Iterative Deterministic Equivalents for the Performance Analysis of Communication Systems

Jakob Hoydis[⋆,§], Romain Couillet[⋆] and Mérouane Debbah[⋆]

Abstract

In this article, we introduce iterative deterministic equivalents as a novel technique for the performance analysis of communication systems whose channels are modeled by complex combinations of independent random matrices. This technique extends the deterministic equivalent approach for the study of functionals of large random matrices to a broader class of random matrix models which naturally arise as channel models in wireless communications. We present two specific applications: First, we consider a multi-hop amplify-and-forward (AF) MIMO relay channel with noise at each stage and derive deterministic approximations of the mutual information after the Kth hop. Second, we study a MIMO multiple access channel (MAC) where the channel between each transmitter and the receiver is represented by the double-scattering channel model. We provide deterministic approximations of the mutual information, the signal-to-interference-plus-noise ratio (SINR) and sum-rate with minimum-mean-square-error (MMSE) detection and derive the asymptotically optimal precoding matrices. In both scenarios, the approximations can be computed by simple and provably converging fixed-point algorithms and are shown to be almost surely tight in the limit when the number of antennas at each node grows infinitely large. Simulations suggest that the approximations are accurate for realistic system dimensions. The technique of iterative deterministic equivalents can be easily extended to other channel models of interest and is, therefore, also a new contribution to the field of random matrix theory.

Parts of this work have been presented at the Asilomar Conference on Signals, Systems, and Computers, CA, US, Nov. 2011.

[⋆]Alcatel Lucent Chair on Flexible Radio, Supélec, 3 rue Joliot Curie, 91192 Gif sur Yvette, France.

[§]Department of Telecommunications, Supélec, 3 rue Joliot Curie, 91192 Gif sur Yvette, France.

{jakob.hoydis,romain.couillet,merouane.debbah}@supelec.fr

I. Introduction

Since the pioneering work of Tse and Hanly [1] on the capacity of code division multiple access (CDMA) technologies assuming long spreading sequences, the theory of large dimensional random matrices (RMT) has drawn an increasing interest from researchers in wireless communications and related fields [2], [3]. RMT is in particular convenient for the study of multiple-input multiple-output (MIMO) channels [4], [5], (random) linear precoders [1], [6], [7], multi-user systems [8], [9], multi-cellular systems [10], [11], [12], etc. In the early contributions, it was systematically assumed that the dimensions of the system under study could grow infinitely large and that the system performance admitted a deterministic limit that RMT can provide [1], [4], [13]. It then became rapidly clear that, for most systems of practical interest, either the former assumption is not natural or the latter condition is not met. However, even for systems of large but finite size, the inherently random performance (e.g. instantaneous mutual information, signal-to-interference-plus-noise ratio (SINR)), can often be well approximated by deterministic quantities. Such quantities are called *deterministic equivalents*, and can be derived by various techniques, such as the Stieltjes transform method [5], [14], the Gaussian method [15], [16], or the replica method [17], [18].

Deterministic equivalents are convenient to study the performance of wireless communication systems when a single system parameter can be modeled by a random matrix, e.g. the fading channel or the spreading codes. In order to tackle the performance analysis of more complex systems, such as multi-hop communications, random beamforming over random fading channels, or double-scattering channels, it is necessary to extend the notion of deterministic equivalents. In the present article, which is inspired by the original idea of [7], where the performance of random isometric precoders over random fading channels is analyzed, we develop a systematic approach to generalize deterministic equivalents to *iterative deterministic equivalents*. To this end, we introduce a generic definition of deterministic equivalents of functionals of random matrices, which we extend, based on the Fubini theorem [19], to a definition of iterative deterministic equivalents.

As application examples, we then provide deterministic equivalents of the mutual information of the multi-hop amplify-and-forward (AF) MIMO relay channel [20], [21], [22] (Section III-A) and of the ergodic capacity as well as the sum-rate with minimum-mean-square-error (MMSE) detection of double-scattering multiple access channels (MACs) [23], [24] (Section III-B). An overview of related research to both topics is provided in the respective sections. Our analysis is based on the Stieltjes transform method, documented in detail in [3].

The remainder of this article is structured as follows. In Section II, we recall the fundamentals

of deterministic equivalents in RMT and develop the notion of iterative deterministic equivalents. In Section III, we study applications of iterative deterministic equivalents to the performance analysis of multi-hop relay channels and double-scattering MACs. The paper is concluded with Section IV. All proofs, related results, and some exemplary Matlab codes are provided in the appendices.

Notations: Boldface lower and upper case symbols represent vectors and matrices, respectively. $\mathbf{I}_N$ is the size-N identity matrix and $\mathrm{diag}(x_1, \ldots, x_N)$ is a diagonal matrix with elements x_i. The trace, transpose, and Hermitian transpose operators are denoted by $\mathrm{tr}\,(\cdot)$, $(\cdot)^{\mathsf{T}}$, and $(\cdot)^{\mathsf{H}}$, respectively. The spectral norm of a matrix $\mathbf{A}$ is denoted by $\|\mathbf{A}\|$, and, for two matrices $\mathbf{A}$ and $\mathbf{B}$, the notation $\mathbf{A} \succ \mathbf{B}$ means that $\mathbf{A} - \mathbf{B}$ is positive-definite. For a vector $\mathbf{x} = [x_1 \ldots x_N]^{\mathsf{T}}$, $\mathbf{x} \geq 0$ denotes $x_i \geq 0$ for all i. The notations $\Rightarrow$ and $\xrightarrow{\text{a.s.}}$ denote weak and almost sure convergence, respectively. We use $\mathcal{CN}\,(\mathbf{m}, \mathbf{R})$ to denote the circular symmetric complex Gaussian distribution with mean $\mathbf{m}$ and covariance matrix $\mathbf{R}$. We denote by $\mathbb{R}^+$ the set $[0, \infty)$, by $\mathbb{R}^-$ the set $(-\infty, 0]$, and by $\mathbf{i} = \sqrt{-1}$. $\mathbb{1}_A(x)$ is the indicator function, i.e., $\mathbb{1}_A(x) = 1$ iff $x \in A$ and $\mathbb{1}_A(x) = 0$ otherwise. $\mathbb{E}\,[\cdot]$ denotes the expectation operator. For $(a_n)_{n\geq 1}$ and $(b_n)_{n\geq 1}$ two sequences of random variables, we denote $a_n \asymp b_n$ the equivalence relation $a_n - b_n \xrightarrow{\text{a.s.}} 0$ for $n \to \infty$.

II. Iterative deterministic equivalents

In this section, we will first recall the notion of deterministic equivalents in probability theory before we explain their connections to RMT and the performance analysis of communication systems. We will then introduce the Fubini theorem, which is the key ingredient to extend classical deterministic equivalents to iterative deterministic equivalents.

A. Deterministic equivalents and random matrices

Definition 1: Consider the probability space $(\Omega, \mathcal{F}, P)$. Let $(f_n)_{n\geq 1}$ be a series of measurable complex-valued functions, $f_n : \Omega \times \mathbb{C} \to \mathbb{C}$, and let $(g_n)_{n\geq 1}$ be a series of complex-valued functions, $g_n : \mathbb{C} \to \mathbb{C}$. Then $(g_n)_{n\geq 1}$ is a *deterministic equivalent* of $(f_n)_{n\geq 1}$ on $D \subset \mathbb{C}$, if there exists a set $A \subset \Omega$ with $P(A) = 1$, such that

$$f_n(\omega, z) - g_n(z) \xrightarrow[n\to\infty]{} 0$$

for all $\omega \in A$ and for all $z \in D$.

Otherwise stated, a deterministic equivalent for $(f_n)_{n\geq 1}$ is a series $(g_n)_{n\geq 1}$ such that $g_n(z)$ approximates $f_n(\omega, z)$ arbitrarily closely as n grows, for every $z \in D$ and almost every ω. In particular, if

$(f_n)_{n\geq 1}$ converges almost surely to a limiting function f, i.e., for all $(\omega, z) \in A \times D$ with $A \subset \Omega$, $P(A) = 1$ and $D \subset \mathbb{C}$, we have

$$f_n(\omega, z) \xrightarrow[n\to\infty]{} f(z) \tag{1}$$

then $(g_n)_{n\geq 1}$ defined by $g_n = f$, for all n, is also a deterministic equivalent of $(f_n)_{n\geq 1}$. In many cases, one can further show that $\int_\Omega f_n(\omega, z) dP(\omega) - g_n(z) \xrightarrow[n\to\infty]{} 0$. Thus g_n is also an approximation of the expected value of f_n.

In the context of large dimensional random matrix theory, one often considers random matrices $\mathbf{H}_n \in \mathbb{C}^{N\times n}$ of growing dimensions $N, n \to \infty$, where in general $N/n = c_n$ is such that

$$0 < \liminf_n c_n \leq \limsup_n c_n < \infty. \tag{2}$$

This simply states that c_n is bounded so that the ratio N/n of the matrix dimensions is never too close to zero or infinity. Formally, to be in line with Definition 1, we will define random matrices in the following as series $(\mathbf{H}_n)_{n\geq 1} = (\mathbf{H}_n(\omega))_{n\geq 1}$ of matrices with growing dimensions which are defined on a probability space $(\Omega, \mathcal{F}, P)$, where every $\omega \in \Omega$ generates the whole sequence $(\mathbf{H}_n(\omega))_{n\geq 1}$ and *not* only a single matrix $\mathbf{H}_n(\omega)$.

In wireless communications, one is often interested in the behavior of functionals $f_n(\mathbf{H}_n, z)$, where $\mathbf{H}_n \in \mathbb{C}^{N\times n}$ is a matrix describing the input-output relation of a wireless channel. In particular, $f_n(\mathbf{H}_n, z) = \frac{1}{N}\log\det(\mathbf{I}_N + z\mathbf{H}_n\mathbf{H}_n^{\mathsf{H}})$, $z \in \mathbb{R}^+$, is the (normalized) mutual information of the MIMO channel $\mathbf{H}_n$ between an n-antenna transmitter and an N-antenna receiver at signal-to-noise ratio (SNR) z. Other quantities of interest are the SINR with linear detectors or precoders and the associated rates. The goal of a large system analysis based on RMT is to provide deterministic approximations of these random quantities, which become arbitrarily tight as the system dimensions grow. Thus, deterministic equivalents provide a *deterministic abstraction of the physical layer*. This is particularly interesting for complex channel models which are intractable by exact analysis.

Deterministic equivalents for functionals of large dimensional random matrices have been considered for a wide range of communication channel models. For instance, in [5], a deterministic equivalent for the ergodic mutual information of the Rician fading channel model $\mathbf{H}_n = \mathbf{X}_n + \mathbf{A}_n$ is provided, where $\mathbf{X}_n \in \mathbb{C}^{N\times n}$ has independent entries with zero mean and a variance profile $\mathbb{E}[|(\mathbf{X}_n)_{ij}|^2] = \sigma_{n,ij}^2$, and $\mathbf{A}_n \in \mathbb{C}^{N\times n}$ is a deterministic matrix. In [25], the deterministic equivalent of [5] is used to determine an asymptotically tight approximation of the ergodic capacity achieving input covariance matrix for the MIMO Rician fading channel. Deterministic equivalents were then extended to broader classes of wireless channel models, such as the capacity of the frequency-selective MIMO channel [15], the MIMO MAC

with Kronecker correlation [14] and the sum rate capacity of linearly precoded broadcast channels under imperfect channel state information [9]. The application of such techniques is therefore very broad as it can simplify the difficult study of communication channels with a various number of random parameters (random channels, unitary precoders, path loss, etc.). Moreover, deterministic equivalents can be used to compute approximate solutions of otherwise intractable optimization problems [12], [9], [25].

All of the works mentioned above consider deterministic equivalents for random matrix models created from sums of independent random matrices. In many cases of practical interest, it is however necessary to consider more complex combinations of matrices, such as products or sums of products. These include for example the multi-hop relay channel (Section III-A) as well as the double-scattering channel model [23] (Section III-B). Another recent example is [7], which considers random beamforming over fading channels, i.e., both the precoding and the channel matrices are assumed to be random. In this work, the authors derive deterministic equivalents of the mutual information and of the SINR with MMSE detection with respect to the random precoding matrices for quasi-static channels. Then, a second set of deterministic equivalents is found, treating *both* precoders and channel matrices as random. This technique relies on a fundamental result of probability theory, the Fubini theorem. In this article, we explain this approach in detail and generalize it to the new notion of *iterative deterministic equivalents*.

B. The Fubini theorem

Theorem 1 ([19]): Let $(\Omega, \mathcal{F}, P)$ and $(\Omega', \mathcal{F}', P')$ be two probability spaces. Denote $(\Omega\times\Omega', \mathcal{F}\times\mathcal{F}', Q)$ their product space. Let $f : \Omega \times \Omega' \to \mathbb{R}$ be $(\mathcal{F} \times \mathcal{F}')$-integrable. Then

$$\begin{aligned}\int_{\Omega\times\Omega'} f(\omega,\omega')dQ(\omega,\omega') &= \int_{\Omega}\left[\int_{\Omega'} f(\omega,\omega')dP'(\omega')\right]dP(\omega)\\ &= \int_{\Omega'}\left[\int_{\Omega} f(\omega,\omega')dP(\omega)\right]dP'(\omega').\end{aligned}$$

In particular, consider a set $A \in \mathcal{F} \times \mathcal{F}'$. Then, we have from Theorem 1 that

$$\begin{aligned}Q(A) &= \int_{\Omega\times\Omega'} \mathbb{1}_A(\omega,\omega')dQ(\omega,\omega')\\ &= \int_{\Omega'}\left[\int_{\Omega} \mathbb{1}_A(\omega,\omega')dP(\omega)\right]dP'(\omega'). \qquad (3)\end{aligned}$$

Equation (3) is the core ingredient for the definition of iterative deterministic equivalents: Let $(\mathbf{H}_n(\omega))_{n\geq 1}$ and $(\mathbf{H}'_n(\omega'))_{n\geq 1}$ be two series of random matrices generated by the spaces $(\Omega, \mathcal{F}, P)$ and $(\Omega', \mathcal{F}', P')$, respectively. As in Theorem 1, call Q the product-space measure. Let $f_n((\mathbf{H}_n(\omega), \mathbf{H}'_n(\omega')), z)$ be a

functional of the matrices $\mathbf{H}_n(\omega)$ and $\mathbf{H}'_n(\omega')$. Assume that there is a function $\tilde{g}_n(\mathbf{H}_n(\omega), z)$, such that, for each $\omega \in A \subset \Omega$ with $P(A) = 1$, there exists a subset $B(\omega) \subset \Omega'$ with $P'(B(\omega)) = 1$, on which

$$f_n((\mathbf{H}_n(\omega), \mathbf{H}'_n(\omega')), z) - \tilde{g}_n(\mathbf{H}_n(\omega), z) \to 0. \tag{4}$$

Although $\tilde{g}_n(\mathbf{H}_n(\omega), z)$ is a random function (as it depends on ω), it is independent of $\mathbf{H}'_n(\omega')$. Thus, we can see $\tilde{g}_n(\mathbf{H}_n(\omega), z)$ as a deterministic equivalent of $f_n((\mathbf{H}_n(\omega), \mathbf{H}'_n(\omega')), z)$ with respect to $(\mathbf{H}'_n(\omega'))_{n\geq 1}$. Now, let us assume that there is a second function $g_n(z)$, such that for $\omega \in C \subset \Omega$ with $P(C) = 1$,

$$\tilde{g}_n(\mathbf{H}_n(\omega), z) - g_n(z) \to 0. \tag{5}$$

Call $D = \{(\omega, \omega') : \omega \in A \cap C \ , \ \omega' \in B(\omega)\} \subset \Omega \times \Omega'$, the space on which $f_n((\mathbf{H}_n, \mathbf{H}'_n), z) - g_n(z) \to 0$. Then, from (3), this space has probability

$$\begin{aligned} Q(D) &= \int_{\Omega} \left[\int_{\Omega'} \mathbb{1}_D(\omega, \omega') dP'(\omega') \right] dP(\omega) \\ &\overset{(a)}{\geq} \int_{A \cap C} \left[\int_{B(\omega)} \mathbb{1}_D(\omega, \omega') dP'(\omega') \right] dP(\omega) \\ &\overset{(b)}{=} \int_{A \cap C} dP(\omega) \\ &\overset{(c)}{=} 1 \end{aligned} \tag{6}$$

where (a) is due to $A \cap B \subset \Omega$ and $B(\omega) \subset \Omega'$, (b) follows since $P'(B(\omega)) = 1$ for $\omega \in A$ and (c) holds since $P(A \cap C) = P(A) + P(C) - P(A \cup C) = 1$.

To summarize, if a deterministic equivalent g_n exists for a functional f_n of a *random* series $(\mathbf{H}'_n)_{n\geq 1}$ and a *deterministic* series $(\mathbf{H}_n)_{n\geq 1}$ of matrices, and if additionally it can be proved that this deterministic equivalent holds true for almost every such $(\mathbf{H}_n)_{n\geq 1}$ generated by a space Ω, then the latter is also a deterministic equivalent for the random series $((\mathbf{H}_n, \mathbf{H}'_n))_{n\geq 1}$.

This is the mathematical key idea behind our method to derive iterative deterministic equivalents of functionals $f_n((\mathbf{H}_n(\omega), \mathbf{H}'_n(\omega')), z)$, of two (or more) random matrices. First, one considers one of the sequences of random matrices, e.g. $(\mathbf{H}_n(\omega))_{n\geq 1}$, to be deterministic and derives a deterministic equivalent with respect to $(\mathbf{H}'_n(\omega'))_{n\geq 1}$. In the example above, this was the role of the functional $\tilde{g}_n(\mathbf{H}_n(\omega), z)$ which is independent of $\mathbf{H}'_n(\omega')$. In a second step, one assumes the matrices $(\mathbf{H}_n(\omega))_{n\geq 1}$ to be random and derives an iterative deterministic equivalent $g_n(z)$ of $\tilde{g}_n(\mathbf{H}_n(\omega), z)$. Of course, this procedure can be carried out for any finite number of random matrices where in each step the "randomness" related to one of the matrices is removed. From the above construction, we will call $(g_n)_{n\geq 1}$ an *iterative deterministic equivalent*.

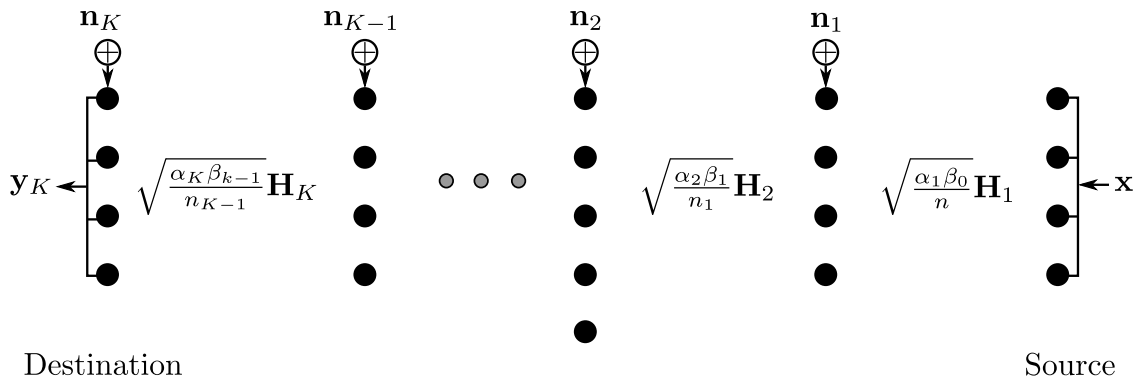

Fig. 1. Multi-hop amplify-and-forward MIMO relay channel.

In the next section, we present two specific examples of iterative deterministic equivalents with applications to the capacity of multi-hop MIMO relay and double-scattering channels. From now on, all matrices and vectors should be understood as sequences of matrices and vectors with growing dimensions. For notational convenience, we drop the index n, e.g. we write $\mathbf{H}$ instead of $(\mathbf{H}_n)_{n\geq 1}$.

III. Applications

A. Multi-hop relay channel

Consider a multi-hop AF MIMO relay channel where a source node communicates via $K-1$ relays with a destination node. There is no direct link between the source and the destination and each relay can only receive data from the preceding hop. This is for example the case if the nodes follow a time-division multiple-access (TDMA) protocol where only one node is transmitting at any given time and the path loss between relay k and $k-2$ is large. Thus each data symbol reaches the destination after K channel uses. The source and destination are respectively equipped with n and n_K antennas while the kth relay has n_k antennas. The relays operate an AF-protocol where each node simply transmits a scaled version of its received signal to the next hop. We will consider a large system limit where $n, n_1, \ldots, n_K$ grow infinitely large at the same speed. Define the following quantities:

$$\begin{aligned} c_1 &= \frac{n}{n_1} \\ c_k &= \frac{n_{k-1}}{n_k}, \qquad k = 2, \ldots, K. \end{aligned} \tag{7}$$

The notation "$n \to \infty$" must be understood from now on as $n \to \infty$, such that $0 < \liminf_n c_k \leq \limsup_n c_k < \infty$ for all k. We denote $\mathbf{y}_k \in \mathbb{C}^{n_k}$ the received base-band signal vector at the kth hop,

given by

$$\mathbf{y}_1 = \sqrt{\alpha_1}\mathbf{H}_1\sqrt{\frac{\beta_0}{n}}\mathbf{x} + \mathbf{n}_1$$
$$\mathbf{y}_k = \sqrt{\alpha_k}\mathbf{H}_k\sqrt{\frac{\beta_{k-1}}{n_{k-1}}}\mathbf{y}_{k-1} + \mathbf{n}_k, \qquad k = 2, \ldots, K \tag{8}$$

where $\mathbf{H}_k \in \mathbb{C}^{n_k \times n_{k-1}}$ is a standard complex Gaussian matrix[1] (let $n_0 \triangleq n$), $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_n)$ is the channel input vector, $\mathbf{n}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{n_k})$ is a noise vector, α_k is a path loss factor, and the parameter β_k is chosen to normalize the transmit power of the kth node according to its power budget $\rho_k > 0$, i.e.,

$$\beta_0 = \frac{\rho_0}{\frac{1}{n}\mathrm{tr}\,\mathbb{E}\left[\mathbf{x}\mathbf{x}^{\mathsf{H}}\right]} = \rho_0$$
$$\beta_k = \frac{\rho_k}{\frac{1}{n_k}\mathrm{tr}\,\mathbb{E}\left[\mathbf{y}_k\mathbf{y}_k^{\mathsf{H}}\right]}, \qquad k = 1, \ldots, K-1. \tag{9}$$

The expectation in the last equation is with respect to the transmit and noise vectors only.[2] The channel matrices $\mathbf{H}_k$ and path loss factors α_k are assumed to be known to the relays and the destination. Since the received signal at each relay is corrupted by noise, the system suffers from noise accumulation. This is in addition to the linear rate loss $\frac{1}{K}$ related to the TDMA protocol. Thus, the capacity decreases quickly with the number of hops K. Note that our system model is different from existing works which consider either no noise [26], or noise only at the destination [22]. An exception is [27], in which the authors consider a similar system model, but do not provide closed-form expressions of the asymptotic mutual information. Several other works deal with the asymptotic capacity of the dual-hop relay channel [28], [29]. Recently, an exact expression of the mutual information of the dual-hop channel for finite channel dimensions was derived in [30]. Here, we will provide an explicit deterministic equivalent of the mutual information at each relay for the general model (8).

Let us introduce the following, recursively defined matrices $\mathbf{R}_k\left(\boldsymbol{\beta}_{k-1}\right)$:

$$\mathbf{R}_0 = \mathbb{E}\left[\mathbf{x}\mathbf{x}^{\mathsf{H}}\right] = \mathbf{I}_n$$
$$\mathbf{R}_k\left(\boldsymbol{\beta}_{k-1}\right) = \mathbb{E}\left[\mathbf{y}_k\mathbf{y}_k^{\mathsf{H}}\right] = \mathbf{I}_{n_k} + \frac{\alpha_k\beta_{k-1}}{n_{k-1}}\mathbf{H}_k\mathbf{R}_{k-1}\left(\boldsymbol{\beta}_{k-2}\right)\mathbf{H}_k^{\mathsf{H}}, \qquad k = 1, \ldots, K \tag{10}$$

and the functionals $\mathcal{J}_k\left(x, \boldsymbol{\beta}_{k-1}\right)$, $x > 0$, which are defined as

$$\mathcal{J}_k\left(x, \boldsymbol{\beta}_{k-1}\right) = \frac{1}{n_k}\log\det\left(\mathbf{I}_{n_k} + x\frac{\alpha_k\beta_{k-1}}{n_{k-1}}\mathbf{H}_k\mathbf{R}_{k-1}\left(\boldsymbol{\beta}_{k-2}\right)\mathbf{H}_k^{\mathsf{H}}\right), \qquad k = 1, \ldots, K \tag{11}$$

[1]A standard complex Gaussian matrix $\mathbf{X}$ has i.i.d. elements $\mathbf{X}_{ij} \sim \mathcal{CN}(0, 1)$.

[2]Under a long-term power constraint, the expectation could be taken also with respect to the matrices $\mathbf{H}_k$. Asymptotically, both constraints are equivalent (see Lemma 1).

where $\boldsymbol{\beta}_k = [\beta_0, \cdots, \beta_k]$. With these definitions, we can express the normalized mutual information $\mathcal{I}_k(\boldsymbol{\beta}_{k-1})$ between $\mathbf{y}_k$ and $\mathbf{x}$ as

$$\mathcal{I}_k(\boldsymbol{\beta}_{k-1}) = \frac{1}{K}\left(\mathcal{J}_k(1, \boldsymbol{\beta}_{k-1}) - \mathcal{J}_k(1, \boldsymbol{\beta}'_{k-1})\right) \tag{12}$$

where $\boldsymbol{\beta}'_k = [0, \beta_1, \cdots, \beta_k]$. Next, we demonstrate by a simple example that (12) holds.

Example 1 (2-hop Relay-channel): The normalized mutual information $\mathcal{I}_2(\boldsymbol{\beta}_1)$ between $\mathbf{x}$ and the channel output after the second hop $\mathbf{y}_2$ is given as

$$\begin{aligned}
\mathcal{I}_2(\boldsymbol{\beta}_1) &= \frac{1}{Kn_2}\log\det\left(\mathbf{I}_{n_2} + \left(\mathbf{I}_{n_2} + \frac{\alpha_2\beta_1}{n_1}\mathbf{H}_2\mathbf{H}_2^{\mathsf{H}}\right)^{-1}\frac{\alpha_2\beta_1\alpha_1\beta_0}{n_1 n}\mathbf{H}_2\mathbf{H}_1\mathbf{H}_1^{\mathsf{H}}\mathbf{H}_2^{\mathsf{H}}\right)\\
&= \frac{1}{Kn_2}\log\det\left(\mathbf{I}_{n_2} + \frac{\alpha_2\beta_1}{n_1}\mathbf{H}_2\left(\mathbf{I}_n + \frac{\alpha_1\beta_0}{n}\mathbf{H}_1\mathbf{H}_1^{\mathsf{H}}\right)\mathbf{H}_2^{\mathsf{H}}\right) - \frac{1}{Kn_2}\log\det\left(\mathbf{I}_{n_2} + \frac{\alpha_2\beta_1}{n_1}\mathbf{H}_2\mathbf{H}_2^{\mathsf{H}}\right)\\
&= \frac{1}{Kn_2}\log\det\left(\mathbf{I}_{n_2} + \frac{\alpha_2\beta_1}{n_1}\mathbf{H}_2\mathbf{R}_1(\beta_0)\mathbf{H}_2^{\mathsf{H}}\right) - \frac{1}{Kn_2}\log\det\left(\mathbf{I}_{n_2} + \frac{\alpha_2\beta_1}{n_1}\mathbf{H}_2\mathbf{R}_1(0)\mathbf{H}_2^{\mathsf{H}}\right)\\
&= \frac{1}{K}\left(\mathcal{J}_2(1, \boldsymbol{\beta}_1) - \mathcal{J}_2(1, \boldsymbol{\beta}'_1)\right).
\end{aligned} \tag{13}$$

In the following, we will derive deterministic equivalents $\bar{\mathcal{I}}_k(\bar{\boldsymbol{\beta}}_{k-1})$ of $\mathcal{I}_k(\boldsymbol{\beta}_{k-1})$. It will turn out that the recursive definition of the matrices $\mathbf{R}_k(\boldsymbol{\beta}_{k-1})$ in (10) allows us to calculate iterative deterministic equivalents of the mutual information after each hop. This is achieved by treating the matrix $\mathbf{R}_{k-1}(\boldsymbol{\beta}_{k-2})$ as deterministic and deriving a deterministic equivalent of $\mathcal{J}_K(x, \boldsymbol{\beta}_{k-1})$ with respect to the matrix $\mathbf{H}_k$. This process can be iterated for $\mathbf{R}_{k-2}(\boldsymbol{\beta}_{k-3}), \mathbf{R}_{k-3}(\boldsymbol{\beta}_{k-4}), \ldots$ and $\mathbf{H}_{k-1}, \mathbf{H}_{k-2}, \ldots$ until the deterministic matrix $\mathbf{R}_0$ is reached. Before we address this problem, we will derive deterministic equivalents $\bar{\beta}_k$ of the power normalization factors β_k:

Lemma 1 (Asymptotic power normalization): Let $\beta_0 = \bar{\beta}_0 = \rho_0$. Then,

$$\beta_k \xrightarrow[n\to\infty]{\text{a.s.}} \bar{\beta}_k = \frac{\rho_k}{1 + \alpha_k\rho_{k-1}}, \qquad k = 1, \ldots, K-1.$$

Proof: Recall the definition of $\beta_k = \frac{\rho_k}{\frac{1}{n_k}\operatorname{tr}\mathbf{R}_k}$, where $\mathbf{R}_k = \mathbf{R}_k\left(\boldsymbol{\beta}_{k-1}\right)$. For $k \geq 1$, we have

$$\begin{aligned}
\frac{1}{n_k}\operatorname{tr}\mathbf{R}_k &= 1 + \frac{\alpha_k\beta_{k-1}}{n_k n_{k-1}}\operatorname{tr}\mathbf{H}_k\mathbf{R}_{k-1}\mathbf{H}_k^{\mathsf{H}}\\
&\overset{(a)}{=} 1 + \alpha_k\beta_{k-1}\frac{1}{n_k}\sum_{j=1}^{n_k}\frac{1}{n_{k-1}}\tilde{\mathbf{h}}_{k,j}^{\mathsf{H}}\mathbf{R}_{k-1}\tilde{\mathbf{h}}_{k,j}\\
&\overset{(b)}{\asymp} 1 + \alpha_k\frac{\rho_{k-1}}{\frac{1}{n_{k-1}}\operatorname{tr}\mathbf{R}_{k-1}}\frac{1}{n_{k-1}}\operatorname{tr}\mathbf{R}_{k-1}\\
&= 1 + \alpha_k\rho_{k-1}
\end{aligned} \tag{14}$$

where (a) is obtained by denoting $\tilde{\mathbf{h}}_{k,j} \in \mathbb{C}^{n_{k-1}}$ the jth row vector of $\mathbf{H}_k$ and (b) is due to Lemma 2 and Lemma 6 in Appendix A and the definition of β_{k-1}. By the continuous mapping theorem [31], we finally have

$$\beta_k = \frac{\rho_k}{\frac{1}{n_k}\mathrm{tr}\,\mathbf{R}_k} \xrightarrow[n\to\infty]{\text{a.s.}} \frac{\rho_k}{1+\alpha_k\rho_{k-1}}. \tag{15}$$

■

In the next theorem, we will build upon Lemma 1, and provide deterministic equivalents of $\mathcal{J}_k(x, \boldsymbol{\beta}_k)$.

Theorem 2: For $k \in \{1, \ldots, K\}$, let $\boldsymbol{\beta}_{k-1} = [\beta_0 \cdots \beta_{k-1}] \geq 0$ be a sequence of random vectors, indexed by n, and $\bar{\boldsymbol{\beta}}_{k-1} = [\bar{\beta}_0 \cdots \bar{\beta}_{k-1}] \geq 0$ be deterministic, such that $\beta_i \xrightarrow[n\to\infty]{\text{a.s.}} \bar{\beta}_i$ for $i = 0, \ldots, k-1$. Then,

$$\mathcal{J}_k\left(x, \boldsymbol{\beta}_{k-1}\right) - \bar{\mathcal{J}}_k\left(x, \bar{\boldsymbol{\beta}}_{k-1}\right) \xrightarrow[n\to\infty]{\text{a.s.}} 0$$

where $\bar{\mathcal{J}}_k\left(x, \bar{\boldsymbol{\beta}}_{k-1}\right)$ is recursively defined for $k \geq 2$ as

$$\bar{\mathcal{J}}_k\left(x, \bar{\boldsymbol{\beta}}_{k-1}\right) = c_k \bar{\mathcal{J}}_{k-1}\left(\frac{x\alpha_k\bar{\beta}_{k-1}}{c_k + x\alpha_k\bar{\beta}_{k-1} + \bar{e}_{k-1}\left(x, \bar{\boldsymbol{\beta}}_{k-1}\right)}, \bar{\boldsymbol{\beta}}_{k-1}\right) + c_k \log\left(1 + \frac{x\alpha_k\beta_{k-1}}{c_k + \bar{e}_{k-1}\left(x, \bar{\boldsymbol{\beta}}_{k-1}\right)}\right)$$
$$+ \log\left(1 + \frac{\bar{e}_{k-1}\left(x, \bar{\boldsymbol{\beta}}_{k-1}\right)}{c_k}\right) - \frac{\bar{e}_{k-1}\left(x, \bar{\boldsymbol{\beta}}_{k-1}\right)}{c_k + \bar{e}_{k-1}\left(x, \bar{\boldsymbol{\beta}}_{k-1}\right)}$$

and $\bar{e}_k\left(x, \bar{\boldsymbol{\beta}}_k\right)$ for $k \geq 0$ is given by Theorem 3. The initial value $\bar{\mathcal{J}}_1\left(x, \bar{\beta}_0\right)$ is given in closed form:

$$\bar{\mathcal{J}}_1\left(x, \bar{\beta}_0\right) = c_1 \log\left(1 + \frac{x\alpha_1\bar{\beta}_0}{c_1 + \bar{e}_0\left(x, \bar{\beta}_0\right)}\right) + \log\left(1 + \frac{\bar{e}_0\left(x, \bar{\beta}_0\right)}{c_1}\right) - \frac{\bar{e}_0\left(x, \bar{\beta}_0\right)}{c_1 + \bar{e}_0\left(x, \bar{\beta}_0\right)}.$$

Proof: The proof is provided in Appendix B. ■

Theorem 2 allows us to compute the quantities $\bar{\mathcal{J}}_k(x, \bar{\boldsymbol{\beta}}_{k-1})$ recursively for any desired relay node k. The values of $\bar{e}_{k-1}\left(x, \bar{\boldsymbol{\beta}}_{k-1}\right)$, needed at each stage, can also be calculated in a recursive manner as shown in the next theorem.

Theorem 3: For $k \in \{1, \ldots, K-1\}$, let $\boldsymbol{\beta}_k = [\beta_0 \cdots \beta_k] \geq 0$ be a sequence of random vectors, indexed by n, and $\bar{\boldsymbol{\beta}}_k = [\bar{\beta}_0 \cdots \bar{\beta}_k] \geq 0$ be deterministic, such that $\beta_i \xrightarrow[n\to\infty]{\text{a.s.}} \bar{\beta}_i$ for $i = 0, \ldots, k$. Let $x > 0$ and denote by $m_k\left(x, \boldsymbol{\beta}_k\right) = \frac{1}{n_{k+1}}\mathrm{tr}\left(\alpha_{k+1}\beta_k\frac{1}{n_k}\mathbf{H}_{k+1}\mathbf{R}_k\left(\boldsymbol{\beta}_{k-1}\right)\mathbf{H}_{k+1}^{\mathsf{H}} + \frac{1}{x}\mathbf{I}_{n_{k+1}}\right)^{-1}$. Then,

$$m_k\left(x, \boldsymbol{\beta}_k\right) - \bar{m}_k\left(x, \bar{\boldsymbol{\beta}}_k\right) \xrightarrow[n\to\infty]{\text{a.s.}} 0$$

where $\bar{m}_k\left(x,\bar{\boldsymbol{\beta}}_k\right)$ is recursively defined for $k\geq 1$ as

$$\bar{m}_k\left(x,\bar{\boldsymbol{\beta}}_k\right)=\frac{xc_{k+1}}{c_{k+1}+\bar{e}_k\left(x,\bar{\boldsymbol{\beta}}_k\right)}$$

and $\bar{e}_k\left(x,\bar{\boldsymbol{\beta}}_k\right)$ is given as the unique positive solution to the following fixed point equation

$$\begin{aligned}\bar{e}_k\left(x,\bar{\boldsymbol{\beta}}_k\right)&=c_{k+1}\left(c_{k+1}+\bar{e}_k\left(x,\bar{\boldsymbol{\beta}}_k\right)\right)\\&\quad-\frac{c_{k+1}\left(c_{k+1}+\bar{e}_k\left(x,\bar{\boldsymbol{\beta}}_k\right)\right)^2}{x\alpha_{k+1}\bar{\beta}_k}\bar{m}_{k-1}\left(\frac{x\alpha_{k+1}\bar{\beta}_k}{c_{k+1}+x\alpha_{k+1}\bar{\beta}_k+\bar{e}_k\left(x,\bar{\boldsymbol{\beta}}_k\right)},\bar{\boldsymbol{\beta}}_{k-1}\right).\end{aligned}$$

The initial values $\bar{m}_0(x,\bar{\beta}_0)$ and $\bar{e}_0(x,\bar{\beta}_0)$ are given in closed form:

$$\begin{aligned}\bar{m}_0(x,\bar{\beta}_0)&=\frac{c_1}{\frac{\alpha_1\bar{\beta}_0}{c_1+\bar{e}_0\left(x,\bar{\beta}_0\right)}+\frac{1}{x}}+(1-c_1)x\\\bar{e}_0\left(x,\bar{\beta}_0\right)&=-\frac{x\alpha_1\bar{\beta}_0(1-c_1)+c_1}{2}+\frac{\sqrt{\left(x\alpha_1\bar{\beta}_0(1-c_1)+c_1\right)^2+4x\alpha_1\bar{\beta}_0c_1^2}}{2}.\end{aligned}$$

Proof: The proof is provided in Appendix C. ■

Remark 3.1: The quantity $m_k\left(x,\boldsymbol{\beta}_k\right)$ can be seen as the Stieltjes transform [3] of the empirical spectral distribution (e.s.d.) of the matrix $\alpha_{k+1}\beta_k\frac{1}{n_k}\mathbf{H}_{k+1}\mathbf{R}_k(\boldsymbol{\beta}_{k-1})\mathbf{H}_{k+1}^{\mathsf{H}}$ evaluated at $-\frac{1}{x}$. One can further show that Theorem 3 implies the weak convergence of the e.s.d. $\alpha_{k+1}\beta_k\frac{1}{n_k}\mathbf{H}_{k+1}\mathbf{R}_k(\boldsymbol{\beta}_{k-1})\mathbf{H}_{k+1}^{\mathsf{H}}$ to a distribution function, whose Stieltjes transform is given by $\bar{m}_k$, for almost every $\mathbf{H}_1,\ldots,\mathbf{H}_K$.

Applying Theorem 2 and Lemma 1 to (12) yields the following corollary which provides a deterministic equivalent of the mutual information $\mathcal{I}_k(\boldsymbol{\beta}_{k-1})$:

Corollary 1 (Asymptotic mutual information of the K-hop AF MIMO Relay channel):

$$\mathcal{I}_k\left(\boldsymbol{\beta}_{k-1}\right)-\bar{\mathcal{I}}_k\left(\bar{\boldsymbol{\beta}}_{k-1}\right)\xrightarrow[n\to\infty]{\text{a.s}}0\qquad k=1,\ldots,K$$

where

$$\bar{\mathcal{I}}_k\left(\bar{\boldsymbol{\beta}}_{k-1}\right)=\frac{1}{K}\left(\bar{\mathcal{J}}_k(1,\bar{\boldsymbol{\beta}}_{k-1})-\bar{\mathcal{J}}_k(1,\bar{\boldsymbol{\beta}}'_{k-1})\right)$$

with $\bar{\boldsymbol{\beta}}_{k-1}=[\bar{\beta}_0\cdots\bar{\beta}_{k-1}]$, $\bar{\boldsymbol{\beta}}'_{k-1}=[0\;\bar{\beta}_1\cdots\bar{\beta}_{k-1}]$ as given by Lemma 1, and $\bar{\mathcal{J}}_k(x,\bar{\boldsymbol{\beta}}_{k-1})$ and $\bar{\mathcal{J}}_k(x,\bar{\boldsymbol{\beta}}'_{k-1})$ as given by Theorem 2.

Remark 3.2: The values of $\bar{\mathcal{I}}_k(\bar{\boldsymbol{\beta}}_{k-1})$ can be very easily numerically computed. We provide the Matlab code which was used to generate the numerical results in this section in Appendix J. Due to the recursive

implementation, the computational complexity grows quickly with k. Calculating $\bar{\mathcal{J}}_k(x, \bar{\boldsymbol{\beta}}_{k-1})$ with high precision for large values of k (> 10) seems therefore impractical.

We would now like to verify our analysis by some numerical results. To this end, we consider a system with three relays, i.e., $K = 4$. We assume $n = n_4 = 4$, $n_1 = n_3 = 8$, $n_2 = 12$, $\rho_1 = \rho_3 = 0.7\rho_0$ and $\rho_2 = 0.5\rho_0$. The last assumption allows us to control the transmit power of all nodes by the transmit SNR ρ_0 of the source node. We further assume the path loss factors $\alpha_1 = 1$, $\alpha_2 = \alpha_4 = 0.7$, $\alpha_3 = 0.5$. Fig. 2 shows the average normalized mutual information $\mathbb{E}\left[\frac{n_k}{n}\mathcal{I}_k\left(\boldsymbol{\beta}_{k-1}\right)\right]$ after each hop ($k = 1, \ldots, 4$) versus the transmit power ρ_0 of the source node. Note that we have re-normalized all results by $\frac{n_k}{n}$ to put them on a common ground for comparison. The deterministic equivalents $\frac{n_k}{n}\bar{\mathcal{I}}_k\left(\bar{\boldsymbol{\beta}}_{k-1}\right)$ as provided by Corollary 1 are drawn by solid lines, simulation results are represented by markers. The error bars represent one standard deviation of the simulation results in each direction. We can observe a very good fit between the asymptotic approximations and the simulation results for all k and the entire range of ρ_0. As expected, the performance decreases rapidly with each hop.

Finally, we would like to remark that, although we have considered a rather simple channel model with neither antenna correlation nor precoding at the nodes, more involved channel models can be treated in a straightforward fashion with the same techniques.

B. Double-scattering MAC

Consider a discrete-time MIMO MAC from K transmitters, equipped with $n_k,\ k = 1, \ldots, K$, antennas, respectively, to a receiver with N antennas. The channel output vector $\mathbf{y} \in \mathbb{C}^N$ reads

$$\mathbf{y} = \sum_{k=1}^{K} \mathbf{H}_k \mathbf{x}_k + \mathbf{n} \tag{16}$$

where $\mathbf{H}_k \in \mathbb{C}^{N \times n_k}$ and $\mathbf{x}_k \in \mathbb{C}^{n_k}$ are the channel matrix and the transmit vector associated with the kth transmitter, $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \frac{1}{\rho}\mathbf{I}_N)$ is a noise vector and $\rho > 0$ denotes the SNR. We assume Gaussian signaling, i.e., $\mathbf{x}_k = [x_{k,1}, \ldots, x_{k,n_k}]^{\mathsf{T}} \sim \mathcal{CN}(\mathbf{0}, \mathbf{Q}_k)$, where $\mathbf{Q}_k \in \mathbb{C}^{n_k \times n_k}$. The channel matrices $\mathbf{H}_k$ are modeled by the double-scattering model [23]

$$\mathbf{H}_k = \frac{1}{\sqrt{N_k n_k}} \mathbf{R}_k^{\frac{1}{2}} \mathbf{W}_{1,k} \mathbf{S}_k^{\frac{1}{2}} \mathbf{W}_{2,k} \mathbf{T}_k^{\frac{1}{2}} \tag{17}$$

where $\mathbf{R}_k \in \mathbb{C}^{N \times N}$, $\mathbf{S}_k \in \mathbb{C}^{N_k \times N_k}$ and $\mathbf{T}_k \in \mathbb{C}^{n_k \times n_k}$ are deterministic correlation matrices, while $\mathbf{W}_{1,k} \in \mathbb{C}^{N \times N_k}$ and $\mathbf{W}_{2,k} \in \mathbb{C}^{N_k \times n_k}$ are independent standard complex Gaussian matrices. Since the distributions of $\mathbf{W}_{1,k}$ and $\mathbf{W}_{2,k}$ are unitarily invariant, we can assume $\mathbf{S}_k = \mathrm{diag}(s_{k,1}, \ldots, s_{k,N_k})$ to be diagonal matrices, without loss of generality for the statistics of $\mathbf{y}$.

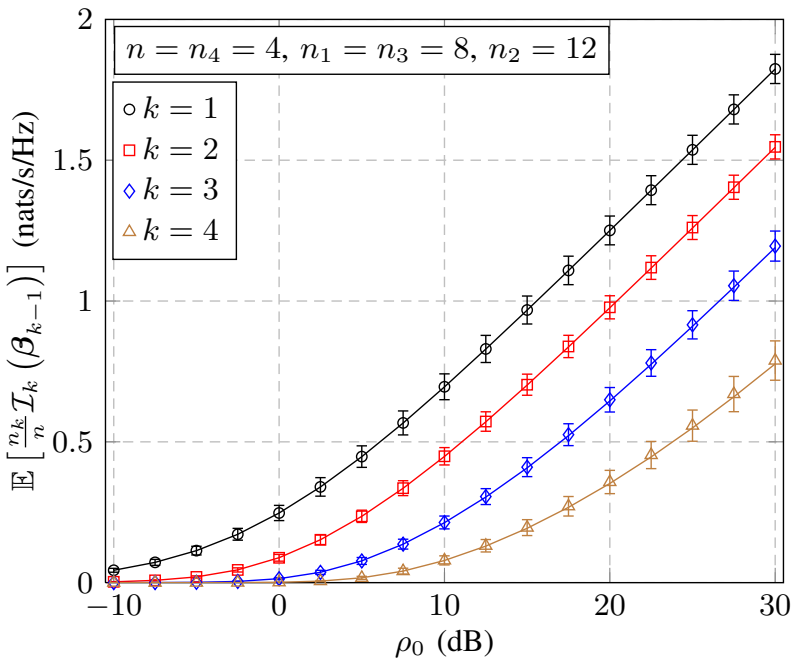

Fig. 2. Average normalized mutual information $\mathbb{E}\left[\frac{n_k}{n}\mathcal{I}_k(\boldsymbol{\beta}_{k-1})\right]$ after the kth hop versus the transmit SNR ρ_0 of the source node. The deterministic equivalents $\frac{n_k}{n}\bar{\mathcal{I}}_k\left(\bar{\boldsymbol{\beta}}_{k-1}\right)$ are drawn by solid lines, the simulation results by markers. The error bars correspond to one standard deviation of the simulation results in each direction.

The double-scattering model [23] was motivated by the observation of low-rank channel matrices, despite low antenna correlation at the transmitter and receiver, see e.g. [24], [32]. A special case of the double-scattering model is the keyhole channel [33], [34], which exhibits null antenna correlation, i.e., $\mathbf{R}_k = \mathbf{I}_N$ and $\mathbf{T}_k = n_k$ for all k, but only a single degree of freedom. The existence of such channels (under laboratory conditions) was confirmed by measurements in [34]. Several theoretical works have studied the double-scattering model so far. The authors of [35] derive capacity upper-bounds for the general model and a closed-form expression for the keyhole channel. An asymptotic study of the outage capacity of the multi-keyhole channel was presented in [36]. The diversity order of the double-scattering model was considered in [37] and it was shown that a MIMO system with t transmit antennas, r receive antennas and s scatterers achieves the diversity of order $trs/\max(t,r,s)$. A closed-from expression of the diversity-multiplexing trade-off (DMT) was derived in [38]. Beamforming along the strongest eigenmode over Rayleigh product MIMO channels, i.e., the double-scattering model without any form of correlation, was considered in [39]. Here, the authors derive exact expressions of the cumulative distribution function (cdf) and the probability density function (pdf) of the largest eigenvalue of the Gramian of the channel

matrix and compute closed-form results for the ergodic capacity, outage probability and SINR distribution. In a later paper [40], the MIMO MAC with double-scattering fading is analyzed. The authors obtain closed-form upper-bounds on the sum-capacity and prove that the transmitters should send their signals along the eigenvectors of the transmit correlation matrices in order to achieve capacity. Despite the significant interest in the double-scattering channel model, little work has been done to study its asymptotic performance when the channel dimensions grow large. We are only aware of [32], in which a model without transmit and receive correlation is studied relying on tools from free probability theory. Implicit expressions of the asymptotic mutual information and the SINR with MMSE detection are found therein.

In the following, we provide deterministic equivalents of the mutual information, the SINR with MMSE-detection and the associated sum-rate. In addition, we derive the precoders which maximize the deterministic equivalent of the mutual information and provide a simple algorithm for their computation. The key idea behind the following proofs is that the double-scattering channel can be interpreted as a Kronecker channel [14] with a *random* receive correlation matrix, which itself is modeled by the Kronecker model. This observation allows us to build upon [14] which provides an asymptotic analysis of the performance of Kronecker channels with deterministic correlation matrices (Theorem 9 in Appendix A). Based on the Fubini theorem, we extend this work by allowing the correlation matrices to be random. A similar technique can be applied to more involved channel models, such as channels with line-of-sight components or MIMO product channels with an arbitrary number of matrices.

Denote $I_N(\rho)$ the instantaneous normalized mutual information of the channel (16), defined as [41]

$$I_N(\rho) = \frac{1}{N} \log\det\left(\mathbf{I}_N + \rho \sum_{k=1}^{K} \mathbf{H}_k \mathbf{Q}_k \mathbf{H}_k^{\mathsf{H}}\right). \tag{18}$$

Moreover, denote $\gamma_{k,j}^N(\rho)$ the SINR at the output of the MMSE detector related to the transmit symbol $x_{k,j}$, given by [42]

$$\gamma_{k,j}^N(\rho) = \mathbf{h}_{k,j}^{\mathsf{H}} \left(\sum_{i=1}^{K} \mathbf{H}_i \mathbf{H}_i^{\mathsf{H}} - \mathbf{h}_{k,j}\mathbf{h}_{k,j}^{\mathsf{H}} + \frac{1}{\rho}\mathbf{I}_N\right)^{-1} \mathbf{h}_{k,j}. \tag{19}$$

We define the normalized sum-rate $R_N(\rho)$ with MMSE detection as

$$R_N(\rho) = \frac{1}{N} \sum_{k=1}^{K} \sum_{j=1}^{n_k} \log\left(1 + \gamma_{k,j}^N(\rho)\right). \tag{20}$$

The notation "$N \to \infty$" will be used to denote that N and all N_k, n_k grow infinitely large, satisfying $0 < \liminf \frac{N_k}{N} \leq \limsup \frac{N_k}{N} < \infty$ and $0 < \liminf \frac{n_k}{N} \leq \limsup \frac{n_k}{N} < \infty$. Additionally, we need the following technical assumption:

A 1: For all k, $\limsup_N \|\mathbf{R}_k\| < \infty$, $\limsup_N \|\mathbf{S}_k\| < \infty$ and $\limsup_N \|\mathbf{T}_k\mathbf{Q}_k\| < \infty$.

Remark 3.3: This assumption implies in particular that the antenna correlation at the transmitter and receiver side cannot grow with the system size, as it would be the case for very dense antenna arrays [43]. Amendments to relax this assumption can be made, following the work in [14]. Moreover, the last constraint, $\limsup_N \|\mathbf{T}_k\mathbf{Q}_k\| < \infty$, implies that no transmitter is allowed to focus an increasing amount of transmit power in a single direction.

Our first theorem introduces a set of $3K$ implicit equations which uniquely determines some quantities $(g_k, \bar{g}_k, \delta_k)$ $(1 \leq k \leq K)$. These will be needed later on to provide deterministic equivalents of $I_N(\rho)$, $\gamma_{k,j}^N\left(\sigma^2\right)$, and $R_N\left(\rho\right)$.

Theorem 4 (Fundamental equations): The following system of $3K$ implicit equations in $\bar{g}_k$, g_k, and δ_k $(1 \leq k \leq K)$:

$$\begin{aligned}
\bar{g}_k &= \frac{1}{n_k}\mathrm{tr}\,\mathbf{T}_k^{\frac{1}{2}}\mathbf{Q}_k\mathbf{T}_k^{\frac{1}{2}}\left(g_k\mathbf{T}_k^{\frac{1}{2}}\mathbf{Q}_k\mathbf{T}_k^{\frac{1}{2}} + \mathbf{I}_{n_k}\right)^{-1} \\
g_k &= \frac{1}{n_k}\sum_{j=1}^{N_k}\frac{s_{k,j}\delta_k}{1+\bar{g}_k s_{k,j}\delta_k} \\
\delta_k &= \frac{1}{N_k}\mathrm{tr}\,\mathbf{R}_k\left(\sum_{i=1}^{K}\frac{n_i}{N_i}\frac{\bar{g}_i g_i}{\delta_i}\mathbf{R}_i + \frac{1}{\rho}\mathbf{I}_N\right)^{-1}
\end{aligned} \tag{21}$$

has a unique solution satisfying $\bar{g}_k, g_k, \delta_k > 0$ for all k and $\rho > 0$.

Proof: The proof is provided in Appendix D. ■

Remark 3.4: The values of $\bar{g}_k$, g_k, and δ_k can be computed by a standard fixed-point algorithm which iteratively computes (21), starting from some arbitrary initialization $\bar{g}_k^{(0)}, g_k^{(0)}, \delta_k^{(0)} > 0$. This algorithm is proved to converge, generally terminates within a few iterations (depending on the system size and the desired accuracy), and does not pose any computational challenge.

The next theorem provides a deterministic equivalent of the (ergodic) mutual information based on the quantities $(g_k, \bar{g}_k, \delta_k)$ as provided by Theorem 4.

Theorem 5 (Mutual information): Assume that **A 1** holds. Then,

$$(i) \quad I_N\left(\rho\right) - \bar{I}_N\left(\rho\right) \xrightarrow[N\to\infty]{\text{a.s.}} 0$$

$$(ii) \quad \mathbb{E}\left[I_N\left(\rho\right)\right] - \bar{I}_N\left(\rho\right) \xrightarrow[N\to\infty]{} 0$$

where

$$\bar{I}_N(\rho) = \frac{1}{N}\log\det\left(\mathbf{I}_N + \rho\sum_{k=1}^{K}\frac{n_k}{N_k}\frac{\bar{g}_k g_k}{\delta_k}\mathbf{R}_k\right) + \frac{1}{N}\sum_{k=1}^{K}\left[\log\det\left(\mathbf{I}_{N_k} + \bar{g}_k\delta_k\mathbf{S}_k\right) + \log\det\left(\mathbf{I}_{n_k} + g_k\mathbf{T}_k^{\frac{1}{2}}\mathbf{Q}_k\mathbf{T}_k^{\frac{1}{2}}\right) - 2n_k g_k\bar{g}_k\right] \quad (22)$$

and $g_k, \bar{g}_k, \delta_k$ are the unique positive solutions to (21).

Proof: The proof is provided in Appendix E. ■

The following result allows us to compute the asymptotically optimal precoding matrices $\mathbf{Q}_k$ which maximize $\bar{I}_N(\rho)$ under individual transmit power constraints.

Theorem 6 (Optimal power allocation): The solution to the following optimization problem:

$$\left(\bar{\mathbf{Q}}_1^\star, \ldots, \bar{\mathbf{Q}}_K^\star\right) = \arg\max_{\mathbf{Q}_1,\ldots,\mathbf{Q}_k} \bar{I}_N(\rho) \quad \text{s.t.} \quad \frac{1}{n_k}\mathrm{tr}\,\mathbf{Q}_k \leq P_k \ \forall k$$

where $\bar{I}_N(\rho)$ is defined in Theorem 5, is given as $\bar{\mathbf{Q}}_k^\star = \mathbf{U}_k\bar{\mathbf{P}}_k^\star\mathbf{U}_k^{\mathsf{H}}$, where $\mathbf{U}_k \in \mathbb{C}^{n_k\times n_k}$ is defined by the spectral decomposition of $\mathbf{T}_k = \mathbf{U}_k\mathrm{diag}(t_{k,1},\ldots,t_{k,n_k})\mathbf{U}_k^{\mathsf{H}}$ and $\bar{\mathbf{P}}_k^\star = \mathrm{diag}(\bar{p}_{k,1}^\star,\ldots,\bar{p}_{k,n_k}^\star)$ is given by the water-filling solution:

$$\bar{p}_{k,j}^\star = \left(\mu_k - \frac{1}{g_k^\star t_{k,j}}\right)^+ \quad (23)$$

where μ_k is chosen to satisfy $\frac{1}{n_k}\mathrm{tr}\,\bar{\mathbf{P}}_k^\star = P_k$ and $g_k^\star$ is given by Theorem 4 for $\mathbf{Q}_k = \bar{\mathbf{Q}}_k^\star$.

Proof: The proof is provided in Appendix F. ■

Remark 3.5: The optimal power allocation matrices $\bar{\mathbf{P}}_k^\star$ can be calculated by the iterative water-filling Algorithm 1 (see [14, Remark 2] and [7, Remark 3] for a discussion on the convergence of this algorithm).

Remark 3.6: Denote by $(\mathbf{Q}_1^\star,\ldots,\mathbf{Q}_K^\star)$ the set of precoding matrices which maximize $\mathbb{E}\left[I_N(\rho)\right]$ for a given set of power constraints. If the condition $\limsup\|\mathbf{T}_k\mathbf{Q}_k^\star\| < \infty$ holds for all k, then $\mathbb{E}\left[I_N(\rho,\mathbf{Q}_1^\star,\ldots,\mathbf{Q}_K^\star)\right] - \bar{I}_N(\rho,\bar{\mathbf{Q}}_1^\star,\ldots,\bar{\mathbf{Q}}_K^\star) \xrightarrow[N\to\infty]{} 0$, by Theorem 5 and the strict concavity of $\bar{I}_N(\rho)$ and $I_N(\rho)$ in the matrices $\mathbf{Q}_k$. However, this condition is difficult to verify and is outside the scope of this paper. See [25] for such a technical discussion in the case of Rician fading channels.

Next, we provide deterministic equivalents of the SINR $\gamma_{k,j}^N(\rho)$ at the output of the MMSE detector and the associated sum-rate $R_N(\rho)$.

Algorithm 1 Iterative water-filling algorithm

Let $\epsilon > 0$, $n = 0$ and $\bar{p}_{k,j}^{\star,0} = P_k$ for all k, j.
repeat
  For all k, compute $g_k^{\star,n}$ according to Theorem 4 with matrices $\mathbf{Q}_k = \mathbf{U}_k \text{diag}\left(\bar{p}_{k,j}^{\star,n}\right)\mathbf{U}_k^{\mathsf{H}}$.
  For all k, j, calculate $\bar{p}_{k,j}^{\star,n+1} = \left(\mu_k - \frac{1}{g_k^{\star,n} t_{k,j}}\right)^+$, with μ_k such that $\frac{1}{n_k}\sum_{j=1}^{n_k} \bar{p}_{k,j}^{\star,n+1} = P_k$.
  $n = n + 1$
until $\max_{k,j} |\bar{p}_{k,j}^{\star,n} - \bar{p}_{k,j}^{\star,n-1}| \leq \epsilon$

Theorem 7 (SINR of the MMSE detector): Let $\mathbf{Q}_k = \text{diag}\left(p_{k,1}, \ldots, p_{k,n_k}\right)$ and $\mathbf{T}_k = \text{diag}(t_{k,1}, \ldots, t_{k,n_k})$ for all k. Assume that **A 1** holds. Then,

$$\gamma_{k,j}^N(\rho) - \bar{\gamma}_{k,j}^N(\rho) \xrightarrow[N\to\infty]{\text{a.s.}} 0$$

where $\bar{\gamma}_{k,j}^N(\rho) = p_{k,j} t_{k,j} g_k$ and g_k is by given by Theorem 4.

Proof: The proof is provided in Appendix G. ■

Remark 3.7: It is easy to see that the theorem is also valid under the more general assumption $\mathbf{T}_k = \mathbf{U}_k \text{diag}(t_{k,1}, \ldots, t_{k,n_k})\mathbf{U}_k^{\mathsf{H}}$ and $\mathbf{Q}_k = \mathbf{U}_k \text{diag}(p_{k,1}, \ldots, p_{k,n_k})\mathbf{U}_k^{\mathsf{H}}$.

Corollary 2 (Sum-rate with MMSE decoding): Let $\mathbf{Q}_k = \text{diag}\left(p_{k,1}, \ldots, p_{k,n_k}\right)$ and $\mathbf{T}_k = \text{diag}(t_{k,1}, \ldots, t_{k,n_k})$ for all k. Assume that **A 1** holds. Then,

$$(i) \quad R_N(\rho) - \bar{R}_N(\rho) \xrightarrow[N\to\infty]{\text{a.s.}} 0$$

$$(ii) \quad \mathbb{E}\left[R_N(\rho)\right] - \bar{R}_N(\rho) \xrightarrow[N\to\infty]{\text{a.s.}} 0$$

where

$$\bar{R}_N(\rho) = \frac{1}{N}\sum_{k=1}^{K}\sum_{j=1}^{n_k} \log\left(1 + \bar{\gamma}_{k,j}^N(\rho)\right)$$

and the $\bar{\gamma}_{k,j}^N(\rho)$ are given by Theorem 7.

Proof: The proof is provided in Appendix H. ■

Remark 3.8: Careful inspection of (22) reveals that the third term of $\bar{I}_N(\rho)$ equals $\bar{R}_N(\rho)$ since

$$\frac{1}{N}\sum_{k=1}^{K} \log\det\left(\mathbf{I}_{n_k} + g_k \mathbf{T}_k^{\frac{1}{2}} \mathbf{Q}_k \mathbf{T}_k^{\frac{1}{2}}\right) = \frac{1}{N}\sum_{k=1}^{K}\sum_{j=1}^{n_k} \log\left(1 + p_{k,j} t_{k,j} g_k\right). \tag{24}$$

Thus, all other terms in (22) correspond to the gains of successive interference cancellation [44] over simple MMSE detection.

A special case of the double-scattering channel is the Rayleigh product MIMO channel [39] which does not exhibit any form of correlation between the transmit and receive antennas or the scatterers. For this model, Theorems 4, 5 and 7 can be given in closed form as shown in the next corollary.

Corollary 3 (Rayleigh product channel): For all k, let $N_k = S$, $n_k = N$ and assume $\mathbf{T}_k = \mathbf{I}_N$, $\mathbf{S}_k = \mathbf{I}_S$, $\mathbf{R}_k = \mathbf{I}_N$, and $\mathbf{Q}_k = \mathbf{I}_N$. Then $\bar{I}_N(\rho)$ and $\bar{\gamma}_{k,j}^N(\rho)$ as defined in Theorems 5 and 7 can be given in closed form as

$$\bar{I}_N(\rho) = \log\left(1 + \rho\frac{NK}{S}\bar{g}\left(\bar{g} + \frac{S}{N} - 1\right)\right) - \frac{KS}{N}\log\left(1 + \frac{N}{S}(\bar{g} - 1)\right) - K\log(\bar{g}) - 2K(1 - \bar{g})$$

and

$$\bar{\gamma}_{k,j}^N(\rho) = \frac{1 - \bar{g}}{\bar{g}}$$

where $\bar{g}$ is the unique root to

$$\bar{g}^3 - \bar{g}^2\left(2 - \frac{S}{N} - \frac{1}{K}\right) + \bar{g}\left(1 - \frac{S}{N} - \frac{1}{K} + \frac{S}{NK}\left(1 + \frac{1}{\rho}\right)\right) - \frac{S}{NK}\frac{1}{\rho} = 0 \tag{25}$$

such that $\bar{g} \in \left(1 - \min\left[\frac{1}{K}, \frac{S}{N}\right], 1\right)$.

Proof: The proof is provided in Appendix I. ■

Note that similar expressions for the asymptotic mutual information and MMSE-SINR have been obtained in [32] by means of free probability theory. However, these results require the numerical solution of a third order differential equation.

As a first numerical example, we consider the "multi-keyhole channel", i.e., $K = 1$, $\mathbf{S}_1 = \mathbf{I}_{N_1}$, $\mathbf{R}_1 = \mathbf{I}_N$, $\mathbf{T}_1 = \mathbf{Q}_1 = \mathbf{I}_{n_1}$, for $N = n_1 = 4$. Fig. 3 depicts the normalized ergodic mutual information $\mathbb{E}[I_N(\rho)]$ and its asymptotic approximation $\bar{I}_N(\rho)$ versus SNR for different numbers of "keyholes" $N_1 \in \{1, 2, 3, 4, 100\}$. Surprisingly, the match between both results is almost perfect although the channel dimensions are very small. As one expects, the multiplexing gain increases linearly with N_1 until $N_1 \geq N = 4$. Larger values of N_1 only change the statistical distribution of the channel matrix while the degrees of freedom are limited by the number of antennas (for $N_1 \to \infty$, $\mathbf{H}_1$ becomes a standard Rayleigh fading channel [23]).

As a second example, we consider a MAC from $K = 3$ transmitters, assuming the double-scattering model in [23]. Under this model, the correlation matrices are given as $\mathbf{R}_k = \mathbf{G}(\phi_{r,k}, d_{r,k}, N_k)$, $\mathbf{S}_k =$

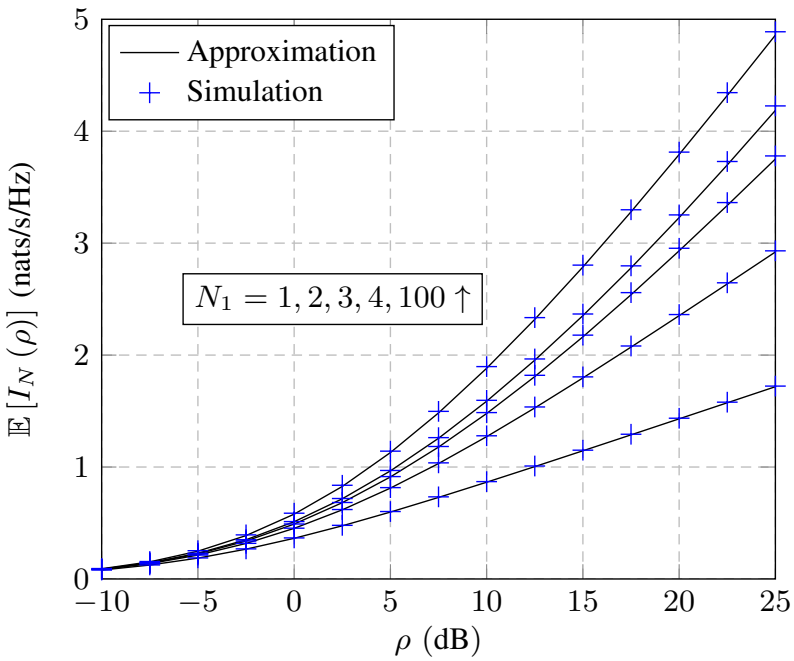

Fig. 3. Ergodic mutual information $\mathbb{E}[I_N(\rho)]$ of the multi-keyhole channel and its deterministic equivalent $\bar{I}_N(\rho)$ versus ρ.

$\mathbf{G}(\phi_{s,k}, d_{s,k}, N_k)$ and $\mathbf{T}_k = \mathbf{G}(\phi_{t,k}, d_{t,k}, N_k)$, where $\mathbf{G}(\phi, d, n)$ is defined as

$$[\mathbf{G}(\phi, d, n)]_{k,l} = \frac{1}{n} \sum_{j=\frac{1-n}{2}}^{\frac{n-1}{2}} \exp\left(\mathbf{i}2\pi d(k-l)\sin\left(\frac{j\phi}{1-n}\right)\right). \tag{26}$$

The values $\phi_{t,k}$ and $\phi_{r,k}$ determine the angular spread of the radiated and received signals, $d_{t,k}$ and $d_{r,k}$ are the antenna spacings at the kth transmitter and receiver in multiples of the signal wavelength, N_k can be seen as the number of scatterers and $d_{s,k}$ as the spacing of the scatterers. For simplicity, we assume $N = 4$, $P_k = 1/n_k$, $N_k = 11$, $n_k = 3$, $d_{t,k} = d_{r,k} = 0.25$ and $d_{s,k} = 50$ for all k. We further assume $\phi_{r,k} = \phi_{t,k}$ for all k, with $\phi_{r,k} \in \{\pi/4, \pi/2, \pi\}$ and $\phi_{s,k} = \pi/8$. Fig. 4 shows $\mathbb{E}[I_N(\rho)]$ and $\bar{I}_N(\rho)$ with uniform and optimal power allocation versus SNR. Again, our asymptotic results yield very tight approximations, even for small system dimensions. Note that we have used the precoding matrices provided by Theorem 6 for the simulations as the optimal precoding matrices are unknown. For comparison, we also provide the sum-rate with MMSE detection $\mathbb{E}[R_N(\rho)]$ and its deterministic approximation $\bar{R}_N(\rho)$. We observe a good fit between both results at low SNR values, but a slight mismatch for higher values. This is due to a slower convergence of the SINR $\gamma_{k,j}^N(\rho)$ to its deterministic approximation $\bar{\gamma}_{k,j}^N(\rho)$, well documented in the RMT literature, e.g. [45].

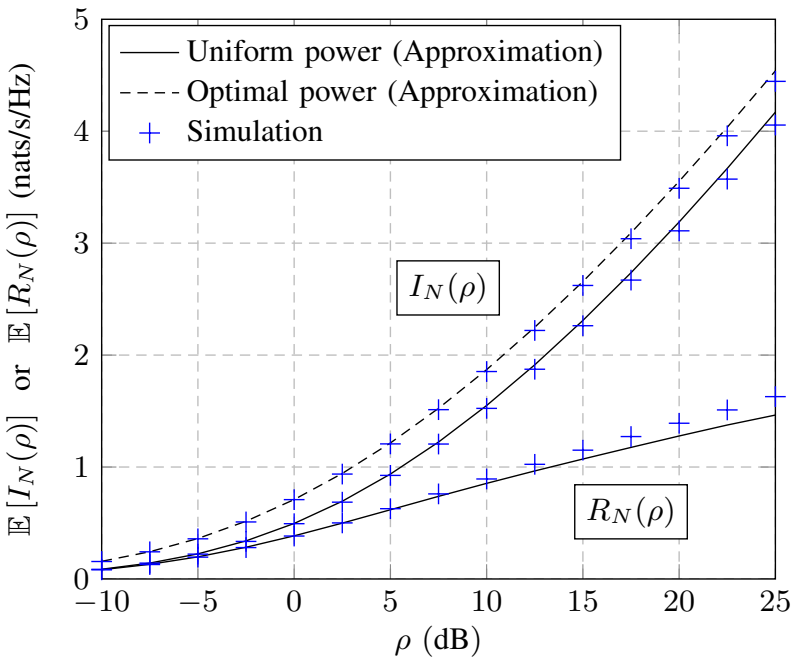

Fig. 4. Ergodic mutual information $\mathbb{E}[I_N(\rho)]$ and sum-rate $\mathbb{E}[R_N(\rho)]$ of the multiple access channel and their asymptotic approximations $\bar{I}_N(\rho)$ and $\bar{R}_N(\rho)$ versus ρ.

IV. Conclusion

In this paper, we have presented a novel tool for the large system analysis of communication systems, called iterative deterministic equivalents. This tool is particularly suited for the analysis of channel models composed of complex combinations of independent random matrices, e.g. products or sums of products of matrices. We have demonstrated the usefulness of this approach with the help of two examples which had not been solved in the literature before. These are the multi-hop AF MIMO relay channel with noise at each stage and the MIMO MAC under the double-scattering channel model. For these channel models, we have provided asymptotically tight deterministic approximations of information theoretic quantities of interest, such as the mutual information and the sum-rate with MMSE detection. These approximations can be easily computed by provably converging fixed-point algorithms and do not require any numerical integration. Our simulation results suggest that the asymptotic performance approximations are very accurate for finite system dimensions with only a few antennas at each node. Finally, the method of iterative deterministic equivalents is applicable to a wide range of channel models of interest (e.g. combinations of correlated i.i.d. and random unitary matrices [7]) which cannot be easily treated so far with other techniques.

APPENDIX A
RELATED RESULTS

Lemma 2 ([46, Lemma 2.7], [9, Lemma 4]): Let $(\mathbf{A}_N)_{N\geq 1}$ be a sequence of random $N\times N$ matrices, satisfying $\limsup_N \|\mathbf{A}_N\| < \infty$, almost surely. Let $(\mathbf{x}_N)_{N\geq 1}$ be a sequence of random N-dimensional vectors of i.i.d. entries with zero mean, variance $1/N$ and 8th order moment of order $O\left(1/N^4\right)$, independent of $\mathbf{A}_N$. Then,

$$\mathbf{x}_N^{\mathsf{H}}\mathbf{A}_N\mathbf{x}_N - \frac{1}{N}\mathrm{tr}\,\mathbf{A}_N \xrightarrow[N\to\infty]{\text{a.s.}} 0.$$

Lemma 3 (Matrix inversion lemma [47, Eq. (2.2)]): Let $\mathbf{A}\in\mathbb{C}^{N\times N}$ be Hermitian invertible. Then, for any vector $\mathbf{x}\in\mathbb{C}^N$ and any scalar $\tau\in\mathbb{C}$ such that $\mathbf{A}+\tau\mathbf{x}\mathbf{x}^{\mathsf{H}}$ is invertible,

$$\mathbf{x}^{\mathsf{H}}\left(\mathbf{A}+\tau\mathbf{x}\mathbf{x}^{\mathsf{H}}\right)^{-1} = \frac{\mathbf{x}^{\mathsf{H}}\mathbf{A}^{-1}}{1+\tau\mathbf{x}^{\mathsf{H}}\mathbf{A}^{-1}\mathbf{x}}.$$

Lemma 4 (Rank-1 perturbation lemma [47]): Let $z<0$, $\mathbf{A}\in\mathbb{C}^{N\times N}$, $\mathbf{B}\in\mathbb{C}^{N\times N}$ with $\mathbf{B}$ Hermitian nonnegative definite, and $\mathbf{v}\in\mathbb{C}^N$. Then,

$$\left|\mathrm{tr}\left((\mathbf{B}-z\mathbf{I}_N)^{-1}-(\mathbf{B}+\mathbf{v}\mathbf{v}^{\mathsf{H}}-z\mathbf{I}_N)^{-1}\right)\mathbf{A}\right| \leq \frac{\|\mathbf{A}\|}{|z|}\ .$$

Lemma 5: Let $\mathbf{R}\in\mathbb{C}^{N\times N}$ be Hermitian with smallest eigenvalue $\lambda_{\min}\geq 1$ and $a,b,c,d>0$. Then

$$\frac{1}{N}\mathrm{tr}\,\mathbf{R}\,(a\mathbf{R}+b\mathbf{I}_N)^{-1}\,\mathbf{R}\,(c\mathbf{R}+d\mathbf{I}_N)^{-1} \geq \frac{1}{(a+b)(c+d)}.$$

Proof: Let $\mathbf{R}=\mathbf{U}\boldsymbol{\Delta}\mathbf{U}^{\mathsf{H}}$, where $\mathbf{U}\in\mathbb{C}^{N\times N}$ is unitary and $\boldsymbol{\Delta}=\mathrm{diag}\,(\delta_1,\ldots,\delta_N)\geq 1$. Thus,

$$\begin{aligned}\frac{1}{N}\mathrm{tr}\,\mathbf{R}\,(a\mathbf{R}+b\mathbf{I}_N)^{-1}\,\mathbf{R}\,(c\mathbf{R}+d\mathbf{I}_N)^{-1} &= \frac{1}{N}\mathrm{tr}\,\boldsymbol{\Delta}^2\,(a\boldsymbol{\Delta}+b\mathbf{I}_N)^{-1}\,(c\boldsymbol{\Delta}+d\mathbf{I}_N)^{-1}\\ &= \frac{1}{(a+b)(c+d)}\frac{1}{N}\sum_{i=1}^{N}\frac{\delta_i^2(a+b)(c+d)}{(a\delta_i+b)(c\delta_i+d)}\\ &\geq \frac{1}{(a+b)(c+d)}. \end{aligned} \tag{27}$$

■

Lemma 6: Let the matrices $\mathbf{R}_k(\boldsymbol{\beta}_{k-1})$, be defined as in (10). Then, almost surely:

$$\limsup_n \|\mathbf{R}_k(\boldsymbol{\beta}_{k-1})\| < \infty, \qquad k=0,\ldots,K.$$

Proof: For $k \in \{1, \ldots, K\}$, denote by $(\Omega_k, \mathcal{F}_k, P_k)$ the probability space generating the sequences of random matrices $\mathbf{H}_k$. By [46], we have on a space $B_k \subset \Omega_k$ with $P_k(B_k) = 1$,

$$\frac{1}{n_{k-1}} \left\| \mathbf{H}_k \mathbf{H}_k^{\mathsf{H}} \right\| - \left(1 + \frac{1}{\sqrt{c_k}}\right)^2 \xrightarrow[n \to \infty]{} 0. \tag{28}$$

Obviously, we have $\|\mathbf{R}_0\| = \|\mathbf{I}_n\| = 1$. Thus, almost surely,

$$\limsup_n \|\mathbf{R}_1(\beta_0)\| \leq 1 + \limsup_n \frac{\alpha_1 \beta_0}{n} \left\| \mathbf{H}_1 \mathbf{H}_1^{\mathsf{H}} \right\| = 1 + \alpha_1 \beta_0 \limsup_n \left(1 + \frac{1}{\sqrt{c_1}}\right)^2 < \infty. \tag{29}$$

Consider now the product space $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \times \mathcal{F}_2, Q_2)$. By the Fubini theorem, we have on a subspace $C_2 \subset \Omega_1 \times \Omega_2$ with $Q_2(C_2) = 1$,

$$\begin{aligned}
\limsup_n \|\mathbf{R}_2(\boldsymbol{\beta}_1)\| &\leq 1 + \limsup_n \frac{\alpha_2 \beta_1}{n_1} \left\| \mathbf{H}_2 \mathbf{R}_1(\beta_0) \mathbf{H}_2^{\mathsf{H}} \right\| \\
&\leq 1 + \limsup_n \alpha_2 \beta_1 \|\mathbf{R}_1(\beta_0)\| \frac{1}{n_1} \left\| \mathbf{H}_2 \mathbf{H}_2^{\mathsf{H}} \right\| \\
&= 1 + \alpha_2 \beta_1 \limsup_n \left(1 + \alpha_1 \beta_0 \left(1 + \frac{1}{\sqrt{c_1}}\right)^2\right) \left(1 + \frac{1}{\sqrt{c_2}}\right)^2 \\
&< \infty.
\end{aligned} \tag{30}$$

Repeating the last step $k - 2$ times concludes the proof. ■

Definition 2 (Standard interference function [48]): A function $\mathbf{h}(x_1, \ldots, x_K) = [h_1(x_1, \ldots, x_K), \ldots, h_K(x_1, \ldots, x_K)]^{\mathsf{T}} \in \mathbb{R}^K$ is said to be standard if it fulfills the following conditions:

1) *Positivity:* for each j, if $x_1, \ldots, x_K \geq 0$, then $h_j(x_1, \ldots, x_K) > 0$.
2) *Monotonicity:* if $x_1 > x_1', \ldots, x_K > x_K'$, then for all j, $h_j(x_1, \ldots, x_K) > h_j(x_1', \ldots, x_K')$.
3) *Scalability:* for all $\alpha > 1$ and for all j, $\alpha h_j(x_1, \ldots, x_K) > h_j(\alpha x_1, \ldots, \alpha x_K)$.

Theorem 8 (Fixed-point theorem [48, Theorem 2]): If a K-variate function $\mathbf{h}(x_1, \ldots, x_K)$ is standard and there exists $(x_1, \ldots, x_K)$ such that for all j, $x_j \geq h_j(x_1, \ldots, x_K)$, then the fixed-point algorithm that consists in setting

$$x_j^{(t+1)} = h_j(x_1^{(t)}, \ldots, x_K^{(t)})$$

for $t \geq 1$ and for any initial values $x_1^{(0)}, \ldots, x_K^{(0)} > 0$ converges to the unique jointly positive solution of the system of K equations

$$x_j = h_j(x_1, \ldots, x_K), \qquad j \in \{1, \ldots, K\}.$$

Theorem 9 ([14, Corollary 1, Theorem 2]): For $k \in \{1,\ldots,K\}$, let $(n_k)_{N\geq 1} = (n_k(N))_{N\geq 1}$ be a sequence of positive integers and let $(\mathbf{R}_{k,N})_{N\geq 1}$, $\mathbf{R}_{k,N} \in \mathbb{C}^{N\times N}$, $(\mathbf{T}_{k,N})_{N\geq 1}$,$\mathbf{T}_{k,N} \in \mathbb{C}^{n_k\times n_k}$, and $(\mathbf{D}_N)_{N\geq 1}$, $\mathbf{D}_N \in \mathbb{C}^{N\times N}$, be three sequences of nonnegative definite Hermitian matrices, satisfying $\limsup_N \|\mathbf{R}_{k,N}\| < \infty$, $\limsup_N \|\mathbf{T}_{k,N}\| < \infty$, and $\limsup_N \|\mathbf{D}_N\| < \infty$. Let $(\mathbf{X}_{k,N})_{N\geq 1}$, $\mathbf{X}_{k,N} \in \mathbb{C}^{N\times n_k}$, be a sequence of random matrices with i.i.d. complex Gaussian entries with zero mean and variance $1/n_k$. Denote $\mathbf{B}_N = \sum_k \mathbf{R}_{k,N}^{\frac{1}{2}}\mathbf{X}_{k,N}\mathbf{T}_{k,N}\mathbf{X}_{k,N}^{\mathsf{H}}\mathbf{R}_{k,N}^{\frac{1}{2}}$ and define the function $V_N(x) = \frac{1}{N}\log\det\left(\mathbf{I}_N + x\mathbf{B}_N\right)$ for $x > 0$. Let $c_k = n_k/N$ and assume that $0 < \liminf_N c_k \leq \limsup_N c_k < \infty$ for all k. Then,

$$(i) \quad \frac{1}{N}\mathrm{tr}\,\mathbf{D}_N\left(\mathbf{B}_N + \frac{1}{x}\mathbf{I}_N\right)^{-1} - \frac{1}{N}\mathrm{tr}\,\mathbf{D}_N\left(\sum_{i=1}^{K}\bar{e}_i\mathbf{R}_{i,N} + \frac{1}{x}\mathbf{I}_N\right)^{-1} \xrightarrow[N\to\infty]{\text{a.s.}} 0$$

$$(ii) \quad V_N(x) - \bar{V}_N(x) \xrightarrow[N\to\infty]{\text{a.s.}} 0$$

where

$$\bar{V}_N(x) = \frac{1}{N}\log\det\left(\mathbf{I}_N + x\sum_{k=1}^{K}\bar{e}_{k,N}\mathbf{R}_{k,N}\right) + \sum_{k=1}^{K}\frac{1}{N}\log\det\left(\mathbf{I}_{n_k} + e_{k,N}\mathbf{T}_{k,N}\right) - \frac{1}{N}\sum_{k=1}^{K}n_k e_{k,N}\bar{e}_{k,N}$$

and where $\bar{e}_{k,N}, e_{k,N},\ k = 1,\ldots,K$, are given as the unique solution to the equations

$$\bar{e}_{k,N} = \frac{1}{n_k}\mathrm{tr}\,\mathbf{T}_{k,N}\left(e_{k,N}\mathbf{T}_{k,N} + \mathbf{I}_{n_k}\right)^{-1}$$

$$e_{k,N} = \frac{1}{n_K}\mathrm{tr}\,\mathbf{R}_{k,N}\left(\sum_{i=1}^{K}\bar{e}_i\mathbf{R}_{i,N} + \frac{1}{x}\mathbf{I}_N\right)^{-1}$$

such that $\bar{e}_{k,N}, e_{k,N} > 0$ for all k.

Corollary 4 (Special case of Theorem 9, see also [47]): Let $(n)_{N\geq 1} = (n(N))_{N\geq 1}$ be a sequence of positive integers. Let $(\mathbf{R}_N)_{N\geq 1}$, $\mathbf{R}_N \in \mathbb{C}^{n\times n}$, be a sequence of nonnegative definite Hermitian matrices, satisfying $\limsup_N \|\mathbf{R}_N\| < \infty$ and let $(\mathbf{X}_N)_{N\geq 1}$, $\mathbf{X}_N \in \mathbb{C}^{N\times n}$, be a sequence of random matrices with i.i.d. complex Gaussian entries with zero mean and variance $1/n$. For $x > 0$, define the following functions $m_N(x) = \frac{1}{N}\mathrm{tr}\left(\mathbf{X}_N\mathbf{R}_N\mathbf{X}_N^{\mathsf{H}} + \frac{1}{x}\mathbf{I}_N\right)^{-1}$ and $J_N(x) = \frac{1}{N}\log\det\left(\mathbf{I}_N + x\mathbf{X}_N\mathbf{R}_N\mathbf{X}_N^{\mathsf{H}}\right)$. Denote $c = \frac{n}{N}$ and assume that $0 < \liminf_N c \leq \limsup_N c < \infty$. Then,

$$(i) \quad m_N(x) - \bar{m}_N(x) \xrightarrow[N\to\infty]{\text{a.s.}} 0, \qquad (ii) \quad J_N(x) - \bar{J}_N(x) \xrightarrow[N\to\infty]{\text{a.s.}} 0$$

where

$$\bar{m}_N(x) = \frac{1}{N}\mathrm{tr}\left(\frac{\mathbf{R}_N}{c + \bar{e}_N} + \frac{1}{x}\mathbf{I}_n\right)^{-1} + (1-c)x$$

$$\bar{J}_N(x) = \frac{1}{N}\log\det\left(\left[c + \bar{e}_N\right]\mathbf{I}_n + x\mathbf{R}_N\right) + (1-c)\log\left(c + \bar{e}_N\right) - \frac{\bar{e}_N}{c + \bar{e}_N} - \log(c)$$

and $\bar{e}_N$ is defined as the unique positive solution to the implicit equation

$$\bar{e}_N = \frac{1}{N}\mathrm{tr}\,\mathbf{R}_N\left(\frac{\mathbf{R}_N}{c+\bar{e}_N}+\frac{1}{x}\mathbf{I}_n\right)^{-1}. \tag{31}$$

APPENDIX B
PROOF OF THEOREM 2:

First, notice that

$$\begin{aligned}
\eta_k &\triangleq \left\|\mathbf{R}_k(\boldsymbol{\beta}_{k-1})-\mathbf{R}_k(\bar{\boldsymbol{\beta}}_{k-1})\right\| \\
&= \left\|\frac{\alpha_k\beta_{k-1}}{n_{k-1}}\mathbf{H}_k\mathbf{R}_{k-1}(\boldsymbol{\beta}_{k-2})\mathbf{H}_k^{\mathsf{H}}-\frac{\alpha_k\bar{\beta}_{k-1}}{n_{k-1}}\mathbf{H}_k\mathbf{R}_{k-1}(\bar{\boldsymbol{\beta}}_{k-2})\mathbf{H}_k^{\mathsf{H}}\right\| \\
&\leq \alpha_k\left\|\frac{\mathbf{H}_k\mathbf{H}_k^{\mathsf{H}}}{n_{k-1}}\right\|\left\|\beta_{k-1}\mathbf{R}_{k-1}(\boldsymbol{\beta}_{k-2})-\bar{\beta}_{k-1}\mathbf{R}_{k-1}(\bar{\boldsymbol{\beta}}_{k-2})\right\| \\
&\leq \alpha_k\left\|\frac{\mathbf{H}_k\mathbf{H}_k^{\mathsf{H}}}{n_{k-1}}\right\|\left[|\beta_{k-1}-\bar{\beta}_{k-1}|\left\|\mathbf{R}_{k-1}(\boldsymbol{\beta}_{k-2})\right\|+\bar{\beta}_{k-1}\left\|\mathbf{R}_{k-1}(\boldsymbol{\beta}_{k-2})-\mathbf{R}_{k-1}(\bar{\boldsymbol{\beta}}_{k-2})\right\|\right].
\end{aligned} \tag{32}$$

Since, almost surely, $\limsup\left\|\mathbf{R}_1(\boldsymbol{\beta}_0)-\mathbf{R}_1(\bar{\boldsymbol{\beta}}_0)\right\|\leq\limsup\alpha_1|\beta_0-\bar{\beta}_0|\left\|\frac{\mathbf{H}_1\mathbf{H}_1^{\mathsf{H}}}{n}\right\|=0$ and $\limsup\left\|\mathbf{R}_{k-1}(\boldsymbol{\beta}_{k-2})\right\|<\infty$ (see proof of Lemma 6), one can iteratively show that $\eta_k\to 0$, almost surely. Thus,

$$\left|\mathcal{J}_k\left(x,\boldsymbol{\beta}_{k-1}\right)-\mathcal{J}_k\left(x,\bar{\boldsymbol{\beta}}_{k-1}\right)\right|\xrightarrow[n\to\infty]{\text{a.s.}}0. \tag{33}$$

This means that we can from now on replace $\boldsymbol{\beta}_k$ by $\bar{\boldsymbol{\beta}}_k$ and focus on $\mathcal{J}_k\left(x,\bar{\boldsymbol{\beta}}_{k-1}\right)$.

As a consequence of Corollary 4, Lemma 6, and the Fubini theorem, we obtain the following relation

$$\mathcal{J}_k(x,\bar{\boldsymbol{\beta}}_{k-1})-\tilde{\mathcal{J}}_k(x,\bar{\boldsymbol{\beta}}_{k-1})\xrightarrow[n\to\infty]{\text{a.s.}}0,\qquad k\geq 1 \tag{34}$$

where

$$\begin{aligned}
\tilde{\mathcal{J}}_k\left(x,\bar{\boldsymbol{\beta}}_{k-1}\right) &= \frac{1}{n_k}\log\det\left(\left[c_k+e_{k-1}\left(x,\bar{\boldsymbol{\beta}}_{k-1}\right)\right]\mathbf{I}_{n_{k-1}}+x\alpha_k\bar{\beta}_{k-1}\mathbf{R}_{k-1}\left(\bar{\boldsymbol{\beta}}_{k-2}\right)\right) \\
&\quad+(1-c_k)\log\left(c_k+e_{k-1}\left(x,\bar{\boldsymbol{\beta}}_{k-1}\right)\right)-\frac{e_{k-1}\left(x,\bar{\boldsymbol{\beta}}_{k-1}\right)}{c_k+e_{k-1}\left(x,\bar{\boldsymbol{\beta}}_{k-1}\right)}-\log\left(c_k\right)
\end{aligned} \tag{35}$$

and $e_{k-1}\left(x,\bar{\boldsymbol{\beta}}_{k-1}\right)$ is given as the unique positive solution to

$$e_{k-1}\left(x,\bar{\boldsymbol{\beta}}_{k-1}\right)=\frac{1}{n_k}\mathrm{tr}\,\alpha_k\bar{\beta}_{k-1}\mathbf{R}_{k-1}\left(\bar{\boldsymbol{\beta}}_{k-2}\right)\left(\frac{\alpha_k\bar{\beta}_{k-1}\mathbf{R}_{k-1}\left(\bar{\boldsymbol{\beta}}_{k-2}\right)}{c_k+e_{k-1}\left(x,\bar{\boldsymbol{\beta}}_{k-1}\right)}+\frac{1}{x}\mathbf{I}_{n_{k-1}}\right)^{-1}. \tag{36}$$

In particular, for $k=1$, we have

$$\tilde{\mathcal{J}}_1\left(x,\bar{\beta}_0\right)=\bar{\mathcal{J}}_1\left(x,\bar{\beta}_0\right) \tag{37}$$

where

$$\bar{\mathcal{J}}_1\left(x,\bar{\beta}_0\right)=c_1\log\left(1+\frac{x\alpha_1\bar{\beta}_0}{c_1+\bar{e}_0\left(x,\bar{\beta}_0\right)}\right)+\log\left(1+\frac{\bar{e}_0\left(x,\bar{\beta}_0\right)}{c_1}\right)-\frac{\bar{e}_0\left(x,\bar{\beta}_0\right)}{c_1+\bar{e}_0\left(x,\bar{\beta}_0\right)} \tag{38}$$

$$\bar{e}_0\left(x,\bar{\beta}_0\right)=-\frac{x\alpha_1\bar{\beta}_0(1-c_1)+c_1}{2}+\frac{\sqrt{\left(x\alpha_1\bar{\beta}_0(1-c_1)+c_1\right)^2+4x\alpha_1\bar{\beta}_0c_1^2}}{2} \tag{39}$$

according to Corollary 4. Note that (31) permits a closed-form solution (39) in this case.

Now, using the recursion $\mathbf{R}_{k-1}\left(\bar{\boldsymbol{\beta}}_{k-2}\right)=\mathbf{I}_{n_{k-1}}+\frac{\alpha_{k-1}\bar{\beta}_{k-2}}{n_{k-2}}\mathbf{H}_{k-1}\mathbf{R}_{k-2}\left(\bar{\boldsymbol{\beta}}_{k-3}\right)\mathbf{H}_{k-1}^{\mathsf{H}}$ (10) in (35), we obtain

$$\begin{aligned}\tilde{\mathcal{J}}_k(x,\bar{\boldsymbol{\beta}}_{k-1})=c_k\mathcal{J}_{k-1}&\left(\frac{x\alpha_k\bar{\beta}_{k-1}}{c_k+x\alpha_k\bar{\beta}_{k-1}+e_{k-1}\left(x,\bar{\boldsymbol{\beta}}_{k-1}\right)},\bar{\boldsymbol{\beta}}_{k-2}\right)+c_k\log\left(1+\frac{x\alpha_k\bar{\beta}_{k-1}}{c_k+e_{k-1}\left(x,\bar{\boldsymbol{\beta}}_{k-1}\right)}\right)\\&+\log\left(1+\frac{e_{k-1}\left(x,\bar{\boldsymbol{\beta}}_{k-1}\right)}{c_k}\right)-\frac{e_{k-1}\left(x,\bar{\boldsymbol{\beta}}_{k-1}\right)}{c_k+e_{k-1}\left(x,\bar{\boldsymbol{\beta}}_{k-1}\right)}.\end{aligned} \tag{40}$$

In the proof of Theorem 3, it is shown that

$$e_{k-1}\left(x,\bar{\boldsymbol{\beta}}_{k-1}\right)-\bar{e}_{k-1}\left(x,\bar{\boldsymbol{\beta}}_{k-1}\right)\xrightarrow[n\to\infty]{\text{a.s.}}0. \tag{41}$$

By the continuous mapping theorem [31], we therefore have

$$\begin{aligned}&\mathcal{J}_{k-1}\left(\frac{x\alpha_k\bar{\beta}_{k-1}}{1+x\alpha_k\bar{\beta}_{k-1}+e_{k-1}\left(x,\bar{\boldsymbol{\beta}}_{k-1}\right)},\bar{\boldsymbol{\beta}}_{k-2}\right)\\&\qquad-\mathcal{J}_{k-1}\left(\frac{x\alpha_k\bar{\beta}_{k-1}}{1+x\alpha_k\bar{\beta}_{k-1}+\bar{e}_{k-1}\left(x,\bar{\boldsymbol{\beta}}_{k-1}\right)},\bar{\boldsymbol{\beta}}_{k-2}\right)\xrightarrow[n\to\infty]{\text{a.s.}}0.\end{aligned} \tag{42}$$

Applying the last result together with Corollary 4, Lemma 6, the continuous mapping theorem and the Fubini theorem to (40) concludes the proof for $k=2$ since $\tilde{\mathcal{J}}_1\left(x,\bar{\beta}_0\right)=\bar{\mathcal{J}}_1\left(x,\bar{\beta}_0\right)$ by (37). The convergence for $k>2$ is shown by successive iterations of the last steps.

Appendix C
Proof of Theorem 3

From standard matrix inequalities and (32), it follows that

$$\left|m_k\left(x,\boldsymbol{\beta}_k\right)-m_k\left(x,\bar{\boldsymbol{\beta}}_k\right)\right|\leq x^2\alpha_{k+1}\left\|\frac{\mathbf{H}_{k+1}\mathbf{H}_{k+1}^{\mathsf{H}}}{n_k}\right\|\left\|\beta_k\mathbf{R}_k\left(\boldsymbol{\beta}_{k-1}\right)-\bar{\beta}_k\mathbf{R}_k\left(\bar{\boldsymbol{\beta}}_{k-1}\right)\right\|\xrightarrow[n\to\infty]{\text{a.s.}}0. \tag{43}$$

Thus, we can replace from now on β_k by $\bar{\beta}_k$, for almost every $(\mathbf{H}_1,\ldots,\mathbf{H}_K)$.

From Corollary 4, Lemma 6 and the Fubini theorem, it follows that

$$m_k\left(x,\bar{\boldsymbol{\beta}}_k\right)-\tilde{m}_k\left(x,\bar{\boldsymbol{\beta}}_k\right)\xrightarrow[n\to\infty]{\text{a.s.}}0 \tag{44}$$

where

$$\tilde{m}_k(x,\bar{\boldsymbol{\beta}}_k) = \frac{1}{n_{k+1}}\mathrm{tr}\left(\frac{\alpha_{k+1}\bar{\beta}_k\mathbf{R}_k\left(\bar{\boldsymbol{\beta}}_{k-1}\right)}{c_{k+1}+e_k\left(x,\bar{\boldsymbol{\beta}}_k\right)}+\frac{1}{x}\mathbf{I}_n\right)^{-1}+(1-c_{k+1})x \tag{45}$$

and $e_k\left(x,\bar{\boldsymbol{\beta}}_k\right)$ is given as the unique positive solution to

$$e_k\left(x,\bar{\boldsymbol{\beta}}_k\right) = \frac{1}{n_{k+1}}\mathrm{tr}\,\alpha_{k+1}\bar{\beta}_k\mathbf{R}_k\left(\bar{\boldsymbol{\beta}}_{k-1}\right)\left(\frac{\alpha_{k+1}\bar{\beta}_k\mathbf{R}_k\left(\bar{\boldsymbol{\beta}}_{k-1}\right)}{c_{k+1}+e_k\left(x,\bar{\boldsymbol{\beta}}_k\right)}+\frac{1}{x}\mathbf{I}_{n_k}\right)^{-1}. \tag{46}$$

In particular, we have $\tilde{m}_0(x,\bar{\boldsymbol{\beta}}_k)=\bar{m}_0(x,\bar{\boldsymbol{\beta}}_k)$, where

$$\bar{m}_0(x,\bar{\beta}_0) = \frac{c_1}{\frac{\alpha_1\bar{\beta}_0}{c_1+\bar{e}_0\left(x,\bar{\beta}_0\right)}+\frac{1}{x}}+(1-c_1)x \tag{47}$$

$$\bar{e}_0\left(x,\bar{\beta}_0\right) = -\frac{x\alpha_1\bar{\beta}_0(1-c_1)+c_1}{2}+\frac{\sqrt{\left(x\alpha_1\bar{\beta}_0(1-c_1)+c_1\right)^2+4x\alpha_1\bar{\beta}_0c_1^2}}{2}. \tag{48}$$

Replacing $\mathbf{R}_k\left(\bar{\boldsymbol{\beta}}_{k-1}\right)$ in (46) by its recursive definition $\mathbf{R}_k\left(\bar{\boldsymbol{\beta}}_{k-1}\right)=\mathbf{I}_{n_k}+\frac{\alpha_k\bar{\beta}_{k-1}}{n_{k-1}}\mathbf{H}_k\mathbf{R}_{k-1}\left(\bar{\boldsymbol{\beta}}_{k-2}\right)\mathbf{H}_k^{\mathsf{H}}$ (10) yields after straightforward calculus

$$\begin{aligned} e_k\left(x,\bar{\boldsymbol{\beta}}_k\right) &= c_{k+1}\left(c_{k+1}+e_k\left(x,\bar{\boldsymbol{\beta}}_k\right)\right) \\ &\quad -\frac{c_{k+1}\left(c_{k+1}+e_k\left(x,\bar{\boldsymbol{\beta}}_k\right)\right)^2}{x\alpha_{k+1}\bar{\beta}_k}m_{k-1}\left(\frac{x\alpha_{k+1}\bar{\beta}_k}{c_{k+1}+x\alpha_{k+1}\bar{\beta}_k+e_k\left(x,\bar{\boldsymbol{\beta}}_k\right)},\bar{\boldsymbol{\beta}}_{k-1}\right). \end{aligned} \tag{49}$$

Similarly, one obtains

$$\tilde{m}_k(x,\bar{\boldsymbol{\beta}}_k) = \frac{c_{k+1}\left(c_{k+1}+e_k\left(x,\bar{\boldsymbol{\beta}}_k\right)\right)}{\alpha_{k+1}\bar{\beta}_k}m_{k-1}\left(\frac{x\alpha_{k+1}\bar{\beta}_k}{c_{k+1}+x\alpha_{k+1}\bar{\beta}_k+e_k\left(x,\bar{\boldsymbol{\beta}}_k\right)},\bar{\boldsymbol{\beta}}_{k-1}\right)+(1-c_{k+1})x. \tag{50}$$

Combining the last two equations leads to

$$\tilde{m}_k(x,\bar{\boldsymbol{\beta}}_k) = \frac{xc_{k+1}}{c_{k+1}+e_k\left(x,\bar{\boldsymbol{\beta}}_k\right)}. \tag{51}$$

Consider now the quantity $\bar{e}_k\left(x,\bar{\boldsymbol{\beta}}_k\right)$, $k\geq 1$, defined as a positive solution to

$$\begin{aligned} \bar{e}_k\left(x,\bar{\boldsymbol{\beta}}_k\right) &= c_{k+1}\left(c_{k+1}+\bar{e}_k\left(x,\bar{\boldsymbol{\beta}}_k\right)\right) \\ &\quad -\frac{c_{k+1}\left(c_{k+1}+\bar{e}_k\left(x,\bar{\boldsymbol{\beta}}_k\right)\right)^2}{x\alpha_{k+1}\bar{\beta}_k}\bar{m}_{k-1}\left(\frac{x\alpha_{k+1}\bar{\beta}_k}{c_{k+1}+x\alpha_{k+1}\bar{\beta}_k+\bar{e}_k\left(x,\bar{\boldsymbol{\beta}}_k\right)},\bar{\boldsymbol{\beta}}_{k-1}\right) \end{aligned} \tag{52}$$

where $\bar{m}_k\left(x,\bar{\boldsymbol{\beta}}_k\right)$ is recursively defined for $k\geq 1$ as

$$\bar{m}_k\left(x,\bar{\boldsymbol{\beta}}_k\right) = \frac{xc_{k+1}}{c_{k+1}+\bar{e}_k\left(x,\bar{\boldsymbol{\beta}}_k\right)}. \tag{53}$$

It remains to show that a unique solution to (52) exists and that $e_k\left(x,\bar{\boldsymbol{\beta}}_k\right)-\bar{e}_k\left(x,\bar{\boldsymbol{\beta}}_k\right)\xrightarrow[n\to\infty]{\text{a.s.}}0$. Let us first define the following functions for $k\geq 1$:

$$f_k(z)=c_{k+1}\left(c_{k+1}+z\right)-\frac{c_{k+1}\left(c_{k+1}+z\right)^2}{x\alpha_{k+1}\bar{\beta}_k}m_{k-1}\left(\frac{x\alpha_{k+1}\bar{\beta}_k}{c_{k+1}+x\alpha_{k+1}\bar{\beta}_k+z},\bar{\boldsymbol{\beta}}_{k-1}\right) \tag{54}$$

$$\bar{f}_k(z)=c_{k+1}\left(c_{k+1}+z\right)-\frac{c_{k+1}\left(c_{k+1}+z\right)^2}{x\alpha_{k+1}\bar{\beta}_k}\bar{m}_{k-1}\left(\frac{x\alpha_{k+1}\bar{\beta}_k}{c_{k+1}+x\alpha_{k+1}\bar{\beta}_k+z},\bar{\boldsymbol{\beta}}_{k-1}\right). \tag{55}$$

From (46) and with the help of Lemma 5 (note that the smallest eigenvalue of $\mathbf{R}_k$ is greater or equal to 1 for all k), one can easily verify that $f_k(z)$ satisfies the following properties for $z\geq 0$:

(i)

$$f_k(z)\geq c_{k+1}(c_{k+1}+z)\left[1-\frac{c_{k+1}+z}{c_{k+1}+z+x\alpha_{k+1}\bar{\beta}_k}\right]>0$$

(ii) for $z>z'\geq 0$,

$$\begin{aligned}
&f_k(z)-f_k(z')\\
&\geq\frac{(z-z')c_{k+1}\alpha_{k+1}^2\bar{\beta}_k^2}{(c_{k+1}+z')(c_{k+1}+z)}\frac{1}{n_k}\operatorname{tr}\mathbf{R}_k\left(\frac{\alpha_{k+1}\bar{\beta}_k\mathbf{R}_k}{c_{k+1}+z}+\frac{1}{x}\mathbf{I}_{n_k}\right)^{-1}\mathbf{R}_k\left(\frac{\alpha_{k+1}\bar{\beta}_k\mathbf{R}_k}{c_{k+1}+z'}+\frac{1}{x}\mathbf{I}_{n_k}\right)^{-1}\\
&\geq\frac{(z-z')c_{k+1}\alpha_{k+1}^2\bar{\beta}_k^2}{\left(\alpha_{k+1}\bar{\beta}_k+\frac{c_{k+1}+z'}{x}\right)\left(\alpha_{k+1}\bar{\beta}_k+\frac{c_{k+1}+z}{x}\right)}\\
&>0
\end{aligned}$$

(iii) for $\alpha>1$,

$$\begin{aligned}
&\alpha f_k(z)-f_k(\alpha z)\\
&\geq\frac{(\alpha-1)c_{k+1}^2\alpha_{k+1}^2\bar{\beta}_k^2}{(c_{k+1}+\alpha z)(\alpha c_{k+1}+\alpha z)}\frac{1}{n_k}\operatorname{tr}\mathbf{R}_k\left(\frac{\alpha_{k+1}\bar{\beta}_k\mathbf{R}_k}{\alpha c_{k+1}+\alpha z}+\frac{1}{\alpha x}\mathbf{I}_{n_k}\right)^{-1}\mathbf{R}_k\left(\frac{\alpha_{k+1}\bar{\beta}_k\mathbf{R}_k}{c_{k+1}+\alpha z}+\frac{1}{x}\mathbf{I}_{n_k}\right)^{-1}\\
&\quad+\frac{(\alpha-1)c_{k+1}\alpha_{k+1}\bar{\beta}_k}{\alpha x}\frac{1}{n_k}\operatorname{tr}\mathbf{R}_k\left(\frac{\alpha_{k+1}\bar{\beta}_k\mathbf{R}_k}{\alpha c_{k+1}+\alpha z}+\frac{1}{\alpha x}\mathbf{I}_{n_k}\right)^{-1}\left(\frac{\alpha_{k+1}\bar{\beta}_k\mathbf{R}_k}{c_{k+1}+\alpha z}+\frac{1}{x}\mathbf{I}_{n_k}\right)^{-1}\\
&\geq\frac{(\alpha-1)c_{k+1}^2\alpha_{k+1}^2\bar{\beta}_k^2}{\left(\alpha_{k+1}\bar{\beta}_k+\frac{c_{k+1}+\alpha z}{x}\right)\left(\alpha_{k+1}\bar{\beta}_k+\frac{\alpha c_{k+1}+\alpha z}{x}\right)}+\frac{(\alpha-1)c_{k+1}\alpha_{k+1}\bar{\beta}_k}{\alpha x\left(\frac{\alpha_{k+1}\bar{\beta}_k}{\alpha c_{k+1}+\alpha z}+\frac{1}{\alpha x}\right)\left(\frac{\alpha_{k+1}\bar{\beta}_k}{c_{k+1}\alpha z}+\frac{1}{x}\right)}\\
&>0
\end{aligned}$$

where $\mathbf{R}_k=\mathbf{R}_k\left(\bar{\boldsymbol{\beta}}_{k-1}\right)$. All properties are independent of $\mathbf{R}_k$ and therefore hold for $n\to\infty$.

Assume now $k=1$. For any sequence of bounded non-negative real numbers z_n, we have by (44) and the continuous mapping theorem [31],

$$f_1(z_n)-\bar{f}_1(z_n)\xrightarrow[n\to\infty]{\text{a.s.}}0. \tag{56}$$

Thus, properties $(i)-(iii)$ of $f_1(z)$ also hold for $\bar{f}_1(z)$. By Definition 2 and Theorem 8, these properties imply the uniqueness of positive solutions to the fixed point equations $z = f_1(z)$ and $y = \bar{f}_1(y)$, and hence the uniqueness of solutions to (52) for $k = 1$. Moreover, note that

$$|f_k(a) - f_k(b)| \leq \frac{\alpha_{k+1}^2 \bar{\beta}_k^2 x^2}{c_{k+1}} \|\mathbf{R}_k\left(\bar{\boldsymbol{\beta}}_{k-1}\right)\|^2 |a-b|. \tag{57}$$

Hence,

$$\begin{aligned}\left|\bar{e}_1\left(x,\bar{\boldsymbol{\beta}}_1\right) - e_1\left(x,\bar{\boldsymbol{\beta}}_1\right)\right| &= \left|\bar{f}_1\left(\bar{e}_1\left(x,\bar{\boldsymbol{\beta}}_1\right)\right) - f_1\left(e_1\left(x,\bar{\boldsymbol{\beta}}_1\right)\right)\right| \\ &\leq \left|\bar{f}_1\left(\bar{e}_1\left(x,\bar{\boldsymbol{\beta}}_1\right)\right) - f_1\left(\bar{e}_1\left(x,\bar{\boldsymbol{\beta}}_1\right)\right)\right| + \left|f_1\left(\bar{e}_1\left(x,\bar{\boldsymbol{\beta}}_1\right)\right) - f_1\left(e_1\left(x,\bar{\boldsymbol{\beta}}_1\right)\right)\right| \\ &\leq \epsilon_n + \frac{\alpha_2^2 \bar{\beta}_1^2 x^2}{c_2} \|\mathbf{R}_1\left(\bar{\beta}_0\right)\|^2 \left|\bar{e}_1\left(x,\bar{\boldsymbol{\beta}}_1\right) - e_1\left(x,\bar{\boldsymbol{\beta}}_1\right)\right| \end{aligned} \tag{58}$$

for some sequence of real numbers ϵ_n, satisfying $\epsilon_n \xrightarrow[n\to\infty]{\text{a.s.}} 0$. By Lemma 6, $\|\mathbf{R}_1\left(\bar{\beta}_0\right)\| < M$, almost surely, for some $M > 0$. Thus, for $x \leq \sqrt{\frac{c_2(1-\delta)}{\alpha_2^2 \bar{\beta}_1^2 M^2}}$ and some $\delta > 0$, we have

$$\left|\bar{e}_1\left(x,\bar{\boldsymbol{\beta}}_1\right) - e_1\left(x,\bar{\boldsymbol{\beta}}_1\right)\right| \leq \frac{\epsilon_n}{\delta} \xrightarrow[n\to\infty]{\text{a.s.}} 0. \tag{59}$$

Since $\bar{e}_1\left(x,\bar{\boldsymbol{\beta}}_1\right)$ and $e_1\left(x,\bar{\boldsymbol{\beta}}_1\right)$ are (almost surely) bounded on any closed subset of $\mathbb{R}^+ \setminus \{0\}$ and have analytic continuations for $x \in \mathbb{C}\setminus\mathbb{R}^-$, we have by Vitali's convergence theorem [49] that the convergence holds for any $x \in \mathbb{R}^+ \setminus \{0\}$.

The last convergence implies by the continuous mapping theorem that,

$$m_1(x,\bar{\boldsymbol{\beta}}_1) - \bar{m}_1\left(x,\bar{\boldsymbol{\beta}}_1\right) \xrightarrow[n\to\infty]{\text{a.s.}} 0. \tag{60}$$

We now assume $k = 2$. The last convergence implies $f_2(z) \to \bar{f}_2(z)$, almost surely. The same steps can therefore be applied to show that $m_2(x,\bar{\boldsymbol{\beta}}_1) - \bar{m}_2\left(x,\bar{\boldsymbol{\beta}}_1\right) \xrightarrow[n\to\infty]{\text{a.s.}} 0$. This terminates the proof as this process can be iterated k times.

Appendix D

Proof of Theorem 4: Fundamental equations

The proof follows essentially the same steps as the proof of Theorem 2 in [7] and will not be given in full detail here. In order to prove the uniqueness of solutions $(\bar{g}_k, g_k, \delta_k)$, it is sufficient to show by Theorem 8 that the K-variate function $\mathbf{h} : (x_1, \ldots, x_K) \mapsto (h_1, \ldots, h_K)$ as defined below, is a standard interference function (see Definition 2). For $k = 1, \ldots, K$, we define

$$h_k(x_1, \ldots, x_K) \mapsto \frac{1}{n_k} \sum_{j=1}^{N_k} \frac{s_{k,j}\delta_k}{1 + \bar{g}_k s_{k,j}\delta_k} \tag{61}$$

where

$$\bar{g}_k = \frac{1}{n_k}\mathrm{tr}\,\mathbf{T}_k^{\frac{1}{2}}\mathbf{Q}_k\mathbf{T}_k^{\frac{1}{2}}\left(x_k\mathbf{T}_k^{\frac{1}{2}}\mathbf{Q}_k\mathbf{T}_k^{\frac{1}{2}}+\mathbf{I}_{n_k}\right)^{-1} \tag{62}$$

and $\delta_k,\ k=1,\ldots,K$, form the unique jointly positive solution to the following fixed-point equations

$$\delta_k = \frac{1}{N_k}\mathrm{tr}\,\mathbf{R}_k\left(\sum_{k=1}^{K}\frac{n_k}{N_k}\frac{\bar{g}_k x_k}{\delta_k}\mathbf{R}_k+\frac{1}{\rho}\mathbf{I}_N\right)^{-1}. \tag{63}$$

The only difference to [7] is the definition of $\bar{g}_k$. In our case, $\bar{g}_k$ is directly defined as a function of x_k, whereas $\bar{b}_k$ in [7] (using their notations) is given as the solution of another fixed point equation. However, the behavior of $\bar{b}_k$ and $\bar{g}_k$ as seen as functions of x_k is identical. In particular, let $x_k > x_k' > 0$ and denote by $\bar{g}_k$ and $\bar{g}_k'$ the corresponding values of (62), respectively. One can easily verify that the following conditions hold:(i) $\bar{g}_k < \bar{g}_k'$ and (ii) $x_k\bar{g}_k > x_k'\bar{g}_k'$. The remaining steps are identical to [7] and will not be repeated here. By showing $\mathbf{h}(x_1,\ldots,x_K)$ to be a standard interference function, we have proved by Theorem 8 that the following fixed-point algorithm, which iteratively computes

$$x_k^{t+1} = h_k(x_1^{(t)},\ldots,x_K^{(t)}), \qquad k=1,\ldots,K \tag{64}$$

for $t\geq 0$ and some set of initial values $x_1^{(0)},\ldots,x_K^{(0)}$, converges as $t\to\infty$ to the unique fixed point $(g_1,\ldots,g_K)$.

Appendix E
Proof of Theorem 5: Mutual information

The key idea is that the double-scattering model can be considered as the Kronecker channel model as considered in Theorem 9 with *random* correlation matrices. Assume now a Kronecker model, for which the matrices $\mathbf{H}_k$ are given as

$$\mathbf{H}_k = \frac{1}{\sqrt{n_k}}\mathbf{Z}_k\mathbf{W}_{2,k}\mathbf{T}_k^{\frac{1}{2}} \tag{65}$$

where $\mathbf{Z}_k\in\mathbb{C}^{N\times N_k}$ is a *deterministic* matrix and $\mathbf{W}_{2,k}$ and $\mathbf{T}_k$ are defined as in (17). Further assume that $\limsup_N\|\mathbf{Z}_k\|<\infty$ for all k. Thus, we can apply Theorem 9 to obtain the following deterministic equivalent $\bar{V}_N(\rho)$ of $I_N(\rho)$:

$$\bar{V}_N(\rho) = \frac{1}{N}\log\det\left(\mathbf{I}_N+\rho\sum_{k=1}^{K}\bar{e}_k\mathbf{Z}_k\mathbf{Z}_k^{\mathsf{H}}\right)+\sum_{k=1}^{K}\frac{1}{N}\log\det\left(\mathbf{I}_{n_k}+e_k\mathbf{T}_k^{\frac{1}{2}}\mathbf{Q}_k\mathbf{T}_k^{\frac{1}{2}}\right)-\frac{1}{N}\sum_{k=1}^{K}n_k e_k\bar{e}_k \tag{66}$$

where $\bar{e}_k, e_k, k = 1, \ldots, K$, are given as the unique solutions to the following equations

$$\bar{e}_k = \frac{1}{n_k} \mathrm{tr}\, \mathbf{T}_k^{\frac{1}{2}} \mathbf{Q}_k \mathbf{T}_k^{\frac{1}{2}} \left(e_k \mathbf{T}_k^{\frac{1}{2}} \mathbf{Q}_k \mathbf{T}_k^{\frac{1}{2}} + \mathbf{I}_{n_k} \right)^{-1}$$
$$e_k = \frac{1}{n_k} \mathrm{tr}\, \mathbf{Z}_k \mathbf{Z}_k^{\mathsf{H}} \left(\sum_{i=1}^{K} \bar{e}_i \mathbf{Z}_i \mathbf{Z}_i^{\mathsf{H}} + \frac{1}{\rho} \mathbf{I}_N \right)^{-1} \tag{67}$$

such that $\bar{e}_k, e_k > 0$ for all k. Recall that the matrices $\mathbf{Q}_k$ are the covariance matrices of the channel inputs $\mathbf{x}_k$. Thus, the channel model is equivalent to a channel $\mathbf{H}_k = \frac{1}{\sqrt{n_k}} \mathbf{Z}_k \mathbf{W}_{2,k} \tilde{\mathbf{T}}_k^{\frac{1}{2}}$, where $\tilde{\mathbf{T}}_k^{\frac{1}{2}} = \mathbf{T}^{\frac{1}{2}} \mathbf{Q}_k^{\frac{1}{2}}$, with uncorrelated channel inputs $\tilde{\mathbf{x}}_k \sim \mathcal{CN}(0, \mathbf{I}_{n_k})$.

For the double-scattering channel, the matrices $\mathbf{Z}_k$ are *random* and defined as

$$\mathbf{Z}_k = \frac{1}{\sqrt{N_k}} \mathbf{R}_k^{\frac{1}{2}} \mathbf{W}_{1,k} \mathbf{S}_k^{\frac{1}{2}}. \tag{68}$$

Let $(\Omega, \mathcal{F}, P)$ be the probability space generating the random sequences of matrices $(\mathbf{W}_{1,k}(\omega))_{N \geq 1}$. There exists $A \subset \Omega$ with $P(A) = 1$, such that for each $\omega \in A$, we have $\limsup_N \|\mathbf{Z}_k(\omega) \mathbf{Z}_k(\omega)^{\mathsf{H}}\| < \infty$ ([46]). Thus, for each of these ω, the matrices $\mathbf{Z}_k \mathbf{Z}_k^{\mathsf{H}}$ satisfy the criteria of the correlation matrices of Theorem 9. Let $(\Omega', \mathcal{F}', P')$ the probability space generating the matrices $\mathbf{W}_{2,k}$. Thus, for every $\omega \in A$, there exist a $A'(\omega) \subset \Omega'$ with $P'(A') = 1$, such that for all $\omega' \in A'(\omega)$, $\bar{V}_N(\rho)$ is a deterministic equivalent of $I_N(\rho)$. Denote by $(\Omega \times \Omega', \mathcal{F} \times \mathcal{F}', Q)$ the product space generating the matrices $\mathbf{W}_{1,k}(\omega)$ and $\mathbf{W}_{1,k}(\omega')$ and denote by $B \subset \Omega \times \Omega'$ the space of all tuples (ω, ω'), such that $\omega \in A$ and $\omega' \in A'(\omega)$. By the Fubini theorem, we have $Q(B) = 1$, which proves that $\bar{V}_N(\rho) - I_n(\rho) \to 0$, almost surely. However, $\bar{V}_N(\rho)$ is a random quantity, which depends on the matrices $\mathbf{Z}_k$. Therefore, we will need to obtain an iterative deterministic equivalent $\bar{I}_N(\rho)$ of $\bar{V}_N(\rho)$.

The first step is to replace the fixed-point equations (67) that depend on $\mathbf{Z}_k$ by deterministic ones. Let us define the quantities $\bar{e}_{k,i,j},\ e_{k,i,j}$, for $i \in \{1, \ldots, K\},\ j \in \{1, \ldots, N_k\}$, which are given as the unique solutions to the following set of fixed-point equations:

$$\bar{e}_{k,i,j} = \frac{1}{n_k} \mathrm{tr}\, \tilde{\mathbf{T}}_k \left(e_{k,i,j} \tilde{\mathbf{T}}_k + \mathbf{I}_{n_k} \right)^{-1}$$
$$e_{k,i,j} = \frac{1}{n_k} \mathrm{tr}\, \mathbf{Z}_{k,i,j} \mathbf{Z}_{k,i,j}^{\mathsf{H}} \left(\sum_{\ell=1}^{K} \bar{e}_{\ell,i,j} \mathbf{Z}_{\ell,i,j} \mathbf{Z}_{\ell,i,j}^{\mathsf{H}} + \frac{1}{\rho} \mathbf{I}_N \right)^{-1} \tag{69}$$

where

$$\mathbf{Z}_{k,i,j} = \begin{cases} \mathbf{Z}_k & , \quad i \neq k \\ [\mathbf{z}_{k,1} \cdots \mathbf{z}_{k,j-1} \mathbf{z}_{k,j+1} \ldots \mathbf{z}_{k,N_k}] & , \quad i = k \end{cases}.$$

Obviously, $\bar{e}_{k,i,j}$ and $e_{k,i,j}$ are independent of the vector $\mathbf{z}_{i,j}$. In addition, we define

$$Z = \max_k \limsup_N \|\mathbf{Z}_k \mathbf{Z}_k^{\mathsf{H}}\|, \qquad T = \max_k \limsup_N \|\tilde{\mathbf{T}}_k\|, \qquad n = \min_k n_k, \qquad c = \frac{N}{n}$$

and

$$\alpha_{i,j} = \max_k |e_{k,i,j} - e_k|, \qquad \bar{\alpha}_{i,j} = \max_k |\bar{e}_{k,i,j} - \bar{e}_k|.$$

Thus, we have for N large,

$$\begin{aligned} |\bar{e}_{k,i,j} - \bar{e}_k| &= \left| \frac{1}{n_k} \mathrm{tr}\, \tilde{\mathbf{T}}_k \left(e_{k,i,j} \tilde{\mathbf{T}}_k + \mathbf{I}_{n_k} \right)^{-1} \left((e_k - e_{k,i,j}) \tilde{\mathbf{T}}_k \right) \left(e_k \tilde{\mathbf{T}}_k + \mathbf{I}_{n_k} \right)^{-1} \right| \\ &\leq \alpha_{i,j} T^2. \end{aligned} \tag{70}$$

Since the right-hand side (RHS) of the last inequality is independent of k, we have

$$\bar{\alpha}_{i,j} \leq \alpha_{i,j} T^2. \tag{71}$$

On the other hand, for $i \neq k$, we have for N sufficiently large,

$$\begin{aligned} &|e_{k,i,j} - e_k| \\ &= \left| \frac{1}{n_k} \mathrm{tr}\, \mathbf{Z}_k \mathbf{Z}_k^{\mathsf{H}} \left(\sum_{\ell=1}^K \bar{e}_{\ell,i,j} \mathbf{Z}_{\ell,i,j} \mathbf{Z}_{\ell,i,j}^{\mathsf{H}} + \frac{1}{\rho} \mathbf{I}_N \right)^{-1} \left(\sum_{\ell=1}^K (\bar{e}_\ell - \bar{e}_{\ell,i,j}) \mathbf{Z}_\ell \mathbf{Z}_\ell^{\mathsf{H}} + \bar{e}_{i,i,j} \mathbf{z}_{i,j} \mathbf{z}_{i,j}^{\mathsf{H}} \right) \left(\sum_{\ell=1}^K \bar{e}_\ell \mathbf{Z}_\ell \mathbf{Z}_\ell^{\mathsf{H}} + \frac{1}{\rho} \mathbf{I}_N \right)^{-1} \right| \\ &\leq cK\rho^2 Z^2 \bar{\alpha}_{i,j} + \frac{\bar{e}_{i,i,j}}{n} \left| \mathbf{z}_{i,j}^{\mathsf{H}} \left(\sum_{\ell=1}^K \bar{e}_\ell \mathbf{Z}_\ell \mathbf{Z}_\ell^{\mathsf{H}} + \frac{1}{\rho} \mathbf{I}_N \right)^{-1} \mathbf{Z}_k \mathbf{Z}_k^{\mathsf{H}} \left(\sum_{\ell=1}^K \bar{e}_{\ell,i,j} \mathbf{Z}_{\ell,i,j} \mathbf{Z}_{\ell,i,j}^{\mathsf{H}} + \frac{1}{\rho} \mathbf{I}_N \right)^{-1} \mathbf{z}_{i,j} \right| \\ &\leq cK\rho^2 Z^2 \bar{\alpha}_{i,j} + \frac{\rho^2 Z^2 T}{n} \end{aligned} \tag{72}$$

where the last inequality is due to $\bar{e}_{k,i,j} \leq T$ and $\left| \mathbf{z}_{i,j}^{\mathsf{H}} \mathbf{A} \mathbf{z}_{i,j} \right| \leq \|\mathbf{z}_{i,j}\|^2 \|\mathbf{A}\| \leq Z \|\mathbf{A}\|$, for any matrix $\mathbf{A}$.

Similarly, one can show that

$$|e_{k,k,j} - e_k| = cK\rho^2 Z^2 \bar{\alpha}_{k,j} + \frac{\rho^2 Z^2 T}{n} + \frac{\rho Z}{n}. \tag{73}$$

It follows from (72), (73) and (71), that

$$\alpha_{i,j} \leq cK\rho^2 Z^2 \bar{\alpha}_{i,j} + \frac{\rho^2 Z^2 T}{n} + \frac{\rho Z}{n} \leq cK\rho^2 Z^2 T^2 \alpha_{i,j} + \frac{\rho^2 Z^2 T}{n} + \frac{\rho Z}{n}. \tag{74}$$

Now, for any $\rho \leq \sqrt{\frac{1-\epsilon}{cKZ^2T^2}}$ and $\epsilon > 0$, we have

$$\alpha_{i,j} \leq \frac{\rho Z}{\epsilon n} (1 + \rho Z T) \qquad , \qquad \bar{\alpha}_{i,j} \leq \frac{\rho Z T^2}{\epsilon n} (1 + \rho Z T).$$

Let $\mu = \max\{\frac{\rho Z}{\epsilon}(1+\rho ZT), \frac{\rho ZT^2}{\epsilon}(1+\rho ZT)\}$, then we finally have

$$\alpha_{i,j} \leq \frac{\mu}{n} \qquad , \qquad \bar{\alpha}_{i,j} \leq \frac{\mu}{n}. \tag{75}$$

The last result establishes that for sufficiently small ρ, the differences between the solutions $(\bar{e}_{k,i,j}, e_{k,i,j})$ to (69) and the solutions $(\bar{e}_k, e_k)$ to (67) are uniformly bounded by μ and vanish as $n \to \infty$. Moreover, e_k and $e_{k,i,j}$ have an analytic continuation on $z = -\frac{1}{\rho} \in \mathbb{C} \setminus \mathbb{R}^+$ and are uniformly bounded on all closed subsets of $z \in \mathbb{C} \setminus \mathbb{R}^+$. Thus, in particular for $\rho \in \mathbb{R}^+ \setminus \{0\}$ and all k, i, j, we have by the Vitali convergence theorem [49]

$$e_k - e_{k,i,j} \xrightarrow[N\to\infty]{} 0 \qquad \text{and hence} \qquad \bar{e}_k - \bar{e}_{k,i,j} \xrightarrow[N\to\infty]{} 0. \tag{76}$$

As a consequence of (76), we can now write

$$\begin{aligned}
e_k &= \frac{1}{n_k}\sum_{j=1}^{N_k} \mathbf{z}_{k,j}^{\mathsf{H}} \left(\sum_{i=1}^{K} \bar{e}_i \mathbf{Z}_i \mathbf{Z}_i^{\mathsf{H}} + \frac{1}{\rho}\mathbf{I}_N\right)^{-1} \mathbf{z}_{k,j} \\
&\overset{(a)}{\asymp} \frac{1}{n_k}\sum_{j=1}^{N_k} \mathbf{z}_{k,j}^{\mathsf{H}} \left(\sum_{i=1}^{K} \bar{e}_{i,k,j} \mathbf{Z}_i \mathbf{Z}_i^{\mathsf{H}} + \frac{1}{\rho}\mathbf{I}_N\right)^{-1} \mathbf{z}_{k,j} \\
&\overset{(b)}{=} \frac{1}{n_k}\sum_{j=1}^{N_k} \frac{\mathbf{z}_{k,j}^{\mathsf{H}} \left(\sum_{i=1}^{K} \bar{e}_{i,k,j} \mathbf{Z}_i \mathbf{Z}_i^{\mathsf{H}} - \bar{e}_{k,k,j}\mathbf{z}_{k,j}\mathbf{z}_{k,j}^{\mathsf{H}} + \frac{1}{\rho}\mathbf{I}_N\right)^{-1} \mathbf{z}_{k,j}}{1 + \bar{e}_{k,k,j}\mathbf{z}_{k,j}^{\mathsf{H}} \left(\sum_{i=1}^{K} \bar{e}_{i,j,k} \mathbf{Z}_i \mathbf{Z}_i^{\mathsf{H}} - \bar{e}_{k,k,j}\mathbf{z}_{k,j}\mathbf{z}_{k,j}^{\mathsf{H}} + \frac{1}{\rho}\mathbf{I}_N\right)^{-1} \mathbf{z}_{k,j}} \\
&\overset{(c)}{\asymp} \frac{1}{n_k}\sum_{j=1}^{N_k} \frac{\frac{s_{k,j}}{N_k}\operatorname{tr}\mathbf{R}_k \left(\sum_{i=1}^{K} \bar{e}_{i,j,k} \mathbf{Z}_i \mathbf{Z}_i^{\mathsf{H}} + \frac{1}{\rho}\mathbf{I}_N\right)^{-1}}{1 + \frac{s_{k,j}\bar{e}_{k,k,j}}{N_k}\operatorname{tr}\mathbf{R}_k \left(\sum_{i=1}^{K} \bar{e}_{i,j,k} \mathbf{Z}_i \mathbf{Z}_i^{\mathsf{H}} + \frac{1}{\rho}\mathbf{I}_N\right)^{-1}} \\
&\overset{(d)}{\asymp} \frac{1}{n_k}\sum_{j=1}^{N_k} \frac{\frac{s_{k,j}}{N_k}\operatorname{tr}\mathbf{R}_k \left(\sum_{i=1}^{K} \bar{e}_{i} \mathbf{Z}_i \mathbf{Z}_i^{\mathsf{H}} + \frac{1}{\rho}\mathbf{I}_N\right)^{-1}}{1 + \frac{s_{k,j}\bar{e}_{k}}{N_k}\operatorname{tr}\mathbf{R}_k \left(\sum_{i=1}^{K} \bar{e}_{i} \mathbf{Z}_i \mathbf{Z}_i^{\mathsf{H}} + \frac{1}{\rho}\mathbf{I}_N\right)^{-1}}
\end{aligned} \tag{77}$$

where (a) follows from (76) since

$$\left|\mathbf{z}_{k,j}^{\mathsf{H}} \left(\sum_{i=1}^{K} \bar{e}_i \mathbf{Z}_i \mathbf{Z}_i^{\mathsf{H}} + \frac{1}{\rho}\mathbf{I}_N\right)^{-1} \mathbf{z}_{k,j} - \mathbf{z}_{k,j}^{\mathsf{H}} \left(\sum_{i=1}^{K} \bar{e}_{i,k,j} \mathbf{Z}_i \mathbf{Z}_i^{\mathsf{H}} + \frac{1}{\rho}\mathbf{I}_N\right)^{-1} \mathbf{z}_{k,j}\right| \leq \frac{\mu K Z \rho^2}{n} \xrightarrow[N\to\infty]{} 0, \tag{78}$$

(b) is due to Lemma 3, (c) is a consequence of Lemmas 2 and 4 and (d) is obtained by applying (76) a second time. Next, we would like to find deterministic equivalents of the terms $\frac{1}{N_k}\operatorname{tr}\mathbf{R}_k \left(\sum_{i=1}^{K} \bar{e}_i \mathbf{Z}_i \mathbf{Z}_i^{\mathsf{H}} + \frac{1}{\rho}\mathbf{I}_N\right)^{-1}$. We cannot directly apply Theorem 9 at this point since the $\bar{e}_k$ are defined as functions of $\mathbf{Z}_k$. However, based on the relations (75) and (78), Theorem 9 (see [14, Theorem 1]) can be shown to hold also for the matrix model under study. Thus,

$$\frac{1}{N_k}\operatorname{tr}\mathbf{R}_k \left(\sum_{i=1}^{K} \bar{e}_i \mathbf{Z}_i \mathbf{Z}_i^{\mathsf{H}} + \frac{1}{\rho}\mathbf{I}_N\right)^{-1} \asymp f_k \tag{79}$$

where f_k for $k \in \{1, \ldots, K\}$ are defined as the unique solution to the following fixed-point equations

$$\bar{f}_k = \frac{1}{N_k} \sum_{j=1}^{N_k} \frac{s_{k,j} \bar{e}_k}{1 + \bar{e}_k s_{k,j} f_k} \tag{80}$$

$$f_k = \frac{1}{N_k} \operatorname{tr} \mathbf{R}_k \left(\sum_{i=1}^{K} \bar{f}_i \mathbf{R}_i + \frac{1}{\rho} \mathbf{I}_N \right)^{-1} \tag{81}$$

such that $f_k > 0$ for all k. Replacing (80) in (81) leads to

$$f_k = \frac{1}{N_k} \operatorname{tr} \mathbf{R}_k \left(\sum_{i=1}^{K} \frac{\bar{e}_i}{N_i} \mathbf{R}_i \sum_{j=1}^{N_i} \frac{s_{i,j}}{1 + s_{i,j} \bar{e}_i f_i} + \frac{1}{\rho} \mathbf{I}_N \right)^{-1}. \tag{82}$$

Thus,

$$e_k = \frac{1}{n_k} \sum_{j=1}^{N_k} \frac{s_{k,j} f_k}{1 + s_{k,j} \bar{e}_k f_k} + \epsilon_k \tag{83}$$

where ϵ_k is a sequence of random variables, satisfying $\epsilon_k \xrightarrow[N \to \infty]{\text{a.s.}} 0$. Consider now the following system of equations

$$\begin{aligned}
\bar{e}_k &= \frac{1}{n_k} \operatorname{tr} \tilde{\mathbf{T}}_k \left(e_k \tilde{\mathbf{T}}_k + \mathbf{I}_{n_k} \right)^{-1} \\
e_k &= \frac{1}{n_k} \sum_{j=1}^{N_k} \frac{s_{k,j} f_k}{1 + s_{k,j} \bar{e}_k f_k} + \epsilon_k \\
f_k &= \frac{1}{N_k} \operatorname{tr} \mathbf{R}_k \left(\sum_{i=1}^{K} \frac{\bar{e}_i}{N_i} \mathbf{R}_i \sum_{j=1}^{N_i} \frac{s_{i,j}}{1 + s_{i,j} \bar{e}_i f_i} + \frac{1}{\rho} \mathbf{I}_N \right)^{-1}
\end{aligned}$$

and its deterministic counterpart

$$\begin{aligned}
\bar{g}_k &= \frac{1}{n_k} \operatorname{tr} \tilde{\mathbf{T}}_k \left(g_k \tilde{\mathbf{T}}_k + \mathbf{I}_{n_k} \right)^{-1} \\
g_k &= \frac{1}{n_k} \sum_{j=1}^{N_k} \frac{s_{k,j} \delta_k}{1 + s_{k,j} \bar{g}_k \delta_k} \\
\delta_k &= \frac{1}{N_k} \operatorname{tr} \mathbf{R}_k \left(\sum_{i=1}^{K} \frac{\bar{g}_i}{N_i} \mathbf{R}_i \sum_{j=1}^{N_i} \frac{s_{i,j}}{1 + s_{i,j} \bar{g}_i \delta_i} + \frac{1}{\rho} \mathbf{I}_N \right)^{-1} = \frac{1}{N_k} \operatorname{tr} \mathbf{R}_k \left(\sum_{i=1}^{K} \frac{n_i}{N_i} \frac{\bar{g}_i g_i}{\delta_i} \mathbf{R}_i + \frac{1}{\rho} \mathbf{I}_N \right)^{-1}.
\end{aligned}$$

Define the quantities:

$$\gamma_1 = \max_k |e_k - g_k|, \qquad \gamma_2 = \max_k |\bar{e}_k - \bar{g}_k|, \qquad \gamma_3 = \max_k |f_k - \delta_k|, \qquad \epsilon = \max_k |\epsilon_k|.$$

Straight-forward calculations lead to the following bounds:

$$\gamma_1 \leq c S \gamma_3 + c S^2 R^2 \rho^2 \gamma_2 + \epsilon, \qquad \gamma_2 \leq \gamma_1 T^2, \qquad \gamma_3 \leq K S R^2 \rho^2 \gamma_2 + K S^2 R^2 T^2 \rho^2 \gamma_1.$$

Combining these results and using the fact that $\epsilon \xrightarrow[N\to\infty]{\text{a.s.}} 0$ yields for ρ sufficiently small,

$$\gamma_1, \gamma_2, \gamma_3 \xrightarrow[N\to\infty]{\text{a.s.}} 0. \tag{84}$$

Since $e_k, g_k, \bar{e}_k, \bar{g}_k, f_k, \delta_k$ are all (almost surely) bounded for ρ in any closed subset of $\mathbb{R}^+ \setminus \{0\}$ and have analytic continuations for $\rho \in \mathbb{C} \setminus \mathbb{R}^-$, we have by Vitali's convergence theorem [49] that (84) holds for any $\rho \in \mathbb{R}^+ \setminus \{0\}$.

Coming now back to $\bar{V}_N(\rho)$ as given in (66), we have from the continuous mapping theorem [31] that

$$\frac{1}{N}\sum_{k=1}^{K}\left[\log\det\left(\mathbf{I}_{n_k} + e_k\tilde{\mathbf{T}}_k\right) - n_k e_k \bar{e}_k\right] - \left[\log\det\left(\mathbf{I}_{n_k} + g_k\tilde{\mathbf{T}}_k\right) - n_k g_k \bar{g}_k\right] \xrightarrow[N\to\infty]{\text{a.s.}} 0. \tag{85}$$

Moreover, since $\|\sum_{k=1}^{K}(\bar{e}_k - \bar{g}_k)\mathbf{Z}_k\mathbf{Z}_k^{\mathsf{H}}\| \xrightarrow[N\to\infty]{\text{a.s.}} 0$, we have

$$\frac{1}{N}\log\det\left(\mathbf{I}_N + \rho\sum_{k=1}^{K}\bar{e}_k\mathbf{Z}_k\mathbf{Z}_k^{\mathsf{H}}\right) - \frac{1}{N}\log\det\left(\mathbf{I}_N + \rho\sum_{k=1}^{K}\bar{g}_k\mathbf{Z}_k\mathbf{Z}_k^{\mathsf{H}}\right) \xrightarrow[N\to\infty]{\text{a.s.}} 0. \tag{86}$$

Applying Theorem 9 to the last term yields

$$\begin{aligned}&\frac{1}{N}\log\det\left(\mathbf{I}_N + \rho\sum_{k=1}^{K}\bar{g}_k\mathbf{Z}_k\mathbf{Z}_k^{\mathsf{H}}\right)\\ &\quad - \frac{1}{N}\log\det\left(\mathbf{I}_N + \rho\sum_{k=1}^{K}\frac{n_k}{N_k}\frac{\bar{g}_k g_k}{\delta_k}\mathbf{R}_k\right) - \frac{1}{N}\sum_{k=1}^{K}\log\det\left(\mathbf{I}_{N_k} + \bar{g}_k\delta_k\mathbf{S}_k\right) + n_k\bar{g}_k g_k \xrightarrow[N\to\infty]{\text{a.s.}} 0.\end{aligned} \tag{87}$$

Combining (85) and (87) finally leads to

$$\bar{V}_N(\rho) - \bar{I}_N(\rho) \xrightarrow[N\to\infty]{\text{a.s.}} 0 \tag{88}$$

where

$$\begin{aligned}\bar{I}_N(\rho) = &\frac{1}{N}\log\det\left(\mathbf{I}_N + \rho\sum_{k=1}^{K}\frac{n_k}{N_k}\frac{\bar{g}_k g_k}{\delta_k}\mathbf{R}_k\right) + \frac{1}{N}\sum_{k=1}^{K}\log\det\left(\mathbf{I}_{N_k} + \bar{g}_k\delta_k\mathbf{S}_k\right) - n_k\bar{g}_k g_k\\ &+ \frac{1}{N}\sum_{k=1}^{K}\log\det\left(\mathbf{I}_{n_k} + g_k\tilde{\mathbf{T}}_k\right) - n_k g_k\bar{g}_k.\end{aligned} \tag{89}$$

This concludes the proof of part (i).

In order to show the convergence in the mean (part (ii)), we will pursue the same approach as in [5], [7]. Define the following functions:

$$m_N(z) = \frac{1}{N}\text{tr}\left(\sum_{k=1}^{K}\mathbf{H}_k\mathbf{H}_k^{\mathsf{H}} - z\mathbf{I}_N\right)^{-1}, \qquad \bar{m}_N(z) = \frac{1}{N}\text{tr}\left(\sum_{k=1}^{K}\frac{n_k}{N_k}\frac{\bar{g}_k g_k}{\delta_k}\mathbf{R}_k - z\mathbf{I}_N\right)^{-1}.$$

One can show that

$$\mathbb{E}\left[I_N(\rho)\right] - \bar{I}_N(\rho) = \int_{\frac{1}{\rho}}^{\infty}\left(\left[\frac{1}{\omega} - \mathbb{E}\left[m_N(-\omega)\right]\right] - \left[\frac{1}{\omega} - \bar{m}_N(-\omega)\right]\right)d\omega. \tag{90}$$

Since both $m_N(\omega)$ and $\bar{m}_N(\omega)$ are uniformly bounded by $\frac{1}{\omega}$, it follows from dominated convergence arguments, Theorem 9 and (84) that, for all $\omega > 0$,

$$\left[\frac{1}{\omega} - \mathbb{E}\left[m_N(-\omega)\right]\right] - \left[\frac{1}{\omega} - \bar{m}_N(-\omega)\right] \to 0. \tag{91}$$

Moreover,

$$\begin{aligned}\left|\left[\frac{1}{\omega} - \mathbb{E}\left[m_N(-\omega)\right]\right] - \left[\frac{1}{\omega} - \bar{m}_N(-\omega)\right]\right| &\leq \left|\left[\frac{1}{\omega} - \mathbb{E}\left[m_N(-\omega)\right]\right]\right| + \left|\left[\frac{1}{\omega} - \bar{m}_N(-\omega)\right]\right| \\ &\leq \frac{1}{\omega^2}\left(\frac{1}{N}\mathrm{tr}\,\mathbb{E}\left[\sum_{k=1}^{K}\mathbf{H}_k\mathbf{H}_k^{\mathsf{H}}\right] + \frac{1}{N}\mathrm{tr}\left(\sum_{k=1}^{K}\frac{n_k}{N_k}\frac{\bar{g}_k g_k}{\delta_k}\mathbf{R}_k\right)\right) \\ &\leq \frac{2KRST}{\omega^2}\end{aligned} \tag{92}$$

where $R = \max_k \limsup\|\mathbf{R}_k\|$, $S = \max_k \limsup\|\mathbf{S}_k\|$, $T = \max_k \limsup\|\mathbf{T}_k\mathbf{Q}_k\|$. Since $\frac{2KRST}{\omega^2}$ is finite and integrable over $[\frac{1}{\rho}, \infty)$, it follows from the dominated convergence theorem that

$$\mathbb{E}\left[I_N(\rho)\right] - \bar{I}_N(\rho) \xrightarrow[N\to\infty]{} 0. \tag{93}$$

Appendix F

Proof of Theorem 6: Optimal power allocation

The proof follows closely those of [25, Proposition 5] and [14, Proposition 3].

We first recall the following property of concave functions (see e.g. [50]):

Property 1: A function $f\left(\mathbf{Q}_1, \ldots, \mathbf{Q}_K\right)$ is strictly concave in the Hermitian nonnegative matrices $\mathbf{Q}_1, \ldots, \mathbf{Q}_K$, if and only if, for any couples $\left(\mathbf{Q}_{1_a}, \mathbf{Q}_{1_b}\right), \ldots, \left(\mathbf{Q}_{K_a}, \mathbf{Q}_{K_b}\right)$ of Hermitian nonnegative matrices, the function

$$\phi(\lambda) = f\left(\lambda\mathbf{Q}_{1_a} + (1-\lambda)\mathbf{Q}_{1b}, \ldots, \lambda\mathbf{Q}_{K_a} + (1-\lambda)\mathbf{Q}_{Kb}\right), \qquad \lambda \in [0,1]$$

is strictly concave.

Consider now $\bar{I}_N(\rho)$ seen as a function of λ for $\mathbf{Q}_k = \lambda\mathbf{Q}_{k_a} - (1-\lambda)\mathbf{Q}_{k_b}$, where $\mathbf{Q}_{k_a}, \mathbf{Q}_{k_b}$ are Hermitian nonnegative definite matrices, for $k = 1, \ldots, K$. Thus, by the chain rule of differentiation,

$$\frac{d\bar{I}_N(\rho)}{d\lambda} = \frac{\partial\bar{I}_N(\rho)}{\partial\lambda} + \sum_{k=1}^{K}\frac{\partial\bar{I}_N(\rho)}{\partial\bar{g}_k}\frac{\partial\bar{g}_k}{\partial\lambda} + \frac{\partial\bar{I}_N(\rho)}{\partial g_k}\frac{\partial g_k}{\partial\lambda} + \frac{\partial\bar{I}_N(\rho)}{\partial\delta_k}\frac{\partial\delta}{\partial\lambda}. \tag{94}$$

One can verify that the partial derivatives of $\bar{I}_N\left(\rho\right)$ with respect to $g_k, \bar{g}_k, \delta_k$, respectively, satisfy

$$\frac{\partial\bar{I}_N\left(\rho\right)}{\partial g_k} = \frac{\partial\bar{I}_N\left(\rho\right)}{\partial\bar{g}_k} = \frac{\partial\bar{I}_N\left(\rho\right)}{\partial\delta_k} = 0, \tag{95}$$

due to the defining relation (21). Thus,

$$\frac{d\bar{I}_N(\rho)}{d\lambda} = \frac{\partial \bar{I}_N(\rho)}{\partial \lambda} = \sum_{k=1}^{K} \frac{1}{N} \mathrm{tr}\, \mathbf{T}_k^{\frac{1}{2}} \left(\mathbf{I}_{n_k} + g_k \mathbf{T}_k^{\frac{1}{2}} \left(\lambda \mathbf{Q}_{k_a} + (1-\lambda) \mathbf{Q}_{k_b} \right) \mathbf{T}_k^{\frac{1}{2}} \right)^{-1} \mathbf{T}_k^{\frac{1}{2}} \left(\mathbf{Q}_{k_a} - \mathbf{Q}_{k_b} \right). \quad (96)$$

The second derivative therefore reads

$$\frac{d^2\bar{I}_N(\rho)}{d\lambda^2} = -\sum_{k=1}^{K} \frac{1}{N} \mathrm{tr} \left[\underbrace{\mathbf{T}_k^{\frac{1}{2}} \left(\mathbf{I}_{n_k} + g_k \mathbf{T}_k^{\frac{1}{2}} \left(\lambda \mathbf{Q}_{k_a} + (1-\lambda) \mathbf{Q}_{k_b} \right) \mathbf{T}_k^{\frac{1}{2}} \right)^{-1} \mathbf{T}_k^{\frac{1}{2}}}_{\triangleq \mathbf{A}_k} \underbrace{\left(\mathbf{Q}_{k_a} - \mathbf{Q}_{k_b} \right)}_{\triangleq \mathbf{B}_k} \right]^2 \quad (97)$$

where $\mathbf{A}_k$ are Hermitian nonnegative definite and $\mathbf{B}_k$ are Hermitian. Let $\mathbf{A}_k = \mathbf{U}_k \mathbf{D}_k \mathbf{U}_k^{\mathsf{H}}$ be the eigenvalue decomposition of $\mathbf{A}_k$, where $\mathbf{U}_k \in \mathbb{C}^{n_k \times n_k}$ are unitary matrices and $\mathbf{D}_k = \mathrm{diag}\left(d_{k,1}, \ldots, d_{k,n_k}\right) \succeq 0$. Moreover, denote $\mathbf{Z}_k = \mathbf{B}_k \mathbf{U}_k = [\mathbf{z}_{k,1} \ldots \mathbf{z}_{k,n_k}]$. Then,

$$\frac{1}{N} \mathrm{tr}\, \left[\mathbf{A}_k \mathbf{B}_k \right]^2 = \frac{1}{N} \mathrm{tr}\, \mathbf{D}_k \mathbf{U}^{\mathsf{H}} \mathbf{B}_k \mathbf{A}_k \mathbf{B}_k \mathbf{U} = \frac{1}{N} \mathrm{tr}\, \mathbf{D}_k \mathbf{Z}_k^{\mathsf{H}} \mathbf{A}_k \mathbf{Z}_k = \frac{1}{N} \sum_{j=1}^{n_k} d_{k,j} \mathbf{z}_{k,j}^{\mathsf{H}} \mathbf{A}_k \mathbf{z}_{k,j} \geq 0. \quad (98)$$

If $\mathbf{T}_k \succ 0$, or equivalently, if $\mathbf{T}_k$ is invertible, we have $\mathbf{A}_k \succ 0$ and (98) holds with strict inequality for $\mathbf{Q}_{k_a} \neq \mathbf{Q}_{k_b}$. Thus, if any of the matrices $\mathbf{T}_k$ is invertible and $\mathbf{Q}_{k_a} \neq \mathbf{Q}_{k_b}$, we have

$$\sum_{k=1}^{K} \frac{1}{N} \mathrm{tr}\, \left[\mathbf{A}_k \mathbf{B}_k \right]^2 > 0 \quad (99)$$

and hence $\frac{d^2\bar{I}_N(\rho)}{d\lambda^2} < 0$. Thus, $\bar{I}_N(\rho)$ is strictly concave in the matrices $\mathbf{Q}_k$. Due to (95) it is then sufficient to maximize

$$\log \det \left(\mathbf{I}_{n_k} + g_k \mathbf{T}_k^{\frac{1}{2}} \mathbf{Q}_k \mathbf{T}_k^{\frac{1}{2}} \right) \quad (100)$$

with respect to $\mathbf{Q}_k$ and with the constraint $\frac{1}{n_k} \mathrm{tr}\, \mathbf{Q}_k \leq P_k$. The solution to this problem is well-known [41] and given by the water-filling solution stated in the theorem. However, since g_k depends on $\mathbf{Q}_k$, this solution must be computed iteratively by Algorithm 1.

Appendix G

Proof of Theorem 7: SINR of the MMSE detector

Similar to the proof of Theorem 5, let us consider the matrix model

$$\mathbf{H}_k = \frac{1}{\sqrt{n_k}} \mathbf{Z}_k \mathbf{W}_{2,k} \mathbf{T}_k^{\frac{1}{2}} \mathbf{Q}_k^{\frac{1}{2}} \quad (101)$$

where

$$\mathbf{Z}_k = \frac{1}{\sqrt{N_k}} \mathbf{R}_k^{\frac{1}{2}} \mathbf{W}_{k,1} \mathbf{S}_k^{\frac{1}{2}}. \quad (102)$$

Then,

$$\gamma_{k,j}^N = p_{k,j}t_{k,j}\frac{1}{n_k}\mathbf{w}_{2,k,j}^{\mathsf{H}}\mathbf{Z}_k^{\mathsf{H}}\left(\sum_{i=1}^K \mathbf{H}_i\mathbf{H}_i^{\mathsf{H}} - p_{k,j}t_{k,j}\frac{1}{n_k}\mathbf{Z}_k\mathbf{w}_{2,k,j}\mathbf{w}_{2,k,j}\mathbf{Z}_k^{\mathsf{H}} + \frac{1}{\rho}\mathbf{I}_N\right)^{-1}\mathbf{Z}_k\mathbf{w}_{2,k,j}. \quad (103)$$

It was shown in the proof of Theorem 5 that, almost surely, $\limsup_N \|\mathbf{Z}_k\mathbf{Z}_k^{\mathsf{H}}\| < 0$. From the Fubini theorem, Lemma 2 and Lemma 4, we therefore have

$$\gamma_{k,j}^N - p_{k,j}t_{k,j}\frac{1}{n_k}\mathrm{tr}\,\mathbf{Z}_k\mathbf{Z}_k^{\mathsf{H}}\left(\sum_{i=1}^K \mathbf{H}_i\mathbf{H}_i^{\mathsf{H}} + \rho\mathbf{I}_{n_k}\right) \xrightarrow[N\to\infty]{\text{a.s}} 0. \quad (104)$$

Applying Theorem 9 (i) leads to

$$\gamma_{k,j}^N - p_{k,j}t_{k,j}\frac{1}{n_k}\mathrm{tr}\,\mathbf{Z}_k\mathbf{Z}_k^{\mathsf{H}}\left(\sum_{i=1}^K \bar{e}_i\mathbf{Z}_i\mathbf{Z}_i^{\mathsf{H}} + \rho\mathbf{I}_{n_k}\right)^{-1} \xrightarrow[N\to\infty]{\text{a.s}} 0 \quad (105)$$

where $\bar{e}_i$ are given as the unique solutions to (67). Notice now from (67) that

$$e_k = \frac{1}{n_k}\mathrm{tr}\,\mathbf{Z}_k\mathbf{Z}_k^{\mathsf{H}}\left(\sum_{i=1}^K \bar{e}_i\mathbf{Z}_i\mathbf{Z}_i^{\mathsf{H}} + \rho\mathbf{I}_{n_k}\right)^{-1} \quad (106)$$

and that $\max_k |e_k - g_k| \xrightarrow[N\to\infty]{\text{a.s}} 0$ by (84). This finally implies

$$\gamma_{k,j}^N - p_{k,j}t_{k,j}g_k \xrightarrow[N\to\infty]{\text{a.s}} 0. \quad (107)$$

Appendix H

Proof of Corollary 2: Sum-rate with MMSE decoding

Part (i) is a simple consequence of Theorem 7 and the continuous mapping theorem.

For Part (ii), first notice that $R_N(\rho) \le I_N(\rho)$ and $\bar{R}_N(\rho) \le \bar{I}_N(\rho)$. Thus,

$$\left| R_N(\rho) - \frac{1}{N}\sum_{k=1}^K\sum_{j=1}^{n_k}\log\left(1 + p_{k,j}t_{k,j}g_k\right)\right| \le I_N(\rho) + \bar{I}_N(\rho) \triangleq \phi_N(\rho). \quad (108)$$

Since $\mathbb{E}\left[\phi_N(\rho)\right] < \infty$ by Theorem 5 (ii), it follows from dominated convergence arguments that

$$\mathbb{E}\left[R_N(\rho)\right] - \bar{R}_N(\rho) \xrightarrow[N\to\infty]{} 0. \quad (109)$$

Appendix I

Proof of Corollary 3: Rayleigh product channel

Under the assumptions of the corollary, the fundamental equations in Theorem 4 reduce to

$$\bar{g} = \frac{1}{1+g} \quad (110)$$

$$g = \frac{S}{N}\frac{\delta}{1+\bar{g}\delta} \quad (111)$$

$$\delta = \frac{1}{K\frac{\bar{g}g}{\delta} + \frac{S}{N}\frac{1}{\rho}} \quad (112)$$

From (110), we have

$$g = \frac{1-\bar{g}}{\bar{g}}. \tag{113}$$

Solving (111) for δ and replacing g by (113) yields

$$\delta = \frac{1-\bar{g}}{\bar{g}\left(\bar{g}+\frac{S}{N}-1\right)}. \tag{114}$$

Solving (112) for δ and replacing g by (113) leads to

$$\delta = \frac{1-K(1-\bar{g})}{\frac{S}{N}\frac{1}{\rho}}. \tag{115}$$

Equating (114) and (115) and rearranging the terms as a polynomial in $\bar{g}$ finally yields

$$\bar{g}^3 - \bar{g}^2\left(2-\frac{S}{N}-\frac{1}{K}\right) + \bar{g}\left(1-\frac{S}{N}-\frac{1}{K}+\frac{S}{NK}\left(1+\frac{1}{\rho}\right)\right) - \frac{S}{NK}\frac{1}{\rho} = 0. \tag{116}$$

By Theorem 4, only one of the roots of this polynomial satisfies $\bar{g}, g, \delta > 0$. Now, (113) implies $\bar{g} < 1$, (114) implies $\bar{g} > 1 - \frac{S}{N}$ (115) implies $\bar{g} > 1 - \frac{1}{K}$. Hence $\bar{g} \in \left(1 - \min\left[\frac{1}{K}, \frac{S}{N}\right], 1\right)$.

Similarly, $\bar{I}_N(\rho)$ reduces under the assumptions of the corollary to

$$\bar{I}_N(\rho) = \log\left(1+\rho\frac{NK}{S}\frac{\bar{g}g}{\delta}\right) + \frac{KS}{N}\log(1+\bar{g}\delta) + K\log(1+g) - 2K\bar{g}g. \tag{117}$$

Replacing $\frac{g}{\delta}$ by $\bar{g}+\frac{S}{N}-1$ in the first term, δ by (114) in the second term, g by (113) in the third term and $\bar{g}g$ by $(1-\bar{g})$ in the last term leads to the desired result.

The simplification of Theorem 7 is immediate since $\bar{\gamma}_{kj}^N = p_{k,j}t_{k,j}g_k = \frac{1-\bar{g}}{\bar{g}}$ by (113).

APPENDIX J

MATLAB CODE RELATED TO THE MULTIHOP AF MIMO RELAY CHANNEL

A. Code for Corollary 1: `corollary1.m`

```
function I_k = corollary1(k,K,a,rho,c)
% Compute deterministic equivalent I_k of the asymptotic mutual information
% Input parameters :
% k                              : denotes which function I_k to compute
% K                              : total number number of hops
% a   = [alpha_1,...,alpha_{K}] : vector containing the path loss factors
% rho = [rho_0,...,rho_{K-1}]   : vector containing the power budgets rho_k
% c   = [c_1,...,c_K]           : vector containing the matrix dimension ratios
% Calculate asymptotic power normalization factors (Lemma 1)
b  = zeros(1,k);
b(1) = rho(1);
for i=2:k
    b(i) = rho(i) / (1+a(i-1)*rho(i-1));
end
% Calculate capacity
I_k = 1/K*(theorem2(k,1,a,b,c) - theorem2(k,1,a,[0,b(2:end)],c));
end
```

B. Code for Theorem 2: `theorem2.m`

```
function J = theorem2(k,x,a,b,c)
% Recursively computes deterministic equivalent of J_k
% Input parameters :
% k                            : denotes which function J_k to compute
% x                            : argument of J_k
% a = [alpha_1,...,alpha_{k}] : vector containing the path loss factors
% b = [beta_0,...,beta_{k-1}] : power normalization factors
% c = [c_1,...,c_K]           : vector containing the matrix dimension ratios
if (k==1) % J_1 is given in closed form
    [~,e] = theorem3(k-1,x,a,b,c);
    J     = c(k)*log(1 + x*a(k)*b(k)/(c(k)+e)) + log(1+e/c(k)) - e/(c(k)+e);
else      % J_k, k>1, must be computed recursively
    [~,e] = theorem3(k-1,x,a,b,c);
    J     = c(k)*theorem2(k-1,x*a(k)*b(k)/(c(k)+x*a(k)*b(k) + e),a,b,c)...
            + c(k)*log(1 + x*a(k)*b(k)/(c(k)+e)) + log(1 + e/c(k)) - e/(c(k)+e);
end
end
```

C. Code for Theorem 3: `theorem3.m`

```
function [m,e] = theorem3(k,x,a,b,c)
% Recursively computes the quantities e_k
% Input parameters :
% k                          : denotes which function e_k to compute
% x                          : argument of e_k
% a = [alpha_1,...,alpha_{k}] : vector containing the path loss factors
% b = [beta_0,...,beta_{k}]   : power normalization factors
% c = [c_1,...,c_K]           : vector containing the matrix dimension ratios
if (k==0) % For k=0, e_o and m_0 are given in closed form
    e = -(x*a(1)*b(1)*(1-c(1))+c(1))/2 + sqrt((x*a(1)*b(1)*(1-c(1))+c(1))^2...
        + 4*x*a(1)*b(1)*c(1)^2)/2;
    m = c(1) / (a(1)*b(1)/(c(1)+e)+1/x) + (1-c(1))*x;
else      % For k>0, e_k is given as the solution to a fixed point equation
    e    = 0;
    eold = 1;
    while abs(e-eold)>1e-6  % the error tolerance 1e-6 can be changed
        eold = e;
        e = c(k+1)*(c(k+1)+eold) - c(k+1)*(c(k+1)+eold)^2/(x*a(k+1)*b(k+1))...
            * theorem3(k-1,x*a(k+1)*b(k+1)/(c(k+1)+x*a(k+1)*b(k+1)+eold),a,b,c);
    end
    m = x*c(k+1)/(c(k+1)+e);
end
end
```

References

[1] D. N. C. Tse and S. V. Hanly, "Linear multiuser receivers: effective interference, effective bandwidth and user capacity," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 641–657, Feb. 1999.

[2] A. M. Tulino and S. Verdú, "Random matrix theory and wireless communications," *Foundations and Trends in Communications and Information Theory*, vol. 1, no. 1, 2004.

[3] R. Couillet and M. Debbah, *Random matrix methods for wireless communications*, 1st ed. New York, NY, USA: Cambridge University Press, 2011, to appear.

[4] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, Feb. 1999.

[5] W. Hachem, P. Loubaton, and J. Najim, "Deterministic equivalents for certain functionals of large random matrices," *Annals of Applied Probability*, vol. 17, no. 3, pp. 875–930, 2007.

[6] M. Debbah, P. Loubaton, and M. de Courville, "The spectral efficiency of linear precoders," in *Proc. IEEE Information Theory Workshop (ITW'03)*, Paris, France, Mar. 2003, pp. 90–93.

[7] R. Couillet, J. Hoydis, and M. Debbah, "Random beamforming over quasi-static and fading channels: A deterministic equivalent approach," *IEEE Trans. Inf. Theory*, 2011, submitted. [Online]. Available: http://arxiv.org/pdf/1011.3717v2

[8] V. K. Nguyen and J. S. Evans, "Multiuser transmit beamforming via regularized channel inversion: A large system analysis," in *Proc. IEEE Global Communications Conference (Globecom08)*, New Orleans, LO, US, Dec. 2008, pp. 1–4.

[9] S. Wagner, R. Couillet, M. Debbah, and D. T. M. Slock, "Large system analysis of linear precoding in MISO broadcast channels with limited feedback," *IEEE Trans. Inf. Theory*, 2011, submitted for publication. [Online]. Available: http://arxiv.org/abs/0906.3682

[10] R. Zakhour and S. Hanly, "Large system analysis of base station cooperation on the downlink," in *Proc. 48th Annual Allerton Conference on Communication, Control, and Computing*, Urbana-Champaign, IL, US, Oct. 2010, pp. 270 –277.

[11] H. Huh, G. Caire, S.-H. Moon, and I. Lee, "Multi-cell MIMO downlink with fairness criteria: The large system limit," in *Proc. International Symposium on Information Theory Proceedings (ISIT'2010)*, Jun. 2010, pp. 2058–2062.

[12] J. Hoydis, M. Kobayashi, and M. Debbah, "Optimal channel training in uplink network MIMO systems," *IEEE Trans. Signal Process.*, vol. 59, no. 6, Jun. 2011.

[13] L. Li, A. M. Tulino, and S. Verdú, "Design of reduced-rank MMSE multiuser detectors using random matrix methods," *IEEE Trans. Inf. Theory*, vol. 50, no. 6, pp. 986–1008, Jun. 2004.

[14] R. Couillet, M. Debbah, and J. W. Silverstein, "A deterministic equivalent for the analysis of correlated MIMO multiple access channels," *IEEE Trans. Inf. Theory*, vol. 57, no. 6, pp. 3493–3514, Jun. 2011.

[15] F. Dupuy and P. Loubaton, "On the capacity achieving covariance matrix for frequency selective MIMO channels using the asymptotic approach," *IEEE Trans. Inf. Theory*, 2010, to appear. [Online]. Available: http://arxiv.org/abs/1001.3102

[16] W. Hachem, O. Khorunzhy, P. Loubaton, J. Najim, and L. A. Pastur, "A new approach for capacity analysis of large dimensional multi-antenna channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 9, 2008.

[17] A. L. Moustakas and S. H. Simon, "Random matrix theory of multi-antenna communications: the Rician channel," *Journal of Physics A: Mathematical and General*, vol. 38, no. 49, pp. 10 859–10 872, Nov. 2005.

[18] ——, "On the outage capacity of correlated multiple-path MIMO channels," *IEEE Trans. Inf. Theory*, vol. 53, no. 11, pp. 3887–3903, 2007.

[19] P. Billingsley, *Probability and Measure*, 3rd ed. Hoboken, NJ: John Wiley and Sons, Inc., 1995.

[20] S. Borade, L. Zheng, and R. Gallager, "Amplify-and-forward in wireless relay networks: rate, diversity, and network size," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3302–3318, 2007.

[21] I. Maric, A. Goldsmith, and M. Médard, "Analog network coding in the high-SNR regime," in *IEEE Wireless Network Coding Conference (WiNC'10)*, Boston, MA, USA, 2010, pp. 1–6.

[22] N. Fawaz, K. Zarifi, M. Debbah, and D. Gesbert, "Asymptotic capacity and optimal precoding in MIMO multi-hop relay networks," *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 2050–2069, 2011.

[23] D. Gesbert, H. Bolcskei, D. A. Gore, and A. J. Paulraj, "Outdoor MIMO wireless channels: models and performance prediction," *IEEE Trans. Commun.*, vol. 50, no. 12, pp. 1926–1934, Dec. 2002.

[24] R. R. Müller and H. Hofstetter, "Confirmation of random matrix model for the antenna array channel by indoor measurements," in *Proc. IEEE Int. Symp. Antennas and Propagation Society*, vol. 1, 2001, pp. 472–475.

[25] J. Dumont, W. Hachem, S. Lasaulce, P. Loubaton, and J. Najim, "On the capacity achieving covariance matrix for Rician MIMO channels: an asymptotic approach," *IEEE Trans. Inf. Theory*, vol. 56, no. 3, pp. 1048–1069, 2010.

[26] G. H. Tucci, "Spectral analysis of the amplify and forward relay network as the number of relay layers increases," in *Proc. 8th International Symposium on Wireless Communication Systems (ISWCS'11)*, Aachen, Germany, Nov. 2011.

[27] S.-P. Yeh and O. Leveque, "Asymptotic capacity of multi-level amplify-and-forward relay networks," in *Proc. IEEE International Symposium on Information Theory (ISIT'07)*, Jun. 2007, pp. 1436–1440.

[28] V. I. Morgenshtern and H. Blcskei, "Random matrix analysis of large relay networks," in *Proc. 44th Allerton Annual Conference on Communications, Control and Computing*, Urbana-Champaign, IL, US, Sep. 2006, pp. 106–112.

[29] J. Wagner, B. Rankov, and A. Wittneben, "Large n analysis of amplify-and-forward MIMO relay channels with correlated Rayleigh fading," *IEEE Trans. Inf. Theory*, vol. 54, no. 12, pp. 5735–5746, Dec. 2008.

[30] S. Jin, M. R. McKay, C. Zhong, and K.-K. Wong, "Ergodic capacity analysis of amplify-and-forward MIMO dual-hop systems," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2204–2224, May 2010.

[31] A. W. van der Vaart, *Asymptotic Statistics (Cambridge Series in Statistical and Probabilistic Mathematics)*. Cambridge University Press, New York, 1998.

[32] R. R. Müller, "A random matrix model of communication via antenna arrays," *IEEE Trans. Inf. Theory*, vol. 48, no. 9, pp. 2495–2506, Sep. 2002.

[33] D. Chizhik, G. J. Foschini, M. J. Gans, and R. A. Valenzuela, "Keyholes, correlations, and capacities of multielement transmit and receive antennas," *IEEE Trans. Wireless Commun.*, vol. 1, no. 2, pp. 361–368, Apr. 2002.

[34] P. Almers, F. Tufvesson, and A. F. Molisch, "Keyhole effect in MIMO wireless channels: Measurements and theory," *IEEE Trans. Wireless Commun.*, vol. 5, no. 12, pp. 3596–3604, Dec. 2006.

[35] H. Shin and J. H. Lee, "Capacity of multiple-antenna fading channels: spatial fading correlation, double scattering, and keyhole," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2636–2647, Oct. 2003.

[36] G. Levin and S. Loyka, "Multi-keyhole MIMO channels: asymptotic analysis of outage capacity," in *Proc. IEEE International Symposium on Information Theory (ISIT'06)*, Seattle, Washington, US, Jul. 2006, pp. 1305–1309.

[37] H. Shin and M. Z. Win, "MIMO diversity in the presence of double scattering," *IEEE Trans. Inf. Theory*, vol. 54, no. 7, pp. 2976–2996, Jul. 2008.

[38] S. Yang and J.-C. Belfiore, "Diversity-multiplexing tradeoff of double scattering MIMO channels," *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 2027–2034, Apr. 2011.

[39] S. Jin, M. R. McKay, K.-K. Wong, and X. Gao, "Transmit beamforming in Rayleigh product MIMO channels: capacity and performance analysis," *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 5204–5221, Oct. 2008.

[40] X. Li, S. Jin, X. Gao, and M. R. McKay, "Capacity bounds and low complexity transceiver design for double-scattering MIMO multiple access channels," *IEEE Trans. Signal Process.*, vol. 58, no. 5, pp. 2809–2822, May 2010.

[41] T. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. New York, NY, USA: John Wiley and Sons, Inc., 2006.

[42] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice-Hall, Inc. Upper Saddle River, NJ, USA, 1993.

[43] D. Gesbert, T. Ekman, and N. Christophersen, "Capacity limits of dense palm-sized MIMO arrays," in *Proc. Global Telecommunications Conference (GLOBECOM '02)*, vol. 2, Taipei, Taiwan, Nov. 2002, pp. 1187–1191.

[44] D. Tse and P. Viswanath, *Fundamentals of Wireless Communications*. New York, NY, USA: Cambridge University Press, 2005.

[45] W. H. A. Kammoun, M. Kharouf and J. Najim, "A central limit theorem for the SINR at the LMMSE estimator output for large dimensional systems," *IEEE Trans. Inf. Theory*, vol. 55, pp. 5048–5063, Nov. 2009.

[46] Z. D. Bai and J. W. Silverstein, "No eigenvalues outside the support of the limiting spectral distribution of large dimensional sample covariance matrices," *The Annals of Probability*, vol. 26, no. 1, pp. 316–345, Jan. 1998.

[47] J. W. Silverstein and Z. D. Bai, "On the empirical distribution of eigenvalues of a class of large dimensional random matrices," *Journal of Multivariate Analysis*, vol. 54, no. 2, pp. 175–192, 1995.

[48] R. Yates, "A framework for uplink power control in cellular radio systems," *IEEE J. Sel. Areas Commun.*, vol. 13, no. 7, pp. 1341–1347, Sep. 1995.

[49] E. C. Titchmarsh, *The Theory of Functions*. Oxford University Press, London, Aug. 1939.

[50] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.