=1}^K\mathbf{D}_k\right)^2 \\
&\geq C\left(\frac{1}{K^2}\mathrm{tr}\ \mathbf{D}_0\sum_{k=1}^K\mathbf{D}_k\right)^2 ,
\end{aligned}$$

where (a) follows from the fact that $(\mathbf{I}_K - \mathbf{A}_K)^{-1} \succcurlyeq \mathbf{0}$ (Lemma 5–(1), and the straightforward inequalities $\boldsymbol{\Delta}^{-1} \succcurlyeq \mathbf{0}$ and $\mathbf{g} \succcurlyeq \mathbf{0}$), (b) follows from Lemma 5–(2) and $\|\boldsymbol{\Delta}\| \leq (1+\frac{N}{K}\frac{\sigma^2_{\max}}{\rho})^2$, (c) follows from the elementary inequality $n^{-1}\sum x_i^2 \geq (n^{-1}\sum x_i)^2$, and (d) is due to Lemma 4–(1). Similar derivations yield:

$$(\mathbb{E}|W_{10}|^4 - 1)\frac{1}{K}\mathrm{tr}\mathbf{D}_0^2\mathbf{T} \quad \geq \quad \frac{\mathbb{E}|W_{10}|^4 - 1}{(\rho+\sigma^2_{\max})^2}\left(\frac{1}{K}\mathrm{tr}\mathbf{D}_0\right)^2 \quad \geq \quad C(\mathbb{E}|W_{10}|^4 - 1)$$

by **A3**. Therefore, if **A4** holds true, then $\liminf_K \Theta_K^2 > 0$ and Theorem 3–(1) is proved.

C. Proof of Theorem 3–(2)

Recall that the SINR β_K is given by Equation (14). The random variable $\frac{\sqrt{K}}{\Theta_K}(\beta_K - \overline{\beta}_K)$ can therefore be decomposed as

$$\begin{aligned}
\frac{\sqrt{K}}{\Theta_K}(\beta_K - \overline{\beta}_K) &= \frac{1}{\sqrt{K}\Theta_K}\left(\mathbf{w}_0^*\mathbf{D}_0^{1/2}\mathbf{Q}\mathbf{D}_0^{1/2}\mathbf{w}_0 - \mathrm{tr}(\mathbf{D}_0\mathbf{Q})\right) + \frac{1}{\sqrt{K}\Theta_K}\left(\mathrm{tr}(\mathbf{D}_0(\mathbf{Q}-\mathbf{T}))\right) \\
&= U_{1,K} + U_{2,K}\ .
\end{aligned} \tag{28}$$

Thanks to Lemma 4–(2) and to the fact that $\liminf_K \Theta_K^2 > 0$, we have $\mathbb{E}U_{K,2}^2 < CK^{-1}$ which implies that $U_{K,2} \to 0$ in probability as $K \to \infty$. Hence, in order to conclude that

$$\frac{\sqrt{K}}{\Theta_K}(\beta_K - \overline{\beta}_K) \xrightarrow[K\to\infty]{} \mathcal{N}(0,1) \quad \text{in distribution} ,$$

it is sufficient by Slutsky's theorem to prove that $U_{1,K} \to \mathcal{N}(0,1)$ in distribution. The remainder of the section is devoted to this point.

Remark 6: Decomposition (28) and the convergence to zero (in probability) of $U_{2,K}$ yield the following interpretation: The fluctuations of $\sqrt{K}(\beta_K - \overline{\beta}_K)$ are mainly due to the fluctuations of vector $\mathbf{w}_0$. Indeed the contribution of the fluctuations[1] of $\frac{1}{K}\mathrm{tr}\mathbf{D}_0\mathbf{Q}$, due to the random nature of $\mathbf{Y}$, is negligible.

Denote by $\mathbb{E}_n$ the conditional expectation $\mathbb{E}_n[\,\cdot\,] = \mathbb{E}[\,\cdot \parallel W_{n,0}, W_{n+1,0}, \ldots, W_{N,0}, \mathbf{Y}]$. Put $\mathbb{E}_{N+1}[\,\cdot\,] = \mathbb{E}[\,\cdot \parallel \mathbf{Y}]$ and note that $\mathbb{E}_{N+1}(\mathbf{w}_0^*\mathbf{D}_0^{1/2}\mathbf{Q}\mathbf{D}_0^{1/2}\mathbf{w}_0) = \mathrm{tr}\mathbf{D}_0\mathbf{Q}$. With these notations at hand, we have:

$$U_{1,K} = \frac{1}{\Theta_K}\sum_{n=1}^{N}(\mathbb{E}_n - \mathbb{E}_{n+1})\frac{\mathbf{w}_0^*\mathbf{D}_0^{1/2}\mathbf{Q}\mathbf{D}_0^{1/2}\mathbf{w}_0}{\sqrt{K}} \triangleq \frac{1}{\Theta_K}\sum_{n=1}^{N} Z_{n,K} \ . \tag{29}$$

Consider the increasing sequence of $\sigma-$fields

$$\mathcal{F}_{N,K} = \sigma(W_{N,0}, \mathbf{Y}) \ , \quad \cdots \ , \quad \mathcal{F}_{1,K} = \sigma(W_{1,0}, \cdots, W_{N,0}, \mathbf{Y}) \ .$$

Then the random variable $Z_{n,K}$ is integrable and measurable with respect to $\mathcal{F}_{n,K}$; moreover it readily satisfies $\mathbb{E}_{n+1}Z_{n,K} = 0$. In particular, the sequence $(Z_{N,K}, \ldots, Z_{1,K})$ is a martingale difference sequence with respect to $(\mathcal{F}_{N,K}, \cdots, \mathcal{F}_{1,K})$. The following CLT for martingales is the key tool to study the asymptotic behavior of $U_{1,K}$:

Theorem 4: Let $X_{N,K}, X_{N-1,K}, \ldots, X_{1,K}$ be a martingale difference sequence with respect to the increasing filtration $\mathcal{G}_{N,K}, \ldots, \mathcal{G}_{1,K}$. Assume that there exists a sequence of real positive numbers s_K^2 such that

$$\frac{1}{s_K^2}\sum_{n=1}^{N}\mathbb{E}\left[X_{n,K}^2 \| \mathcal{G}_{n+1,K}\right] \xrightarrow[K\to\infty]{} 1$$

in probability. Assume further that the Lyapunov condition holds:

$$\exists \alpha > 0, \quad \frac{1}{s_K^{2(1+\alpha)}}\sum_{n=1}^{N}\mathbb{E}\left|X_{n,K}\right|^{2+\alpha} \xrightarrow[K\to\infty]{} 0 \ ,$$

[1]In fact, one may prove that the fluctuation of $\frac{1}{K}\mathrm{tr}\mathbf{D}_0(\mathbf{Q}-\mathbf{T})$ are of order K, i.e. $\mathrm{tr}\mathbf{D}_0(\mathbf{Q}-\mathbf{T})$ asymptotically behaves as a Gaussian random variable. Such a speed of fluctuations already appears in [21], when studying the fluctuations of the mutual information.

Then $s_K^{-1}\sum_{n=1}^N X_{n,K}$ converges in distribution to $\mathcal{N}(0,1)$ as $K \to \infty$.

Remark 7: This theorem is proved in [24], gathering Theorem 35.12 (which is expressed under the weaker Lindeberg condition) together with the arguments of Section 27 (where it is proved that Lyapunov's condition implies Lindeberg's condition).

In order to prove that

$$U_{1,K} = \frac{1}{\Theta_K}\sum_{n=1}^N Z_{n,K} \xrightarrow[K\to\infty]{} \mathcal{N}(0,1) \quad \text{in distribution} , \tag{30}$$

we shall apply Theorem 4 to the sum $\frac{1}{\Theta_K}\sum_{n=1}^N Z_{n,K}$ and the filtration $(\mathcal{F}_{n,K})$. The proof is carried out into four steps:

Step 1: We first establish Lyapunov's condition. Due to the fact that $\liminf_K \Theta_K^2 > 0$, we only need to show that

$$\exists\ \alpha > 0, \quad \sum_{n=1}^N \mathbb{E}|Z_{n,K}|^{2+\alpha} \xrightarrow[K\to\infty]{} 0\ . \tag{31}$$

Step 2: We prove that $V_K = \sum_{n=1}^N \mathbb{E}_{n+1} Z_{n,K}^2$ satisfies

$$V_K - \left(\frac{\left(\mathbb{E}|W_{10}|^4 - 2\right)}{K} \mathrm{tr}\left(\mathbf{D}_0^2 (\mathrm{diag}(\mathbf{Q}))^2\right) + \frac{1}{K}\mathrm{tr}(\mathbf{D}_0\mathbf{Q}\mathbf{D}_0\mathbf{Q}) \right) \xrightarrow[K\to\infty]{} 0 \quad \text{in probability} . \tag{32}$$

Step 3: We first show that

$$\frac{1}{K}\mathrm{tr}\mathbf{D}_0^2(\mathrm{diag}(\mathbf{Q}))^2 - \frac{1}{K}\mathrm{tr}\mathbf{D}_0^2\mathbf{T}^2 \xrightarrow[K\to\infty]{} 0 \quad \text{in probability.} \tag{33}$$

In order to study the asymptotic behavior of $\frac{1}{K}\mathrm{tr}(\mathbf{D}_0\mathbf{Q}\mathbf{D}_0\mathbf{Q})$, we introduce the random variables $U_\ell = \frac{1}{K}\mathrm{tr}(\mathbf{D}_0\mathbf{Q}\mathbf{D}_\ell\mathbf{Q})$ for $0 \le \ell \le K$ (the one of interest being U_0). We then prove that the U_ℓ's satisfy the following system of equations:

$$U_\ell = \sum_{k=1}^K c_{\ell k} U_k + \frac{1}{K}\mathrm{tr}\mathbf{D}_0\mathbf{D}_\ell\mathbf{T}^2 + \epsilon_\ell, \quad 0 \le \ell \le K, \tag{34}$$

where

$$c_{\ell k} = \frac{1}{K}\frac{\frac{1}{K}\mathrm{tr}\mathbf{D}_\ell\mathbf{D}_k\mathbf{T}(-\rho)^2}{\left(1+\frac{1}{K}\mathrm{tr}\mathbf{D}_k\mathbf{T}(-\rho)\right)^2}, \quad 0 \le \ell \le K,\ 1 \le k \le K \tag{35}$$

and the perturbations ϵ_ℓ satisfy $\mathbb{E}|\epsilon_\ell| \le CK^{-\frac{1}{2}}$ where we recall that C is independent of ℓ.

Step 4: We prove that $U_0 = \frac{1}{K}\mathrm{tr}\mathbf{D}_0\mathbf{Q}\mathbf{D}_0\mathbf{Q}$ satisfies

$$U_0 = \frac{1}{K}\mathrm{tr}\mathbf{D}_0^2\mathbf{T}^2 + \frac{1}{K}\mathbf{g}^{\mathrm{T}}\left(\mathbf{I}-\mathbf{A}\right)^{-1}\mathbf{\Delta}^{-1}\mathbf{g} + \epsilon \tag{36}$$

with $\mathbb{E}|\epsilon| \le CK^{-\frac{1}{2}}$. This equation combined with (32) and (33) yields $\sum_n \mathbb{E}_{n+1}Z_{n,K}^2 - \Theta_K^2 \to 0$ in probability. As $\liminf_K \Theta_K^2 > 0$, this implies $\frac{1}{\Theta_K}\sum_n \mathbb{E}_{n+1}Z_{n,K}^2 \to 1$ in probability, which proves (30) and thus ends the proof of Theorem 3.

Write $\mathbf{B} = [b_{ij}]_{i,j=1}^N = \mathbf{D}_0^{1/2}\mathbf{Q}\mathbf{D}_0^{1/2}$ and recall from (29) that $Z_{n,K} = \frac{1}{\sqrt{K}}(\mathbb{E}_n - \mathbb{E}_{n+1})\mathbf{w}_0^*\mathbf{B}\mathbf{w}_0$. We have

$$\mathbb{E}_n\mathbf{w}_0^*\mathbf{B}\mathbf{w}_0 = \sum_{\ell=1}^{n-1} b_{\ell\ell} + \sum_{\ell_1,\ell_2=n}^{N} W_{\ell_1 0}^* W_{\ell_2 0} b_{\ell_1\ell_2} \ .$$

Hence

$$Z_{n,K} = \frac{1}{\sqrt{K}}\left(\left(|W_{n0}|^2-1\right)b_{nn} + W_{n0}^*\sum_{\ell=n+1}^{N} W_{\ell 0}b_{n\ell} + W_{n0}\sum_{\ell=n+1}^{N} W_{\ell 0}^* b_{\ell n}\right) \ . \tag{37}$$

Step 1: Validation of the Lyapunov condition: The following inequality will be of help to check Lyapunov's condition.

Lemma 6 (Burkholder's inequality): Let X_k be a complex martingale difference sequence with respect to the increasing sequence of σ–fields $\mathcal{F}_k$. Then for $p \ge 2$, there exists a constant C_p for which

$$\mathbb{E}\left|\sum_k X_k\right|^p \le C_p\left(\mathbb{E}\left(\sum_k \mathbb{E}\left[|X_k|^2 \| \mathcal{F}_{k-1}\right]\right)^{p/2} + \mathbb{E}\sum_k |X_k|^p\right) \ .$$

Recall Assumption **A1**. Eq. (37) yields:

$$\begin{aligned} |Z_{n,K}|^4 &\le \frac{1}{K^2}\left(\frac{|W_{n0}|^2+1}{\rho\sigma_{\max}^2} + 2\left|W_{n0}\sum_{\ell=n+1}^{N} W_{\ell 0}b_{n\ell}\right|\right)^4 \\ &\le \frac{2^3}{K^2}\left(\left(\frac{|W_{n0}|^2+1}{\rho\sigma_{\max}^2}\right)^4 + 2^4\left|W_{n0}\sum_{\ell=n+1}^{N} W_{\ell 0}b_{n\ell}\right|^4\right) \end{aligned} \tag{38}$$

where we use the fact that $|b_{nn}| \le (\rho\sigma_{\max}^2)^{-1}$ (cf. Lemma 2–(1)) and the convexity of $x \mapsto x^4$. Due to Assumption **A1**, we have:

$$\mathbb{E}\left(|W_{n0}|^2+1\right)^4 \le 2^3\left(\mathbb{E}|W_{n0}|^8+1\right) < \infty \ . \tag{39}$$

Considering the second term at the right-hand side of (38), we write

$$\begin{aligned}
\mathbb{E}\left|W_{n0}\sum_{\ell=n+1}^{N}W_{\ell 0}b_{n\ell}\right|^4 &= \mathbb{E}\left|W_{n0}\right|^4\mathbb{E}\left|\sum_{\ell=n+1}^{N}W_{\ell 0}b_{n\ell}\right|^4 ,\\
&\overset{(a)}{\leq} C\left(\mathbb{E}\left(\sum_{\ell=n+1}^{N}(\mathbb{E}|W_{\ell 0}|^2)|b_{n\ell}|^2\right)^2+\sum_{\ell=n+1}^{N}(\mathbb{E}|W_{\ell 0}|^4)(\mathbb{E}|b_{n\ell}|^4)\right) ,\\
&\overset{(b)}{\leq} C\left(\mathbb{E}\left(\sum_{\ell=n+1}^{N}|b_{n\ell}|^2\right)^2+\sum_{\ell=n+1}^{N}\mathbb{E}|b_{n\ell}|^2\right) ,
\end{aligned}$$

where (a) follows from Lemma 6 (Burkholder's inequality), the filtration being $\mathcal{F}_{N,K},\dots,\mathcal{F}_{n+1,K}$ and (b) follows from the bound $|b_{n\ell}|^4\leq|b_{n\ell}|^2\max|b_{n\ell}|^2\leq|b_{n\ell}|^2(\sigma^2_{\max}\rho^{-1})^2$ (cf. Lemma 2–(1)). Now, notice that

$$\sum_{\ell=n+1}^{N}|b_{n\ell}|^2<\sum_{\ell=1}^{N}|b_{n\ell}|^2=\left[\mathbf{D}_0^{1/2}\mathbf{Q}\mathbf{D}_0\mathbf{Q}\mathbf{D}_0^{1/2}\right]_{nn}\leq\|\mathbf{D}_0^{1/2}\mathbf{Q}\mathbf{D}_0\mathbf{Q}\mathbf{D}_0^{1/2}\|\leq\frac{\sigma^4_{\max}}{\rho^2} .$$

This yields $\mathbb{E}|W_{n0}\sum_{\ell=n+1}^{N}W_{\ell\,0}b_{n\ell}|^4\leq C$. Gathering this result with (39), getting back to (38), taking the expectation and summing up finally yields:

$$\sum_{n=1}^{N}\mathbb{E}|Z_{n,K}|^4\ \leq\ \frac{C}{K}\ \xrightarrow[K\to\infty]{}\ 0$$

which establishes Lyapunov's condition (31) with $\alpha=2$.

Step 2: Proof of (32)*:* Eq. (37) yields:

$$\begin{aligned}
\mathbb{E}_{n+1}Z^2_{n,K} &= \frac{1}{K}\Bigg(\left(\mathbb{E}|W_{10}|^4-1\right)b^2_{nn}+\mathbb{E}_{n+1}\left(W^*_{n0}\sum_{\ell=n+1}^{N}W_{\ell 0}b_{n\ell}+W_{n0}\sum_{\ell=n+1}^{N}W^*_{\ell 0}b_{\ell n}\right)^2\\
&\quad +2b_{nn}\left(\mathbb{E}\,W^*_{10}|W_{10}|^2\right)\sum_{\ell=n+1}^{N}W_{\ell 0}b_{n\ell}+2b_{nn}\left(\mathbb{E}\,W_{10}|W_{10}|^2\right)\sum_{\ell=n+1}^{N}W^*_{\ell 0}b_{\ell n}\Bigg) .
\end{aligned}$$

Note that the second term of the right-hand side writes:

$$\mathbb{E}_{n+1}\left(W^*_{n0}\sum_{\ell=n+1}^{N}W_{\ell 0}b_{n\ell}+W_{n0}\sum_{\ell=n+1}^{N}W^*_{\ell 0}b_{\ell n}\right)^2=2\sum_{\ell_1,\ell_2=n+1}^{N}W_{\ell_1 0}W^*_{\ell_2 0}b_{n\ell_1}b_{\ell_2 n} .$$

Therefore, $V_K=\sum_{n=1}^{N}\mathbb{E}_{n+1}Z^2_{n,K}$ writes:

$$\begin{aligned}
V_K=\frac{\left(\mathbb{E}|W_{10}|^4-1\right)}{K}\sum_{n=1}^{N}b^2_{nn}+\frac{2}{K}\sum_{n=1}^{N}\sum_{\ell_1,\ell_2=n+1}^{N}W_{\ell_1 0}W^*_{\ell_2 0}b_{n\ell_1}b_{\ell_2 n}&\\
+\frac{2}{K}\Re\left(\left(\mathbb{E}\,W^*_{10}|W_{10}|^2\right)\sum_{n=1}^{N}b_{nn}\sum_{\ell=n+1}^{N}W_{\ell 0}b_{n\ell}\right) &,
\end{aligned}$$

where $\Re$ denotes the real part of a complex number. We introduce the following notations:

$$\mathbf{R} = (r_{ij})_{i,j=1}^N \stackrel{\triangle}{=} (b_{ij}\mathbf{1}_{i>j})_{i,j=1}^N \qquad \text{and} \qquad \Gamma_K = \frac{1}{K}\sum_{n=1}^N b_{nn} \sum_{\ell=n+1}^N W_{\ell 0} b_{n\ell} \ .$$

Note in particular that $\mathbf{R}$ is the strictly lower triangular matrix extracted from $\mathbf{D}_0^{1/2}\mathbf{Q}\mathbf{D}_0^{1/2}$. We can now rewrite V_K as:

$$V_K = \frac{\left(\mathbb{E}|W_{10}|^4 - 1\right)}{K}\mathrm{tr}\left(\mathbf{D}_0^2(\mathrm{diag}(\mathbf{Q}))^2\right) + \frac{2}{K}\mathbf{w}_0^*\mathbf{R}\mathbf{R}^*\mathbf{w}_0 + 2\Re\left(\Gamma_K \mathbb{E}W_{10}^*|W_{10}|^2\right) \ . \tag{40}$$

We now prove that the third term of the right-hand side vanishes, and find an asymptotic equivalent for the second one. Using Lemma 2, we have:

$$\begin{aligned}
\mathbb{E}_{N+1}|\Gamma_K|^2 &= \frac{1}{K^2}\sum_{n,m=1}^N b_{nn}b_{mm}\sum_{\ell=1}^N b_{n\ell}b_{m\ell}^*\mathbf{1}_{\ell>n}\mathbf{1}_{\ell>m} = \frac{1}{K^2}\mathrm{tr}\left(\mathrm{diag}(\mathbf{B})\mathbf{R}^*\mathbf{R}\mathrm{diag}(\mathbf{B})\right)\\
&= \frac{1}{K^2}\mathrm{tr}\left(\mathbf{D}_0^{1/2}\mathrm{diag}(\mathbf{Q})\mathbf{D}_0^{1/2}\mathbf{R}^*\mathbf{R}\mathbf{D}_0^{1/2}\mathrm{diag}(\mathbf{Q})\mathbf{D}_0^{1/2}\right)\\
&\le \frac{1}{K^2}\|\mathbf{D}_0\|^2\|\mathbf{Q}\|^2\mathrm{tr}(\mathbf{R}^*\mathbf{R}) \quad \le \quad \frac{1}{K^2}\|\mathbf{D}_0\|^2\|\mathbf{Q}\|^2\mathrm{tr}(\mathbf{B}^2) \quad \le \quad \frac{1}{K^2}\|\mathbf{D}_0\|^4\|\mathbf{Q}\|^2\mathrm{tr}(\mathbf{Q}^2)\\
&\le \frac{1}{K}\|\mathbf{D}_0\|^2\|\mathbf{Q}\|^4 \quad \le \quad \frac{1}{K}\frac{\sigma_{\max}^4}{\rho^4} \quad \xrightarrow[K\to\infty]{} \quad 0 \ .
\end{aligned}$$

In particular, $\mathbb{E}|\Gamma_K|^2 \to 0$ and

$$\Re\left(\left(\mathbb{E}W_{10}^*|W_{10}|^2\right)\Gamma_K\right) \xrightarrow[K\to\infty]{} 0 \quad \text{in probability} \ . \tag{41}$$

Consider now the second term of the right-hand side of Eq. (40). We prove that:

$$\frac{1}{K}\mathbf{w}_0^*\mathbf{R}\mathbf{R}^*\mathbf{w}_0 - \frac{1}{K}\mathrm{tr}(\mathbf{R}\mathbf{R}^*) \xrightarrow[K\to\infty]{} 0 \quad \text{in probability.} \tag{42}$$

By Lemma 1 (Ineq. (12)), we have

$$\mathbb{E}\left(\frac{1}{K}\mathbf{w}_0^*\mathbf{R}\mathbf{R}^*\mathbf{w}_0 - \frac{1}{K}\mathrm{tr}(\mathbf{R}\mathbf{R}^*)\right)^2 \le \frac{C}{K^2}(\mathbb{E}|W_{10}|^4)\mathrm{tr}(\mathbf{R}\mathbf{R}^*\mathbf{R}\mathbf{R}^*) \ .$$

Notice that $\mathrm{tr}(\mathbf{R}\mathbf{R}^*\mathbf{R}\mathbf{R}^*) = \|\mathbf{R}\|_4^4$ where $\|\mathbf{R}\|_4$ is the Schatten ℓ_4-norm of $\mathbf{R}$. Using Lemma 3, we have:

$$\|\mathbf{R}\|_4^4 \le C\|\mathbf{D}_0^{1/2}\mathbf{Q}\mathbf{D}_0^{1/2}\|_4^4 \le NC\|\mathbf{D}_0^{1/2}\mathbf{Q}\mathbf{D}_0^{1/2}\|^4 \le N\frac{C\sigma_{\max}^8}{\rho^4} \ .$$

Therefore,

$$\mathbb{E}\left(\frac{1}{K}\mathbf{w}_0^*\mathbf{R}\mathbf{R}^*\mathbf{w}_0 - \frac{1}{K}\mathrm{tr}(\mathbf{R}\mathbf{R}^*)\right)^2 \le C\frac{N}{K^2} \xrightarrow[K\to\infty]{} 0$$

which implies (42). Now, due to the fact that $\mathbf{B} = \mathbf{B}^*$, we have

$$\begin{aligned}
\frac{2}{K}\mathrm{tr}\,\mathbf{R}\mathbf{R}^* &= \frac{2}{K}\sum_{n=1}^{N}\sum_{\ell=n+1}^{N}|b_{n\ell}|^2 \\
&= \frac{1}{K}\sum_{n,\ell=1}^{N}|b_{n\ell}|^2 - \frac{1}{K}\sum_{n=1}^{N}|b_{nn}|^2 \\
&= \frac{1}{K}\mathrm{tr}\,\mathbf{D}_0\mathbf{Q}\mathbf{D}_0\mathbf{Q} - \frac{1}{K}\mathrm{tr}\,\mathbf{D}_0^2(\mathrm{diag}(\mathbf{Q}))^2 \qquad (43)
\end{aligned}$$

Gathering (40–43), we obtain (32). Step 2 is proved.

Step 3: Proof of (33) *and* (34)*:* We begin with some identities. Write $\mathbf{Q}(z) = [q_{ij}(z)]_{i,j=1}^N$ and $\widetilde{\mathbf{Q}}(z) = [\tilde{q}_{ij}(z)]_{i,j=1}^K$. Denote by $\mathbf{y}_k$ the column number k of $\mathbf{Y}$ and by $\boldsymbol{\xi}_n$ the row number n of $\mathbf{Y}$. Denote by $\mathbf{Y}^k$ the matrix that remains after deleting column k from $\mathbf{Y}$ and by $\mathbf{Y}_n$ the matrix that remains after deleting row n from $\mathbf{Y}$. Finally, write $\mathbf{Q}_k(z) = (\mathbf{Y}^k\mathbf{Y}^{k^*} - z\mathbf{I})^{-1}$ and $\widetilde{\mathbf{Q}}_n(z) = (\mathbf{Y}_n^*\mathbf{Y}_n - z\mathbf{I})^{-1}$. The following formulas can be established easily (see for instance [28, §0.7.3. and §0.7.4]):

$$q_{nn}(-\rho) = \frac{1}{\rho(1+\boldsymbol{\xi}_n\widetilde{\mathbf{Q}}_n(-\rho)\boldsymbol{\xi}_n^*)}, \quad \tilde{q}_{kk}(-\rho) = \frac{1}{\rho(1+\mathbf{y}_k^*\mathbf{Q}_k(-\rho)\mathbf{y}_k)}, \qquad (44)$$

$$\mathbf{Q} = \mathbf{Q}_k - \frac{\mathbf{Q}_k\mathbf{y}_k\mathbf{y}_k^*\mathbf{Q}_k}{1+\mathbf{y}_k^*\mathbf{Q}_k\mathbf{y}_k} \qquad (45)$$

Lemma 7: The following hold true:

1) (*Rank one perturbation inequality*) The resolvent $\mathbf{Q}_k(-\rho)$ satisfies $|\mathrm{tr}\mathbf{A}(\mathbf{Q}-\mathbf{Q}_k)| \le \|\mathbf{A}\|/\rho$ for any $N\times N$ matrix $\mathbf{A}$.
2) Let Assumptions **A1**–**A3** hold. Then,

$$\max_{1\le n\le N}\mathbb{E}(q_{nn}(-\rho) - t_n(-\rho))^2 \le \frac{C}{K}\ . \qquad (46)$$

The same conclusion holds true if q_{nn} and t_n are replaced with $\tilde{q}_{kk}$ and $\tilde{t}_k$ respectively.

We are now in position to prove (33). First, notice that:

$$\begin{aligned}
\mathbb{E}\left|q_{nn}^2 - t_n^2\right| &= \mathbb{E}\left|q_{nn}-t_n\right|(q_{nn}+t_n) \\
&\le \sqrt{\mathbb{E}(q_{nn}-t_n)^2}\sqrt{\mathbb{E}(q_{nn}+t_n)^2} \quad \le \quad \frac{2}{\rho}\sqrt{\mathbb{E}(q_{nn}-t_n)^2}\ . \qquad (47)
\end{aligned}$$

Now,

$$\begin{aligned}
\frac{1}{K}\mathbb{E}\left|\mathrm{tr}\,\mathbf{D}_0^2(\mathrm{diag}(\mathbf{Q})^2 - \mathbf{T}^2)\right| &\le \frac{1}{K}\sum_{n=1}^{N}\sigma_{0,n}^4\mathbb{E}\left|q_{nn}^2 - t_n^2\right| \quad \le \quad \frac{\sigma_{\max}^4 N}{K}\max_{1\le n\le N}\mathbb{E}\left|q_{nn}^2 - t_n^2\right| \\
&\le \frac{2\sigma_{\max}^4 N}{\rho K}\sqrt{\max_{1\le n\le N}\mathbb{E}(q_{nn}-t_n)^2} \quad \xrightarrow[K\to\infty]{} \quad 0\ ,
\end{aligned}$$

where the last inequality follows from (47) together with Lemma 7–(2). Convergence (33) is established.

We now establish the system of equations (34). Our starting point is the identity

$$\mathbf{Q} = \mathbf{T} + \mathbf{T}(\mathbf{T}^{-1} - \mathbf{Q}^{-1})\mathbf{Q} = \mathbf{T} + \frac{\rho}{K}\mathbf{T}\,\mathrm{diag}(\mathrm{tr}\widetilde{\mathbf{D}}_1\widetilde{\mathbf{T}}, \dots, \mathrm{tr}\widetilde{\mathbf{D}}_N\widetilde{\mathbf{T}})\mathbf{Q} - \mathbf{T}\mathbf{Y}\mathbf{Y}^*\mathbf{Q} \ .$$

Using this identity, we develop $U_\ell = \frac{1}{K}\mathrm{tr}\mathbf{D}_0\mathbf{Q}\mathbf{D}_\ell\mathbf{Q}$ as

$$\begin{aligned} U_\ell &= \frac{1}{K}\mathrm{tr}\mathbf{D}_0\mathbf{Q}\mathbf{D}_\ell\mathbf{T} + \frac{\rho}{K^2}\mathrm{tr}\mathbf{D}_0\mathbf{Q}\mathbf{D}_\ell\mathbf{T}\mathrm{diag}(\mathrm{tr}\widetilde{\mathbf{D}}_1\widetilde{\mathbf{T}}, \dots, \mathrm{tr}\widetilde{\mathbf{D}}_N\widetilde{\mathbf{T}})\mathbf{Q} - \frac{1}{K}\mathrm{tr}\mathbf{D}_0\mathbf{Q}\mathbf{D}_\ell\mathbf{T}\mathbf{Y}\mathbf{Y}^*\mathbf{Q} \\ &\stackrel{\triangle}{=} X_1 + X_2 - X_3 \ . \end{aligned} \tag{48}$$

Lemma 4–(2) with $\mathbf{S} = \mathbf{D}_0\mathbf{D}_\ell\mathbf{T}$ yields:

$$X_1 = \frac{1}{K}\mathrm{tr}\mathbf{D}_0\mathbf{D}_\ell\mathbf{T}^2 + \epsilon_1 \tag{49}$$

where $\mathbb{E}|\epsilon_1| \le \sqrt{\mathbb{E}\epsilon_1^2} \le C/K$. Consider now the term $X_3 = \frac{1}{K}\sum_{k=1}^K \mathrm{tr}\mathbf{D}_0\mathbf{Q}\mathbf{D}_\ell\mathbf{T}\mathbf{y}_k\mathbf{y}_k^*\mathbf{Q}$. Using (44) and (45), we have

$$\mathbf{y}_k^*\mathbf{Q} = \left(1 - \frac{\mathbf{y}_k^*\mathbf{Q}\mathbf{y}_k}{1 + \mathbf{y}_k^*\mathbf{Q}\mathbf{y}_k}\right)\mathbf{y}_k^*\mathbf{Q}_k = \rho\,\tilde{q}_{kk}\,\mathbf{y}_k^*\mathbf{Q}_k \ .$$

Hence

$$\begin{aligned} X_3 &= \frac{\rho}{K}\sum_{k=1}^K \tilde{q}_{kk}\mathbf{y}_k^*\mathbf{Q}_k\mathbf{D}_0\mathbf{Q}\mathbf{D}_\ell\mathbf{T}\mathbf{y}_k \\ &= \frac{\rho}{K}\sum_{k=1}^K \tilde{t}_k\mathbf{y}_k^*\mathbf{Q}_k\mathbf{D}_0\mathbf{Q}\mathbf{D}_\ell\mathbf{T}\mathbf{y}_k + \frac{\rho}{K}\sum_{k=1}^K(\tilde{q}_{kk} - \tilde{t}_k)\mathbf{y}_k^*\mathbf{Q}_k\mathbf{D}_0\mathbf{Q}\mathbf{D}_\ell\mathbf{T}\mathbf{y}_k \\ &\stackrel{\triangle}{=} X_3' + \epsilon_2 \ . \end{aligned} \tag{50}$$

By Cauchy-Schwartz inequality,

$$\mathbb{E}|\epsilon_2| \le \frac{\rho}{K}\sum_{k=1}^K \sqrt{\mathbb{E}(\tilde{q}_{kk} - \tilde{t}_k)^2}\sqrt{\mathbb{E}(\mathbf{y}_k^*\mathbf{Q}_k\mathbf{D}_0\mathbf{Q}\mathbf{D}_\ell\mathbf{T}\mathbf{y}_k)^2} \ .$$

We have $\mathbb{E}(\mathbf{y}_k^*\mathbf{Q}_k\mathbf{D}_0\mathbf{Q}\mathbf{D}_\ell\mathbf{T}\mathbf{y}_k)^2 \le \sigma_{\max}^8\rho^{-6}\mathbb{E}\|\mathbf{y}_k\|^4 \le C$. Using in addition Lemma 7–(2), we obtain

$$\mathbb{E}|\epsilon_2| \le \frac{C}{\sqrt{K}} \ .$$

Consider X_3'. From (44) and (45), we have $\mathbf{Q} = \mathbf{Q}_k - \rho\tilde{q}_{kk}\mathbf{Q}_k\mathbf{y}_k\mathbf{y}_k^*\mathbf{Q}_k$. Hence, we can develop X_3' as

$$\begin{aligned} X_3' &= \frac{\rho}{K}\sum_{k=1}^K \tilde{t}_k\mathbf{y}_k^*\mathbf{Q}_k\mathbf{D}_0\mathbf{Q}_k\mathbf{D}_\ell\mathbf{T}\mathbf{y}_k - \frac{\rho^2}{K}\sum_{k=1}^K \tilde{t}_k\tilde{q}_{kk}\mathbf{y}_k^*\mathbf{Q}_k\mathbf{D}_0\mathbf{Q}_k\mathbf{y}_k\mathbf{y}_k^*\mathbf{Q}_k\mathbf{D}_\ell\mathbf{T}\mathbf{y}_k \\ &\stackrel{\triangle}{=} X_4 + X_5 \ . \end{aligned} \tag{51}$$

Consider X_4. Notice that $\mathbf{y}_k$ and $\mathbf{Q}_k$ are independent. Therefore, by Lemma 1, we obtain

$$\mathbf{y}_k^*\mathbf{Q}_k\mathbf{D}_0\mathbf{Q}_k\mathbf{D}_\ell\mathbf{T}\mathbf{y}_k = \frac{1}{K}\mathrm{tr}\mathbf{D}_k\mathbf{Q}_k\mathbf{D}_0\mathbf{Q}_k\mathbf{D}_\ell\mathbf{T} + \epsilon_3 = \frac{1}{K}\mathrm{tr}\mathbf{D}_k\mathbf{Q}\mathbf{D}_0\mathbf{Q}\mathbf{D}_\ell\mathbf{T} + \epsilon_3 + \epsilon_4$$

where $\mathbb{E}\epsilon_3^2 < CK^{-1}$ by Ineq. (13). Applying twice Lemma 7–(1) to $\epsilon_4 = \frac{1}{K}(\mathrm{tr}\mathbf{D}_k\mathbf{Q}_k\mathbf{D}_0\mathbf{Q}_k\mathbf{D}_\ell\mathbf{T} - \mathrm{tr}\mathbf{D}_k\mathbf{Q}\mathbf{D}_0\mathbf{Q}\mathbf{D}_\ell\mathbf{T})$ yields $|\epsilon_4| < CK^{-1}$. Note in addition that $\sum \tilde{t}_k\mathbf{D}_k = \mathrm{diag}(\mathrm{tr}\widetilde{\mathbf{D}}_1\widetilde{\mathbf{T}}, \ldots, \mathrm{tr}\widetilde{\mathbf{D}}_N\widetilde{\mathbf{T}})$. Thus, we obtain

$$\begin{aligned} X_4 &= \frac{\rho}{K^2}\mathrm{tr}\left(\sum_{k=1}^K \tilde{t}_k\mathbf{D}_k\right)\mathbf{Q}\mathbf{D}_0\mathbf{Q}\mathbf{D}_\ell\mathbf{T} + \epsilon_5 \\ &= X_2 + \epsilon_5 \ , \end{aligned} \tag{52}$$

where $\epsilon_5 = \epsilon_3 + \epsilon_4$, which yields $\mathbb{E}|\epsilon_5| \le CK^{-\frac{1}{2}}$.

We now turn to X_5. First introduce the following random variable:

$$\epsilon_6 = \tilde{t}_k\tilde{q}_{kk}\mathbf{y}_k^*\mathbf{Q}_k\mathbf{D}_0\mathbf{Q}_k\mathbf{y}_k\mathbf{y}_k^*\mathbf{Q}_k\mathbf{D}_\ell\mathbf{T}\mathbf{y}_k - \tilde{t}_k\tilde{q}_{kk}\left(\frac{1}{K}\mathrm{tr}\mathbf{D}_k\mathbf{Q}_k\mathbf{D}_0\mathbf{Q}_k\right)\left(\frac{1}{K}\mathrm{tr}\mathbf{D}_k\mathbf{Q}_k\mathbf{D}_\ell\mathbf{T}\right)$$

Then

$$\begin{aligned} |\epsilon_6| \le \frac{1}{\rho^2}\mathbf{y}_k^*\mathbf{Q}_k\mathbf{D}_0\mathbf{Q}_k\mathbf{y}_k & \left|\mathbf{y}_k^*\mathbf{Q}_k\mathbf{D}_\ell\mathbf{T}\mathbf{y}_k - \frac{1}{K}\mathrm{tr}\mathbf{D}_k\mathbf{Q}_k\mathbf{D}_\ell\mathbf{T}\right| \\ & + \frac{1}{\rho^2}\left|\mathbf{y}_k^*\mathbf{Q}_k\mathbf{D}_0\mathbf{Q}_k\mathbf{y}_k^* - \frac{1}{K}\mathrm{tr}\mathbf{D}_k\mathbf{Q}_k\mathbf{D}_0\mathbf{Q}_k\right|\frac{1}{K}\mathrm{tr}\mathbf{D}_k\mathbf{Q}_k\mathbf{D}_\ell\mathbf{T} \end{aligned}$$

and one can prove that $\mathbb{E}|\epsilon_6| < CK^{-\frac{1}{2}}$ with help of Lemma 1, together with Cauchy-Schwarz inequality. In addition, we can prove with the help of Lemma 7 that:

$$\begin{aligned} \tilde{t}_k\tilde{q}_{kk}\left(\frac{1}{K}\mathrm{tr}\mathbf{D}_k\mathbf{Q}_k\mathbf{D}_0\mathbf{Q}_k\right)\left(\frac{1}{K}\mathrm{tr}\mathbf{D}_k\mathbf{Q}_k\mathbf{D}_\ell\mathbf{T}\right) &= \tilde{t}_k^2\left(\frac{1}{K}\mathrm{tr}\mathbf{D}_k\mathbf{Q}\mathbf{D}_0\mathbf{Q}\right)\left(\frac{1}{K}\mathrm{tr}\mathbf{D}_k\mathbf{Q}\mathbf{D}_\ell\mathbf{T}\right) + \epsilon_7 \\ &= \tilde{t}_k^2\left(\frac{1}{K}\mathrm{tr}\mathbf{D}_k\mathbf{Q}\mathbf{D}_0\mathbf{Q}\right)\left(\frac{1}{K}\mathrm{tr}\mathbf{D}_k\mathbf{D}_\ell\mathbf{T}^2\right) + \epsilon_7 + \epsilon_8 \end{aligned}$$

where ϵ_7 and ϵ_8 are random variables satisfying $\mathbb{E}|\epsilon_7| < CK^{-\frac{1}{2}}$ by Lemma 7, and $\max_{k,\ell}\mathbb{E}|\epsilon_8| \le \max_{k,\ell}\sqrt{\mathbb{E}|\epsilon_8|^2} \le CK^{-\frac{1}{2}}$ by Lemma 4–(2). Using the fact that $\rho^2\tilde{t}_k^2 = (1 + \frac{1}{K}\mathrm{tr}\mathbf{D}_k\mathbf{T})^{-2}$, we end up with

$$X_5 = -\frac{\rho^2}{K}\sum_{k=1}^K \tilde{t}_k^2\left(\frac{1}{K}\mathrm{tr}\mathbf{D}_k\mathbf{Q}\mathbf{D}_0\mathbf{Q}\right)\left(\frac{1}{K}\mathrm{tr}\mathbf{D}_k\mathbf{D}_\ell\mathbf{T}^2\right) + \epsilon_9 = -\sum_{k=1}^K c_{\ell k}U_k + \epsilon_9 \tag{53}$$

where $c_{\ell k}$ is given by (35), and where $\mathbb{E}|\epsilon_9| < CK^{-\frac{1}{2}}$.

Plugging Eq. (49)–(53) into (48), we end up with $U_\ell = \sum_{k=1}^K c_{\ell k}U_k + \frac{1}{K}\mathrm{tr}\mathbf{D}_0\mathbf{D}_\ell\mathbf{T}^2 + \epsilon$ with $\mathbb{E}|\epsilon| < CK^{-\frac{1}{2}}$. Step 3 is established.

Step 4 : Proof of (36)*:* We rely on results of Section V-B, in particular on Lemma 5.

Define the following $(K+1)\times 1$ vectors:

$$\mathbf{u} = [U_k]_{k=0}^K, \ \mathbf{d} = \left[\frac{1}{K}\mathrm{tr}\mathbf{D}_0\mathbf{D}_k\mathbf{T}^2\right]_{k=0}^K, \ \boldsymbol{\epsilon} = [\epsilon_k]_{k=0}^K \ ,$$

where the U_k's and ϵ_k's are defined in (34). Recall the definition of the $c_{\ell k}$'s for $0 \le \ell \le K$ and $1 \le k \le K$, define $c_{\ell\,0} = 0$ for $0 \le \ell \le K$ and consider the $(K+1) \times (K+1)$ matrix $\mathbf{C} = [c_{\ell k}]_{\ell,k=0}^{K}$.

With these notations, System (34) writes

$$(\mathbf{I}_{K+1} - \mathbf{C})\,\mathbf{u} = \mathbf{d} + \boldsymbol{\epsilon}\ . \tag{54}$$

Let $\boldsymbol{\alpha} = \frac{1}{K}\mathrm{tr}\,\mathbf{D}_0^2\mathbf{T}^2$ and $\boldsymbol{\beta} = (1 + \frac{1}{K}\mathrm{tr}\,\mathbf{D}_0\mathbf{T})^2$. We have in particular

$$\mathbf{d} = \begin{bmatrix} \boldsymbol{\alpha} \\ \mathbf{g} \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 0 & \frac{1}{K}\mathbf{g}^T\boldsymbol{\Delta}^{-1} \\ \mathbf{0} & \mathbf{A}^T \end{bmatrix}$$

(recall that $\mathbf{A}$, $\boldsymbol{\Delta}$ and $\mathbf{g}$ are defined in the statement of Theorem 3).

Consider a square matrix $\mathbf{X}$ which first column is equal to $[1, 0, \ldots, 0]^{\mathrm{T}}$, and partition $\mathbf{X}$ as $\mathbf{X} = \begin{bmatrix} 1 & \mathbf{x}_{01}^T \\ \mathbf{0} & \mathbf{X}_{11} \end{bmatrix}$. Recall that the inverse of $\mathbf{X}$ exists if and only if $\mathbf{X}_{11}^{-1}$ exists, and in this case the first row $[\mathbf{X}^{-1}]_0$ of $\mathbf{X}^{-1}$ is given by

$$\left[\mathbf{X}^{-1}\right]_0 = \left[1 \quad -\mathbf{x}_{01}^T\mathbf{X}_{11}^{-1}\right]$$

(see for instance [28]). We now apply these results to the system (54). Due to (54), U_0 can be expressed as

$$U_0 = [(\mathbf{I} - \mathbf{C})^{-1}]_0(\mathbf{d} + \boldsymbol{\epsilon})\ .$$

By Lemma 5–(1), $(\mathbf{I}_K - \mathbf{A}^T)^{-1}$ exists hence $(\mathbf{I} - \mathbf{C})^{-1}$ exists,

$$\left[(\mathbf{I}_{K+1} - \mathbf{C})^{-1}\right]_0 = \left[1 \quad \frac{1}{K}\mathbf{g}^T\boldsymbol{\Delta}^{-1}(\mathbf{I}_K - \mathbf{A}^T)^{-1}\right]\ ,$$

and

$$U_0 = \boldsymbol{\alpha} + \frac{1}{K}\mathbf{g}^T\boldsymbol{\Delta}^{-1}\left(\mathbf{I} - \mathbf{A}^T\right)^{-1}\mathbf{g} + \epsilon_0 + \frac{1}{K}\mathbf{g}^T\boldsymbol{\Delta}^{-1}\left(\mathbf{I} - \mathbf{A}^T\right)^{-1}\boldsymbol{\epsilon}'$$

with $\boldsymbol{\epsilon}' = [\epsilon_1, \ldots, \epsilon_K]^T$. Gathering the estimates of Section V-B together with the fact that $\|\mathbb{E}\boldsymbol{\epsilon}\|_\infty \le CK^{-\frac{1}{2}}$, we get (36). Step 4 is established, so is Theorem 3.

APPENDIX

A. *Proof of Lemma 4*

Let us establish (23). The lower bound immediately follows from the representation

$$t_n = \frac{1}{\rho + \frac{1}{K}\sum_{k=1}^{K}\frac{\sigma_{nk}^2}{1+\frac{1}{K}\sum_{\ell=1}^{N}\sigma_{\ell k}^2 t_\ell}} \overset{(a)}{\ge} \frac{1}{\rho + \sigma_{\max}^2}$$

where (a) follows from **A2** and $t_\ell(-\rho) \geq 0$. The upper bound requires an extra argument: As proved in [26, Theorem 2.4], the t_n's are Stieltjes transforms of probability measures supported by $\mathbb{R}_+$, i.e. there exists a probability measure μ_n over $\mathbb{R}_+$ such that $t_n(z) = \int \frac{\mu_n(dt)}{t-z}$. Thus

$$t_n(-\rho) = \int_0^\infty \frac{\mu_n(dt)}{t+\rho} \leq \frac{1}{\rho} \ ,$$

and (23) is proved.

We now briefly justify (24). We have $\mathbb{E}\,|\mathrm{tr}\,\mathbf{S}(\mathbf{Q}-\mathbf{T})|^2 = \mathbb{E}\,|\mathrm{tr}\mathbf{S}(\mathbf{Q}-\mathbb{E}\mathbf{Q})|^2 + |\mathrm{tr}\mathbf{S}(\mathbb{E}\mathbf{Q}-\mathbf{T})|^2$. In [21, Lemma 6.3] it is stated that $\sup_K \mathbb{E}\,|\mathrm{tr}\mathbf{S}(\mathbf{Q}-\mathbb{E}\mathbf{Q})|^2 < \infty$. Furthermore, in the proof of [21, Theorem 3.3] it is shown that $\sup_K K\|\mathbb{E}\mathbf{Q}-\mathbf{T}\| < \infty$, hence $|\mathrm{tr}\mathbf{S}(\mathbb{E}\mathbf{Q}-\mathbf{T})| \leq K\|\mathbf{S}(\mathbb{E}\mathbf{Q}-\mathbf{T})\| \leq K\|\mathbb{E}\mathbf{Q}-\mathbf{T}\|\|\mathbf{S}\| < \infty$ by Lemma 2–(2). The result follows.

B. *Proof of Corollary 1*

Recall that in the separable case, $\mathbf{D}_k = \tilde{d}_k\mathbf{D}$ and $\widetilde{\mathbf{D}}_n = d_n\widetilde{\mathbf{D}}$. Let $\tilde{\mathbf{d}}$ be the $K\times 1$ vector $\tilde{\mathbf{d}} = [\tilde{d}_k]_{k=1}^K$. In the separable case, Eq. (20) is written

$$\frac{\Theta^2}{\tilde{d}_0^2} = \frac{1}{K\tilde{d}_0^2}\mathbf{g}^T(\mathbf{I}-\mathbf{A})^{-1}\boldsymbol{\Delta}^{-1}\mathbf{g} + \gamma(\mathbb{E}|W_{10}|^4 - 1) \ , \tag{55}$$

where γ is defined in statement of the corollary. Here, vector $\mathbf{g}$ and matrix $\mathbf{A}$ are given by

$$\mathbf{g} = \gamma\tilde{d}_0\tilde{\mathbf{d}} \quad \text{and} \quad \mathbf{A} = \left[\frac{1}{K}\frac{\frac{1}{K}\mathrm{tr}\mathbf{D}_\ell\mathbf{D}_m\mathbf{T}^2}{\left(1+\frac{1}{K}\mathrm{tr}\mathbf{D}_\ell\mathbf{T}\right)^2}\right]_{\ell,m=1}^K = \frac{\gamma}{K}\boldsymbol{\Delta}^{-1}\tilde{\mathbf{d}}\tilde{\mathbf{d}}^T \ .$$

By the matrix inversion lemma [28], we have

$$\begin{aligned} \frac{1}{K\tilde{d}_0^2}\mathbf{g}^T(\mathbf{I}-\mathbf{A})^{-1}\boldsymbol{\Delta}^{-1}\mathbf{g} &= \frac{\gamma^2}{K}\tilde{\mathbf{d}}^T\left(\boldsymbol{\Delta} - \frac{\gamma}{K}\tilde{\mathbf{d}}\tilde{\mathbf{d}}^T\right)^{-1}\tilde{\mathbf{d}} \\ &= \frac{\gamma^2}{K}\tilde{\mathbf{d}}^T\left(\boldsymbol{\Delta}^{-1} + \frac{\gamma}{K}\frac{1}{1-\frac{\gamma}{K}\tilde{\mathbf{d}}^T\boldsymbol{\Delta}^{-1}\tilde{\mathbf{d}}}\boldsymbol{\Delta}^{-1}\tilde{\mathbf{d}}\tilde{\mathbf{d}}^T\boldsymbol{\Delta}^{-1}\right)\tilde{\mathbf{d}} \ . \end{aligned}$$

Noticing that

$$\frac{1}{K}\tilde{\mathbf{d}}^T\boldsymbol{\Delta}^{-1}\tilde{\mathbf{d}} = \frac{1}{K}\sum_{k=1}^K \frac{\tilde{d}_k^2}{\left(1+\frac{1}{K}\mathrm{tr}\mathbf{D}_k\mathbf{T}\right)^2} = \frac{\rho^2}{K}\sum_{k=1}^K \tilde{d}_k^2\tilde{t}_k^2 = \rho^2\tilde{\gamma} \ ,$$

we obtain

$$\frac{1}{K\tilde{d}_0^2}\mathbf{g}^T(\mathbf{I}-\mathbf{A})^{-1}\boldsymbol{\Delta}^{-1}\mathbf{g} = \gamma\frac{\rho^2\gamma\tilde{\gamma}}{1-\rho^2\gamma\tilde{\gamma}} \ .$$

Plugging this equation into (55), we obtain (22).

C. *Proof of Lemma 7*

The proof of Part 1 can be found in [21, Proof of Lemma 6.3] (see also [14, Lemma 2.6]). Let us prove Part 2. We have from Equations (11) and (44)

$$\begin{aligned} |q_{nn}(-\rho) - t_n(-\rho)| &= \frac{1}{\rho(1+\frac{1}{K}\mathrm{tr}\widetilde{\mathbf{D}}_n\widetilde{\mathbf{T}})(1+\boldsymbol{\xi}_n\widetilde{\mathbf{Q}}_n\boldsymbol{\xi}_n^*)} \left| \boldsymbol{\xi}_n\widetilde{\mathbf{Q}}_n\boldsymbol{\xi}_n^* - \frac{1}{K}\mathrm{tr}\widetilde{\mathbf{D}}_n\widetilde{\mathbf{T}} \right| \\ &\leq \frac{1}{\rho}\left| \boldsymbol{\xi}_n\widetilde{\mathbf{Q}}_n\boldsymbol{\xi}_n^* - \frac{1}{K}\mathrm{tr}\widetilde{\mathbf{D}}_n\widetilde{\mathbf{T}} \right| . \end{aligned}$$

Hence,

$$\mathbb{E}(q_{nn} - t_n)^2 \quad \leq \quad \frac{2}{\rho}\mathbb{E}\left(\boldsymbol{\xi}_n\widetilde{\mathbf{Q}}_n\boldsymbol{\xi}_n^* - \frac{1}{K}\mathrm{tr}\widetilde{\mathbf{D}}_n\widetilde{\mathbf{Q}} \right)^2 + \frac{2}{\rho K^2}\mathbb{E}\left(\mathrm{tr}\widetilde{\mathbf{D}}_n(\widetilde{\mathbf{Q}} - \widetilde{\mathbf{T}})\right)^2 \quad \leq \quad \frac{C}{K}$$

by Lemma 1 and Lemma 4–(2), which proves (46).

REFERENCES

[1] D.N.C Tse and S. Hanly, "Linear multi-user receiver: Effective interference, effective bandwidth and user capacity," *IEEE Trans. on Information Theory*, vol. 45, no. 2, pp. 641–657, Mar. 1999.

[2] S. Verdú and Sh. Shamai, "Spectral efficiency of CDMA with random spreading," *IEEE Trans. on Information Theory*, vol. 45, no. 2, pp. 622–640, Mar. 1999.

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

[4] E. Biglieri, G. Caire, and G. Taricco, "CDMA system design through asymptotic analysis," *IEEE Trans. on Communications*, vol. 48, no. 11, pp. 1882–1896, Nov. 2000.

[5] W. Phoel and M.L. Honig, "Performance of coded DS-CDMA with pilot-assisted channel estimation and linear interference suppression," *IEEE Trans. on Communications*, vol. 50, no. 5, pp. 822–832, May 2002.

[6] J.-M. Chaufray, W. Hachem, and Ph. Loubaton, "Asymptotic Analysis of Optimum and Sub-Optimum CDMA Downlink MMSE Receivers," *IEEE Trans. on Information Theory*, vol. 50, no. 11, pp. 2620–2638, Nov. 2004.

[7] L. Li, A.M. Tulino, and S. Verdú, "Design of Reduced-Rank MMSE Multiuser Detectors Using Random Matrix Methods," *IEEE Trans. on Information Theory*, vol. 50, no. 6, pp. 986–1008, June 2004.

[8] A.M. Tulino, L. Li, and S. Verdú, "Spectral Efficiency of Multicarrier CDMA," *IEEE Trans. on Information Theory*, vol. 51, no. 2, pp. 479–505, Feb. 2005.

[9] M.J.M. Peacock, I.B. Collings, and M.L. Honig, "Asymptotic spectral efficiency of multiuser multisignature CDMA in frequency-selective channels," *IEEE Trans. on Information Theory*, vol. 52, no. 3, pp. 1113–1129, Mar. 2006.

[10] J.H. Kotecha and A.M. Sayeed, "Transmit signal design for optimal estimation of correlated MIMO channels," *IEEE Trans. on Signal Processing*, vol. 52, no. 2, pp. 546–557, Feb. 2004.

[11] A.M. Sayeed, "Deconstructing multiantenna fading channels," *IEEE Trans. on Signal Processing*, vol. 50, no. 10, pp. 2563–2579, Oct. 2002.

[12] Shiu D.-S., G.J. Foschini, M.J. Gans, and J.M. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Trans. on Communications*, vol. 48, no. 3, pp. 502–513, Mar. 2000.

[13] V. L. Girko, *Theory of Random Determinants*, vol. 45 of *Mathematics and its Applications (Soviet Series)*, Kluwer Academic Publishers Group, Dordrecht, 1990.

[14] 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.

[15] D. Shlyakhtenko, "Random Gaussian band matrices and freeness with amalgamation," *Internat. Math. Res. Notices*, , no. 20, pp. 1013–1025, 1996.

[16] D.N.C. Tse and O. Zeitouni, "Linear multiuser receivers in random environments," *IEEE Trans. on Information Theory*, vol. 46, no. 1, pp. 171–188, Jan. 2000.

[17] J.W. Silverstein, "Weak convergence of random functions defined by the eigenvectors of sample covariance matrices," *Ann. Probab.*, vol. 18, no. 3, pp. 1174–1194, 1990.

[18] G.-M. Pan, M.-H Guo, and W. Zhou, "Asymptotic distributions of the Signal-to-Interference Ratios of LMMSE detection in multiuser communications," *Ann. Appl. Probab.*, vol. 17, no. 1, pp. 181–206, 2007.

[19] F. Götze and A. Tikhomirov, "Asymptotic distributions of quadratic forms and applications," *J. Theoret. Probab.*, vol. 15, pp. 424–475, 2002.

[20] W. Hachem, O. Khorunzhiy, P. Loubaton, J. Najim, and L. Pastur, "A new approach for capacity analysis of large dimensional multi-antenna channels," Accepted for publication in IEEE trans. Inf. Theory, available at http://arxiv.org/abs/cs.IT/0612076, 2006.

[21] W. Hachem, Ph. Loubaton, and J. Najim, "A CLT for information-theoretic statistics of Gram random matrices with a given variance profile," accepted for publication in *Ann. Appl. Probab.*, 2007. `arXiv:0706.0166`.

[22] A.L. Moustakas, S.H. Simon, and A.M. Sengupta, "MIMO capacity through correlated channels in the presence of correlated interferers and noise: A (not so) large N analysis," *IEEE Trans. on Information Theory*, vol. 49, no. 10, pp. 2545–2561, Oct. 2003.

[23] A.L. Moustakas and S.H. Simon, "On the outage capacity of correlated multiple-path MIMO channels," *IEEE Trans. on Information Theory*, vol. 53, no. 11, pp. 3887–3903, Nov. 2007.

[24] P. Billingsley, *Probability and Measure*, John Wiley, 3rd edition, 1995.

[25] R.J. Bhansali, L. Giraitis, and P.S. Kokoszka, "Convergence of quadratic forms with nonvanishing diagonal," *Stat. Probab. Letters*, vol. 77, pp. 726–734, 2007.

[26] W. Hachem, P. Loubaton, and J. Najim, "Deterministic equivalents for certain functionals of large random matrices," *Ann. Appl. Probab.*, vol. 17, no. 3, pp. 875–930, 2007.

[27] 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.

[28] R. Horn and C. Johnson, *Matrix Analysis*, Cambridge Univ. Press, 1994.

[29] N.K. Nikolski, *Operators, Functions and Systems: An Easy Reading. Vol. 2: Model Operators and Systems*, Mathematical Surveys and Monographs. AMS, 2002.