Random Beamforming over Quasi-Static and Fading Channels: A Deterministic Equivalent Approach

Romain Couillet, Jakob Hoydis, Mérouane Debbah

Random Beamforming over Quasi-Static and Fading Channels: A Deterministic Equivalent Approach

Romain Couillet[†,‡], Jakob Hoydis[*,‡] and Mérouane Debbah[‡]

Abstract

In this work, we study the performance of random isometric precoding over quasi-static and correlated fading channels. We derive deterministic approximations of the mutual information and the signal-to-interference-plus-noise ratio (SINR) at the output of the minimum-mean-square-error (MMSE) receiver and provide simple provably converging fixed-point algorithms for their computation. Although the deterministic approximations are only asymptotically exact, almost surely, we show by simulations that they are very accurate for small system dimensions. The analysis is based on the Stieltjes transform method which enables the derivation of deterministic equivalents of functionals of large-dimensional random matrices. In contrast to previous works, our analysis does not rely on arguments from free probability theory which allows us to consider random matrix models for which asymptotic freeness does not hold. Thus, the results of this work are also a novel contribution to the field of random matrix theory and are shown to be applicable to a wide spectrum of practical systems. In this article, we specifically characterize the performance of multi-cellular communication systems, multiple-input multiple-output multiple-access channels (MIMO-MAC), and MIMO interference channels.

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

[‡]ST-Ericsson, 505 Route des Lucioles, 06560 Sophia Antipolis, France.

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

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

I. INTRODUCTION

Consider the following discrete time wireless channel model

$$\mathbf{y} = \sum_{k=1}^{K} \mathbf{H}_k \mathbf{W}_k \mathbf{P}_k^{\frac{1}{2}} \mathbf{x}_k + \mathbf{n} \tag{1}$$

where

(i) $\mathbf{y} \in \mathbb{C}^N$ is the channel output vector,

(ii) $\mathbf{H}_k \in \mathbb{C}^{N \times N_k}$, $k \in \{1, \ldots, K\}$, are complex channel matrices, satisfying either of the following properties:

(ii-a) The matrix $\mathbf{H}_k \in \mathbb{C}^{N \times N_k}$ is deterministic. In this case, we will denote $\mathbf{R}_k = \mathbf{H}_k \mathbf{H}_k^{\mathsf{H}}$.

(ii-b) The matrix $\mathbf{H}_k \in \mathbb{C}^{N \times N_k}$ is a random channel matrix whose jth column vector $\mathbf{h}_{kj} \in \mathbb{C}^N$ is modeled as

$$\mathbf{h}_{kj} = \mathbf{R}_{kj}^{\frac{1}{2}} \mathbf{z}_{kj}, \qquad j \in \{1, \ldots, N_k\} \tag{2}$$

where $\mathbf{R}_{kj} \in \mathbb{C}^{N \times N}$ are Hermitian nonnegative definite matrices and the vectors $\mathbf{z}_{kj} \in \mathbb{C}^N$ have independent and identically distributed (i.i.d.) elements with zero mean, variance $1/N$ and $4+\epsilon$ moment of order $\mathcal{O}(1/N^{2+\varepsilon/2})$, for some common $\epsilon > 0$.

(iii) $\mathbf{W}_k \in \mathbb{C}^{N_k \times n_k}$, $k \in \{1, \ldots, K\}$, are complex (signature or precoding) matrices which contain each $n_k \leq N_k$ orthonormal columns of independent $N_k \times N_k$ Haar-distributed random unitary matrices,

(iv) $\mathbf{P}_k \in \mathbb{R}^{n_k \times n_k}$, $k \in \{1, \ldots, K\}$, are diagonal (power loading) matrices with nonnegative entries,

(v) $\mathbf{x}_k \sim \mathcal{CN}(0, \mathbf{I}_{n_k})$, $k \in \{1, \ldots, K\}$, are random independent Gaussian transmit vectors,

(vi) $\mathbf{n} \sim \mathcal{CN}(0, \sigma^2 \mathbf{I}_N)$ is a white Gaussian noise vector.

In addition, we define the ratios of the matrix dimensions $c_i \triangleq \frac{n_i}{N_i}$ and $\bar{c}_i \triangleq \frac{N_i}{N}$ for $i \in \{1, \ldots, K\}$.

Remark 1: The statistical model (2) of the channel $\mathbf{H}_k$ under assumption (ii-b) generalizes several well-known fading channel models of interest (see [1], [2] for examples). These models comprise in particular the *Kronecker* channel model with transmit and receive correlation matrices [3], [4], where the matrices $\mathbf{H}_k$ are given by

$$\mathbf{H}_k = \mathbf{R}_k^{\frac{1}{2}} \mathbf{Z}_k \mathbf{T}_k^{\frac{1}{2}} \tag{3}$$

with $\mathbf{Z}_k \in \mathbb{C}^{N \times N_k}$ a random matrix whose elements are independent $\mathcal{CN}(0, 1/N)$ and $\mathbf{R}_k \in \mathbb{C}^{N \times N}$, $\mathbf{T}_k \in \mathbb{C}^{N_k \times N_k}$ antenna correlation matrices. Since both $\mathbf{Z}_k$ and $\mathbf{W}_k$ are unitarily invariant, we can assume without loss of generality for the statistical properties of $\mathbf{y}$ that $\mathbf{T}_k = \mathrm{diag}(t_{k1}, \ldots, t_{kN_k})$. Defining the matrices $\mathbf{R}_{kj} = t_{kj} \mathbf{R}_k$ for $j \in \{1, \ldots, N_k\}$, we fall back to the channel model in (2). Taking instead all $\mathbf{R}_{kj}$ to be diagonal matrices makes the entries of $\mathbf{H}_k$ independent with $[\mathbf{H}_k]_{ij}$ of zero mean and variance $[\mathbf{R}_{kj}]_{ii}/N$. This corresponds to a centered variance profile model, studied extensively in [5], [6], [7].

The objective of this work is to study the performance of the communication channel (1) in the large dimensional regime where $N, N_1, \ldots, N_K, n_1, \ldots, n_K$ are simultaneously large. In the following, we will consider both the *quasi-static channel scenario* which assumes hypotheses (i), (ii-a), (iii)-(vi), and the *fading channel scenario* which assumes (i), (ii-b), (iii)-(vi). The study of the latter naturally arises as an extension of the study of the quasi-static channel scenario. The respective application contexts of both scenarios are described below.

A. Quasi-static channel scenario (hypothesis (ii-a))

Possible applications of the channel model (1) under assumptions (i), (ii-a), (iii)-(vi) arise in the study of direct-sequence (DS) or multi-carrier (MC) code-division multiple-access (CDMA) systems with isometric signatures over frequency-selective fading channels or space-division multiple-access (SDMA) systems with isometric precoding matrices over flat-fading channels. More precisely, for DS-CDMA systems, the matrices $\mathbf{H}_k$ are either Toeplitz or circular matrices (if a cyclic prefix is used) constructed from the channel impulse response; for MC-CDMA, the matrices $\mathbf{H}_k$ are diagonal and represent the channel frequency response on each sub-carrier; for flat fading SDMA systems, the matrices $\mathbf{H}_k$ can be of arbitrary form and their elements represent the complex channel gains between the transmit and receive antennas. In all cases, the diagonal entries of the matrices $\mathbf{P}_k$ determine the transmit power of each signature (CDMA) or transmit stream (SDMA).

The large system analysis of random i.i.d. and random orthogonal precoded systems with optimal and sub-optimal linear receivers has been the subject of numerous publications. The asymptotic performance of minimum-mean-square-error (MMSE) receivers for the channel model (1) for the case $K = 1, \mathbf{P}_1 = \mathbf{I}_{n_1}$ and $\mathbf{H}_1$ diagonal with i.i.d. elements has been studied in [8] relying on results from free probability theory. This result was extended to frequency-selective fading channels and sub-optimal receivers in [9]. Although not published, the associated mutual information was evaluated in [10] (this result is recalled in [11, Theorem 4.11]). The case of i.i.d. and isometric MC-CDMA over Rayleigh fading channels with multiple signatures per user terminal, i.e., $K \geq 1$ and $\mathbf{H}_k$ diagonal with i.i.d. complex Gaussian entries, was considered in [12], where approximate solutions of the signal-to-noise-plus-interference-ratio (SINR) at the output of the MMSE receiver were provided. Asymptotic expressions for the spectral efficiency of the same model were then derived in [13]. DS-CDMA over flat-fading channels, i.e., $K \geq 1$, $n_k = N$ and $\mathbf{H}_k = \mathbf{I}_N$ for all k, was studied in [14], where the authors derived deterministic equivalents of the Shannon- and η-transform based on the asymptotic freeness [11, Section 3.5] of the matrices $\mathbf{W}_k\mathbf{P}_k\mathbf{W}_k^{\mathsf{H}}$. Besides, a sum-rate maximizing power-allocation algorithm was proposed. Finally, a different approach via incremental matrix expansion [15] led to the exact characterization of the asymptotic SINR of the MMSE receiver for the general channel model (1). However, the previously mentioned works share the underlying assumption that the spectral distributions of the matrices $\mathbf{H}_k$ and $\mathbf{P}_k$ converge to some limiting distributions or the matrices $\mathbf{H}_k\mathbf{H}_k^{\mathsf{H}}$ are jointly diagonalizable.[1] Also, the computation of the asymptotic SINR requires the computation of rather complicated implicit equations. These can be solved in most cases by standard fixed-point algorithms but a proof of convergence to the correct solution was not provided. Finally, a closed-form expression for the asymptotic spectral efficiency is missing, although an approximate solution which requires numerical integration was presented in [13].

The above works assume non-random communication channels and can therefore be only applied to the performance analysis of static or slow fading channels. Turning the matrices $\mathbf{H}_k$ into random matrices instead allows for the study of the ergodic performance of fast fading channels with isometric precoders. The next section discusses the practical applications in this broader context.

[1]That is, there exists a unitary matrix $\mathbf{V}$ such that $\mathbf{V}\mathbf{H}_k\mathbf{H}_k^{\mathsf{H}}\mathbf{V}^{\mathsf{H}}$ is diagonal for all k.

B. Fading channel scenario (hypothesis (ii-b))

The second scenario considers the channel model (1) under assumptions (i), (ii-b), (iii)-(vi). In contrast to the first scenario, the $\mathbf{H}_k$ matrices are now assumed to be random. Thus, we aim at evaluating both the instantaneous performance for a random channel realization and the ergodic performance of these channels. These are appropriate performance measures in fast fading environments.

Of particular interest in this setting is the evaluation of the MIMO channel capacity under random beamforming. In point-to-point MIMO channels, the ergodic channel capacity has been the object of numerous works and is by now well understood [16], [17]. However, the ergodic sum-rate of more involved models, such as the MIMO MAC [4] under individual or sum power constraints, has been studied only recently through the scope of random matrix theory. As a by-product of this work, we will extend the results of [4] to the transmit covariance optimization in the class of scaled identity matrices under sum power constraints. More fundamental is the capacity of MIMO channels with co-channel interference, for which much less is known about the optimal transmission strategies [18], [19]. The first interesting question relates to the problem of how many antennas should be used for transmission and how many independent data streams should be sent, which are the same problem when the channels have i.i.d. entries. With transmit antenna correlation, however, it makes a difference which antennas are selected for transmission and the question of the optimal number of antennas to be used becomes a combinatorial problem. To circumvent this issue, random beamforming can be used. The remaining question is then how many orthogonal streams should be sent, using *all* available antennas. This is one of the key motivations of this article, as our results enable the evaluation of the sum-rate of systems composed of multiple transmitter-receiver pairs, each applying random isotropic beamforming.

In summary, regardless of the specific application scenario of the model (1), unitary precoders have gained significant interest in wireless communications [20] (see also the recent work on spatial multiplexing systems [21] and limited feedback beamforming solutions in future wireless standards [22]). Thus, the performance evaluation of isometric precoded systems is compulsory and a field of active research [23].

C. Contributions

The object of this article is to propose a new framework for the analysis of large random matrix models involving Haar matrices using the *Stieltjes-transform method*. This method is considered today as one of the most practical and powerful tools for handling large random matrices in wireless communications research. Our analysis is fundamentally based on a trace lemma for Haar matrices first provided in [8] and recalled in Lemma 5 (Appendix F). Unlike previous contributions, we dismiss most of the practical constraints of free probability theory, combinatorial and incremental matrix expansion methods, such as the need for spectral limits of the deterministic matrices in the model to exist, or the need for the matrices $\mathbf{H}_k\mathbf{H}_k^{\mathsf{H}}$ to be diagonalizable in a common eigenvector basis. The expressions we derive appear to be very similar to previously derived expressions when the precoding matrices $\mathbf{W}_k$ have i.i.d. entries instead of being Haar distributed (see in particular Remark 2). This allows for a unified understanding of both models with i.i.d. or Haar matrices. As a consequence, we believe that the generality of the theoretical results presented in this article, supported by a large scope of application contexts, might stimulate further related research.

Before summarizing our main contributions, we introduce some definitions which will be of repeated use. The central object of interest is the matrix $\mathbf{B}_N \in \mathbb{C}^{N\times N}$, defined as

$$\mathbf{B}_N = \sum_{k=1}^{K} \mathbf{H}_k \mathbf{W}_k \mathbf{P}_k \mathbf{W}_k^{\mathsf{H}} \mathbf{H}_k^{\mathsf{H}}.$$

We denote by $I_N(\sigma^2)$ the normalized mutual information of the channel (1), given by [24]

$$I_N(\sigma^2) = \frac{1}{N} \log\det\left(\mathbf{I}_N + \frac{1}{\sigma^2}\mathbf{B}_N\right) \quad \text{(nats/s/Hz)}.$$

We further denote by $\gamma_{kj}^N(\sigma^2)$ the SINR at the output of the linear MMSE detector for the jth component of the transmit vector $\mathbf{x}_k$, which reads [25]

$$\gamma_{kj}^N(\sigma^2) = p_{kj} \mathbf{w}_{kj}^{\mathsf{H}} \mathbf{H}_k^{\mathsf{H}} \left(\mathbf{B}_{N(k,j)} + \sigma^2 \mathbf{I}_N\right)^{-1} \mathbf{H}_k \mathbf{w}_{kj}$$

where $\mathbf{B}_{N(k,j)} = \mathbf{B}_N - p_{kj}\mathbf{H}_k \mathbf{w}_{kj}\mathbf{w}_{kj}^{\mathsf{H}}\mathbf{H}_k^{\mathsf{H}}$ and $\mathbf{w}_{kj}$ is the jth column of $\mathbf{W}_k$. We then define the normalized sum-rate with MMSE detection as

$$R_N(\sigma^2) = \frac{1}{N}\sum_{k=1}^{K}\sum_{j=1}^{n_k} \log\left(1 + \gamma_{kj}^N(\sigma^2)\right).$$

Depending on whether we consider the quasi-static channel scenario (ii-a) or the fading channel scenario (ii-b), we will sometimes differentiate between $I_N^{(a)}(\sigma^2)$ and $I_N^{(b)}(\sigma^2)$, the mutual information under (ii-a) and (ii-b), respectively. The same holds for $\gamma_{kj}^N(\sigma^2)$ and $R_N(\sigma^2)$.

The technical contributions of this paper are as follows: We derive deterministic approximations $\bar{I}_N(\sigma^2)$, $\bar{\gamma}_{kj}^N(\sigma^2)$, and $\bar{R}_N(\sigma^2)$ of $I_N(\sigma^2)$, $\gamma_{jk}^N(\sigma^2)$, and $R_N(\sigma^2)$, respectively, which are (almost surely) asymptotically tight as the system dimensions N, N_i, n_i grow large at the same rate (denoted simply $N \to \infty$). These approximations, called *deterministic equivalents*, are easy to compute as they are shown to be the limits of simple (provably converging) fixed-point algorithms, they are given in closed form and do not require any numerical integration, and they require only very general conditions on the matrices $\mathbf{H}_k$ and $\mathbf{P}_k$.

We then present several applications of our results to wireless communications. First, we consider a cellular uplink orthogonal SDMA communication model with inter-cell interference, assuming independent codes in adjacent cells and quasi-static channels at all communication pairs. We then study a MIMO multiple access channel (MAC) from several multi-antenna transmitters to a multi-antenna receiver under the fading channel scenario (hypothesis (ii-b)). The transmitters are unaware of the channel realizations and send an arbitrary number of independent data streams using isometric random beamforming vectors. The receiver is assumed to be aware of all instantaneous channel realizations and beamforming vectors. Under this setting, we derive the optimal power allocation under individual or sum-power constraints which can be computed by an iterative water-filling algorithm. Finally, we address the problem of finding the optimal number of independent streams to be transmitted in a two-by-two interference channel. Although the use of deterministic approximations in this context requires an exhaustive search over all possible stream-configurations, it is computationally much less expensive than Monte Carlo simulations. Extensions to more than two transmit-receive pairs and possible different objective functions, e.g. weighted sum-rate or sum-rate with MMSE decoding, are straightforward and not presented.

For all these applications, numerical simulations show that the deterministic approximations are very tight even for small system dimensions. In the interference channel model, these simulations suggest in particular that, at low SNR, it is optimal to use all streams while, at high SNR, stream-control, i.e. transmitting less than the maximal number of streams, is beneficial.

Our work also constitutes a novel contribution to the field of random matrix theory, as we introduce new proof techniques based on the Stieltjes transform method in the context of random isometric matrices. Namely, we provide in Theorem 7 (Appendix A) a deterministic equivalent $\bar{F}_N$ of the empirical spectral distribution (e.s.d.) F_N of $\mathbf{B}_N$ (see Appendix A for a definition of e.s.d.). That is, $\bar{F}_N$ is such that, as $N \to \infty$, $F_N - \bar{F}_N \Rightarrow 0$, this convergence being valid almost surely. Although deterministic equivalents of e.s.d. are by now more or less standard and have been developed for rather involved random matrix models [5], [4], [1], results for the case of isometric (Haar) matrices are still an exception. In particular, most results on Haar matrices are based on the assumption of *asymptotic freeness* of the underlying matrices, a requirement which is rarely met for the matrices in the channel model (1) of interest here. The approach taken in this work is therefore novel as it does not rely on free probability theory [26], [27] and we do not require any of the matrices in (1) to be asymptotically free. Interestingly, a very recent extension of free probability theory, coined free deterministic equivalents [28], has come as a response to the present article in which free probability tools are developed to tackle the aforementioned limitations.

The remainder of this article is structured as follows: in Section II, we introduce the main results of this work, the proofs of which are postponed to the appendices. In Section III, the results are applied to the practical wireless communication models discussed above. Section IV then concludes the article.

Notations: Boldface lower and upper case symbols represent vectors and matrices, respectively. $\mathbf{I}_N$ is the size-N identity matrix and diag$(x_1, \ldots, x_N)$ is a diagonal matrix with elements x_i. The trace, transpose and Hermitian transpose operators are denoted by 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. 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 $[0, \infty)$ and by $\mathbb{C}_+$ the set $\{z \in \mathbb{C}, \text{Im}[z] > 0\}$. Denote by $\mathcal{C}(X, Y)$ the set of continuous functions from $X \subset \mathbb{C}$ to $Y \subset \mathbb{C}$, by $\mathcal{H}(X, Y)$ the set of holomorphic functions from $X \subset \mathbb{C}$ to $Y \subset \mathbb{C}$, and by $\mathcal{S}(X)$ the class *Stieltjes transforms* of finite measures supported by $X \subset \mathbb{R}$ (see Definition 1).

II. Main results

In this section, we present the main results of the article. All proofs are deferred to the appendix. We will distinguish the results for the quasi-static and the fading channel scenarios. Since we will make limiting considerations as the system dimensions grow large, some technical assumptions will be necessary:

A1 The notation $N \to \infty$ denotes the simultaneous growth of N, N_i, n_i for all i, in such a way that $0 \leq c_i = \frac{n_i}{N_i} \leq 1$ and $0 < \liminf_N \bar{c}_i = \frac{N_i}{N} \leq \limsup_N \bar{c}_i < \infty$.

For all convergence results in this paper (as $N \to \infty$), the matrices $\mathbf{P}_k = \mathbf{P}_k(N) \in \mathbb{R}_+^{n_k \times n_k}$, $\mathbf{H}_k = \mathbf{H}_k(N) \in \mathbb{C}^{N \times N_k}$ (as well as the $\mathbf{R}_{kj} = \mathbf{R}_{kj}(N) \in \mathbb{C}^{N \times N}$ under assumption (ii-a)), and $\mathbf{W}_k = \mathbf{W}_k(N) \in \mathbb{C}^{N_K \times n_k}$ should be understood as sequences of (random) matrices with growing dimensions. Wherever this is clear from the context, we drop the dependence on N to simplify the notations.

In order to control the power loading matrices as the system grows large, we need the following assumption:

A2 There exists $P > 0$ such that, for all k, $\limsup_N \|\mathbf{P}_k\| \leq P$.

Under (ii-a), the channel gains will need to remain bounded for all large N:

A3-a There exists $R > 0$ such that $\max_k \limsup_N \|\mathbf{R}_k\| \leq R$, where we recall that $\mathbf{R}_k = \mathbf{H}_k \mathbf{H}_k^{\mathsf{H}}$.

The equivalent constraint under (ii-b) is that the channel correlations remain bounded for all large N:

A3-b There exists $R > 0$ such that $\limsup_N \|\mathbf{R}_{kj}\| \leq R$ for all j, k.

Due to some technical issues, it will be sometimes necessary to require the following condition:

A4 For all random matrices $\mathbf{H}_k$ within a set of probability one, there exists $M > 0$ such that $\max_k \|\mathbf{H}_k \mathbf{H}_k^{\mathsf{H}}\| < M$ for all large N.

Assumption **A4** is met in particular in the situation when there exists $m > 0$, such that for all k, j, N, $\mathbf{R}_{kj} \in \mathcal{R}_N$ with $\mathcal{R}_N$ a discrete set of cardinality $|\mathcal{R}_N| < m$ for all N (see the arguments in [4]). For example, this holds true for the scenario of a common correlation matrix at each receiver, i.e., $\mathbf{R}_{kj} = \bar{\mathbf{R}}_k$ are equal for all j.

A. Fundamental Equations

We first introduce the *fundamental equations* for model (1). These equations provide the core deterministic quantities that will define the deterministic equivalents for $I_N(\sigma^2)$, $\gamma_{ij}^N(\sigma^2)$, and $R_N(\sigma^2)$.

Theorem 1 (Fundamental equations under (ii-a)): Consider the system model (1) under assumptions (i), (ii-a), (iii)-(vi). Let $\sigma^2 > 0$. Then the following system of implicit equations

$$\bar{a}_k(\sigma^2) = \frac{1}{N} \operatorname{tr} \mathbf{P}_k \left(a_k(\sigma^2) \mathbf{P}_k + [\bar{c}_k - a_k(\sigma^2)\bar{a}_k(\sigma^2)] \mathbf{I}_{n_k} \right)^{-1}$$
$$a_k(\sigma^2) = \frac{1}{N} \operatorname{tr} \mathbf{R}_k \left(\sum_{j=1}^{K} \bar{a}_j(\sigma^2) \mathbf{R}_j + \sigma^2 \mathbf{I}_N \right)^{-1} \tag{4}$$

with $k \in \{1, \dots, K\}$, admits a unique solution such that, for all k, $a_k(\sigma^2), \bar{a}_k(\sigma^2) \geq 0$, and $0 \leq a_k(\sigma^2)\bar{a}_k(\sigma^2) < c_k \bar{c}_k$. Moreover, this solution is obtained explicitly by the following fixed-point algorithm

$$\bar{a}_k(\sigma^2) = \lim_{t \to \infty} \bar{a}_k^{(t)}(\sigma^2), \quad \bar{a}_k^{(t)}(\sigma^2) = \lim_{l \to \infty} \bar{a}_k^{(t,l)}(\sigma^2)$$

where, for $k \in \{1, \ldots, K\}$,

$$a_k^{(t)}(\sigma^2) = \frac{1}{N} \operatorname{tr} \mathbf{R}_k \left(\sum_{j=1}^{K} \bar{a}_j^{(t-1)}(\sigma^2) \mathbf{R}_j + \sigma^2 \mathbf{I}_N \right)^{-1}$$

$$\bar{a}_k^{(t,l)}(\sigma^2) = \frac{1}{N} \operatorname{tr} \mathbf{P}_k \left(a_k^{(t)}(\sigma^2) \mathbf{P}_k + [\bar{c}_k - a_k^{(t)}(\sigma^2) \bar{a}_k^{(t,l-1)}(\sigma^2)] \mathbf{I}_{n_k} \right)^{-1}$$

with arbitrary initial values $\bar{a}_k^{(t,0)}(\sigma^2) \in [0, c_k \bar{c}_k / a_k^{(t)}(\sigma^2))$ and $a_k^{(0)}(\sigma^2) = 1$.

Proof: The proof is provided in Appendix A. ■

Remark 2: Assume $\bar{c}_k = 1$ for every k (e.g., when $\mathbf{H}_k$ is a Toeplitz matrix as in the CDMA case). Then, extending every $\mathbf{P}_k \in \mathbb{C}^{n_k \times n_k}$ into $N \times N$ matrices filled with zeros, we can assume $c_k = 1$ without affecting the final result. In this scenario, the fundamental equations (1) under (ii-a) become

$$\bar{a}_k(\sigma^2) = \frac{1}{N} \operatorname{tr} \mathbf{P}_k \left(a_k(\sigma^2) \mathbf{P}_k + [1 - a_k(\sigma^2) \bar{a}_k(\sigma^2)] \mathbf{I}_N \right)^{-1} \tag{5}$$

$$a_k(\sigma^2) = \frac{1}{N} \operatorname{tr} \mathbf{R}_k \left(\sum_{j=1}^{K} \bar{a}_j(\sigma^2) \mathbf{R}_j - z \mathbf{I}_N \right)^{-1}.$$

This can be compared to the scenario where the matrices $\mathbf{W}_k$, instead of being Haar matrices, have i.i.d. entries of variance $1/N$. The fundamental equations of this model were derived in [4, Corollary 1] and are given as follows:

$$\underline{\bar{a}}_k(\sigma^2) = \frac{1}{N} \operatorname{tr} \mathbf{P}_k \left(\underline{a}_k(\sigma^2) \mathbf{P}_k + \mathbf{I}_N \right)^{-1} \tag{6}$$

$$\underline{a}_k(\sigma^2) = \frac{1}{N} \operatorname{tr} \mathbf{R}_k \left(\sum_{j=1}^{K} \underline{\bar{a}}_j(\sigma^2) \mathbf{R}_j + \sigma^2 \mathbf{I}_N \right)^{-1}$$

such that $\underline{a}_k(\sigma^2)$ is positive for all k. The scalars $\underline{a}_k(\sigma^2)$ and $\underline{\bar{a}}_k(\sigma^2)$ are also defined as the limits of a classical fixed-point algorithm. The only difference between the two sets of equations lies in the additional term $-a_k \bar{a}_k \mathbf{I}_N$ in (5), not present in (6).

We now turn to the fundamental equations in the fading channel context.

Theorem 2 (Fundamental equations under (ii-b)): Consider the system model (1) under assumptions (i), (ii-b), (iii)-(vi). Let $\sigma^2 > 0$. Then, the following system of implicit equations

$$\bar{b}_k(\sigma^2) = \frac{1}{N} \operatorname{tr} \mathbf{P}_k \Big(b_k(\sigma^2) \mathbf{P}_k + \left[\bar{c}_k - b_k(\sigma^2) \bar{b}_k(\sigma^2) \right] \mathbf{I}_{n_k} \Big)^{-1}$$

$$b_k(\sigma^2) = \frac{1}{N} \sum_{j=1}^{N_k} \frac{\zeta_{kj}(\sigma^2)}{1 + \bar{b}_k(\sigma^2) \zeta_{kj}(\sigma^2)}$$

$$\zeta_{kj}(\sigma^2) = \frac{1}{N} \operatorname{tr} \mathbf{R}_{kj} \left(\frac{1}{N} \sum_{k=1}^{K} \sum_{j=1}^{N_k} \frac{\bar{b}_k(\sigma^2) \mathbf{R}_{k,j}}{1 + \bar{b}_k(\sigma^2) \zeta_{kj}(\sigma^2)} + \sigma^2 \mathbf{I}_N \right)^{-1}, \quad j \in \{1, \ldots, N_k\}$$

with $k \in \{1, \ldots, K\}$, admits a unique solution satisfying $\zeta_{kj}(\sigma^2), b_k(\sigma^2), \bar{b}_k(\sigma^2) \geq 0$ and $0 \leq b_k(\sigma^2) \bar{b}_k(\sigma^2) < c_k \bar{c}_k$ for all k, j. Moreover, this solution is given explicitly by the following fixed-point algorithm

$$b_k(\sigma^2) = \lim_{t \to \infty} g_k^{(t)}(\sigma^2), \quad \bar{b}_k(\sigma^2) = \lim_{t \to \infty} \bar{b}_k^{(t)}(\sigma^2), \quad \zeta_{kj}(\sigma^2) = \lim_{t \to \infty} \zeta_{kj}^{(t)}(\sigma^2)$$

where

$$\bar{b}_k^{(t)}(\sigma^2) = \lim_{l\to\infty} \bar{b}_k^{(t,l)}(\sigma^2), \quad \zeta_{kj}^{(t)}(\sigma^2) = \lim_{l\to\infty} \zeta_{kj}^{(t,l)}(\sigma^2)$$

$$b_k^{(t)}(\sigma^2) = \frac{1}{N}\sum_{j=1}^{N_k} \frac{\zeta_{kj}^{(t)}(\sigma^2)}{1+\bar{b}_k^{(t-1)}(\sigma^2)\zeta_{kj}^{(t)}(\sigma^2)}$$

$$\bar{b}_k^{(t,l)}(\sigma^2) = \frac{1}{N}\mathrm{tr}\mathbf{P}_k\left(b_k^{(t-1)}(\sigma^2)\mathbf{P}_k + \left[\bar{c}_k - b_k^{(t-1)}(\sigma^2)\bar{b}_k^{(t,l-1)}(\sigma^2)\right]\mathbf{I}_{n_k}\right)^{-1}$$

$$\zeta_{kj}^{(t,l)}(\sigma^2) = \frac{1}{N}\mathrm{tr}\mathbf{R}_{kj}\left(\frac{1}{N}\sum_{k=1}^{K}\sum_{j=1}^{N_k}\frac{\bar{b}_k^{(t-1)}(\sigma^2)\mathbf{R}_{k,j}}{1+\bar{b}_k^{(t-1)}(\sigma^2)\zeta_{kj}^{(t,l-1)}(\sigma^2)} + \sigma^2\mathbf{I}_N\right)^{-1}$$

with the initial values $\zeta_{kj}^{(t,0)}(\sigma^2) = 1/\sigma^2$, $\bar{b}_k^{(t,0)} \in [0, c_k\bar{c}_k/b_k^{(t-1)}(\sigma^2))$ and $b_k^{(0)}(\sigma^2) = 1$ for all k, j.

Proof: The proof is provided in Appendix D. ■

B. System performance

The following results are all based on the fundamental equations of Theorem 1 and Theorem 2.

Theorem 3 (Mutual information under (ii-a)): Consider the system model (1) under assumptions (i), (ii-a), (iii)-(vi), and denote, for $\sigma^2 > 0$,

$$I_N^{(a)}(\sigma^2) = \frac{1}{N}\log\det\left(\mathbf{I}_N + \frac{1}{\sigma^2}\mathbf{B}_N\right).$$

Assume **A1**, **A2**, and **A3**-a. Then, as $N \to \infty$,

$$\mathbb{E}I_N^{(a)}(\sigma^2) - \bar{I}_N^{(a)}(\sigma^2) \to 0$$
$$I_N^{(a)}(\sigma^2) - \bar{I}_N^{(a)}(\sigma^2) \xrightarrow{\text{a.s.}} 0$$

where

$$\bar{I}_N^{(a)}(\sigma^2) = \frac{1}{N}\log\det\left(\mathbf{I}_N + \frac{1}{\sigma^2}\sum_{k=1}^{K}\bar{a}_k\mathbf{R}_k\right) + \sum_{k=1}^{K}\left[\frac{1}{N}\log\det\left([\bar{c}_k - a_k\bar{a}_k]\mathbf{I}_{n_k} + a_k\mathbf{P}_k\right) + (1-c_k)\bar{c}_k\log(\bar{c}_k - a_k\bar{a}_k) - \bar{c}_k\log(\bar{c}_k)\right] \quad (7)$$

with $a_k = a_k(\sigma^2)$, $\bar{a}_k = \bar{a}_k(\sigma^2)$, $k \in \{1,\ldots,K\}$, given by Theorem 1.

Proof: The proof is provided in Appendix B. ■

Theorem 4 (Mutual information under (ii-b)): Consider the system model (1) under assumptions (i), (ii-b), (iii)-(vi), and denote, for $\sigma^2 > 0$

$$I_N^{(b)}(\sigma^2) = \frac{1}{N}\log\det\left(\mathbf{I}_N + \frac{1}{\sigma^2}\mathbf{B}_N\right).$$

Assume **A1**, **A2**, **A3**-b, and **A4**. Let $\bar{b}_k = \bar{b}_k(\sigma^2)$, $b_k = b_k(\sigma^2)$ and $\zeta_{kj} = \zeta_{kj}(\sigma^2)$ for all k, j be defined as in Theorem 2. Then, as $N \to \infty$,

$$\mathbb{E}I_N^{(b)}(\sigma^2) - \bar{I}_N^{(b)}(\sigma^2) \to 0$$
$$I_N^{(b)}(\sigma^2) - \bar{I}_N^{(b)}(\sigma^2) \xrightarrow{\text{a.s.}} 0$$

where

$$\bar{I}_N^{(b)}(\sigma^2) = \bar{V}_N(\sigma^2) + \frac{1}{N}\sum_{k=1}^{K} \log\det\left(\left[\bar{c}_k - b_k\bar{b}_k\right]\mathbf{I}_{n_k} + b_k\mathbf{P}_k\right) + \sum_{k=1}^{K}(1-c_k)\bar{c}_k\log(\bar{c}_k - b_k\bar{b}_k) - \bar{c}_k\log(\bar{c}_k)$$

$$\bar{V}_N(\sigma^2) = \frac{1}{N}\log\det\left(\mathbf{I}_N + \frac{1}{\sigma^2}\frac{1}{N}\sum_{k=1}^{K}\sum_{j=1}^{N_k}\frac{\bar{b}_k\mathbf{R}_{k,j}}{1+\bar{b}_k\zeta_{kj}}\right) - \sum_{k=1}^{K}\bar{b}_k b_k + \frac{1}{N}\sum_{k=1}^{K}\sum_{j=1}^{N_k}\log\left(1+\bar{b}_k\zeta_{kj}\right). \tag{8}$$

Proof: The proof is provided in Appendix E. ■

Before moving to approximations of the SINR at the output of the MMSE receiver, we provide hereafter a mutual information maximizing power allocation scheme of practical interest in multiple access channels. We will in particular differentiate between the scenario (I) in which each transmitter has its own power allocation policy, i.e. the total power transmitted across its antennas cannot exceed a given threshold, and the scenario (II) in which the total power transmitted across all antennas is less than a threshold.

Proposition 1 (Optimal power allocation): Consider the system model (1) under assumptions (i), (ii-b), (iii)-(vi). Let $\sigma^2 > 0$ and $\bar{I}^{(b)}(\sigma^2)$ be defined as in Theorem 4 and let $P, P_1, \ldots, P_K \geq 0$. Then, the solution to the following optimization problem:

$$\begin{aligned}(\bar{\mathbf{P}}_1^\star, \ldots, \bar{\mathbf{P}}_K^\star) &= \arg\max_{\mathbf{P}_1,\ldots,\mathbf{P}_K} \bar{I}_N(\sigma^2) \\ &\quad \text{s.t.} \begin{cases} \frac{1}{n_k}\mathrm{tr}\mathbf{P}_k \leq P_k \ \forall k & (I) \\ \sum_{k=1}^{K}\frac{1}{n_k}\mathrm{tr}\mathbf{P}_k \leq P & (II)\end{cases}\end{aligned}$$

is given by

$$\bar{\mathbf{P}}_k^\star = \bar{p}_k^\star \mathbf{I}_{n_K}$$

where

$$\bar{p}_k^\star = \begin{cases} P_k & (I) \\ \left(\bar{b}_k^\star - \frac{\bar{c}_k}{b_k^\star} + \frac{c_k\bar{c}_k}{\lambda}\right)^+ & (II)\end{cases}$$

for all k, with $b_k^\star = b_k^\star(\sigma^2), \bar{b}_k^\star = \bar{b}_k^\star(\sigma^2)$ given by Theorem 2 when $\mathbf{P}_k = \bar{\mathbf{P}}_k^\star$, and λ in (II) is chosen such that $\sum_{k=1}^{K}\frac{1}{n_k}\mathrm{tr}\bar{\mathbf{P}}_k^\star = P$. Moreover, let

$$\begin{aligned}(\mathbf{P}_1^\star, \ldots, \mathbf{P}_K^\star) &= \arg\max_{\mathbf{P}_1,\ldots,\mathbf{P}_K} \mathbb{E}I_N(\sigma^2) \\ &\quad \text{s.t.} \begin{cases} \frac{1}{n_k}\mathrm{tr}\mathbf{P}_k \leq P_k \ \forall k & (I) \\ \sum_{k=1}^{K}\frac{1}{n_k}\mathrm{tr}\mathbf{P}_k \leq P & (II)\end{cases}\end{aligned} \tag{9}$$

and assume **A1**, **A2**, **A3**-b, and **A4**. Then,

$$\mathbb{E}I_N^{(b)}(\mathbf{P}_1^\star, \ldots, \mathbf{P}_K^\star) - \bar{I}_N^{(b)}(\bar{\mathbf{P}}_1^\star, \ldots, \bar{\mathbf{P}}_K^\star) \xrightarrow{\text{a.s.}} 0.$$

Proof: The proof is provided in Appendix E. ■

Remark 3: The optimal power allocation matrices $\bar{\mathbf{P}}_k^\star$ under a sum-power constraint (II) can be computed by the iterative water-filling algorithm below. Although we cannot prove the sure convergence of this algorithm (see [7] for a related discussion), we know that if it converges, it achieves the correct solution. In our simulations, we could not create a case in which it did not converge.

Algorithm 1 Iterative water-filling algorithm

1: Let $\epsilon > 0$, $t = 0$ and $\bar{p}_{kj}^{(0)} = P_k$ for all k, j.
2: **repeat**
3: For all k, compute $\bar{b}_k^{(t)}$ and $b_k^{(t)}$ according to Theorem 2 for the matrices $\mathbf{P}_k = \mathrm{diag}\left(\bar{p}_{kj}^{(t)}\right)$.
4: For all k, j, calculate $\bar{p}_{kj}^{(t+1)} = \left(\bar{b}_k^{(t)} - \frac{\bar{c}_k}{b_k^{(t)}} + \frac{c_k \bar{c}_k}{\lambda}\right)^+$, with λ such that $\sum_{k=1}^K \frac{1}{n_k} \sum_{j=1}^{n_k} \bar{p}_{kj}^{(t+1)} = P$.
5: $t = t + 1$
6: **until** $\max_{k,j} |\bar{p}_{k,j}^{(t)} - \bar{p}_{k,j}^{(t-1)}| \leq \epsilon$

Remark 4: The optimal power allocation also shows that sending as many independent data streams as transmit antennas is optimal to maximize the ergodic mutual information. In this scenario, $\mathbf{P}_k$ becomes a scaled identity matrix. The precoder $\mathbf{W}_k$ is now of no practical use and we fall back to a standard MAC channel model. In this case, Proposition 1 can be seen as the optimal power allocation in the class of scaled identity matrices. This optimality is no longer valid in the case of interference channels as will be discussed later on.

Theorem 5 (SINR of the MMSE detector under (ii-a)): Consider the system model (1) under assumptions (i), (ii-a), (iii)-(vi) and, for $\sigma^2 > 0$, denote

$$\gamma_{kj}^{N(a)}(\sigma^2) = p_{kj} \mathbf{w}_{kj}^{\mathsf{H}} \mathbf{H}_k^{\mathsf{H}} \left(\mathbf{B}_{N(k,j)} + \sigma^2 \mathbf{I}_N\right)^{-1} \mathbf{H}_k \mathbf{w}_{kj}. \tag{10}$$

Assume **A1**, **A2**, and **A3**-a. Then, as $N \to \infty$,

$$\gamma_{kj}^{N(a)}(\sigma^2) - \bar{\gamma}_{kj}^{N(a)}(\sigma^2) \xrightarrow{\text{a.s.}} 0$$

where

$$\bar{\gamma}_{kj}^{N(a)}(\sigma^2) = \frac{p_{kj} a_k}{\bar{c}_k - a_k \bar{a}_k}$$

with $a_k = a_k(\sigma^2)$ and $\bar{a}_k = \bar{a}_k(\sigma^2)$ defined in Theorem 1.

Proof: The proof is provided in Appendix C. ■

As an (almost immediate) corollary, we have

Corollary 1: Under the conditions of Theorem 5, denote

$$R_N^{(a)}(\sigma^2) = \frac{1}{N} \sum_{k=1}^{K} \sum_{j=1}^{n_k} \log\left(1 + \gamma_{kj}^{N(a)}(\sigma^2)\right).$$

Then,

$$\mathbb{E}R_N^{(a)}(\sigma^2) - \bar{R}_N^{(a)}(\sigma^2) \to 0$$
$$R_N^{(a)}(\sigma^2) - \bar{R}_N^{(a)}(\sigma^2) \xrightarrow{\text{a.s.}} 0$$

where

$$\bar{R}_N^{(a)}(\sigma^2) = \frac{1}{N}\sum_{k=1}^{K}\sum_{j=1}^{n_k} \log\left(1 + \bar{\gamma}_{kj}^{N(a)}(\sigma^2)\right).$$

Proof: The proof is provided in Appendix C. ■

Theorem 6 (SINR of the MMSE detector under (ii-b)): Consider the system model (1) under assumptions (i), (ii-b), (iii)-(vi) and, for $\sigma^2 > 0$, denote

$$\gamma_{kj}^{N(b)}(\sigma^2) = p_{kj}\mathbf{w}_{kj}^{\mathsf{H}}\mathbf{H}_k^{\mathsf{H}}\left(\mathbf{B}_{N(k,j)} + \sigma^2\mathbf{I}_N\right)^{-1}\mathbf{H}_k\mathbf{w}_{kj}.$$

Assume **A1**, **A2**, **A3**-b, and **A4**. Then, as $N \to \infty$,

$$\gamma_{kj}^{N(b)}(\sigma^2) - \bar{\gamma}_{kj}^{N(b)}(\sigma^2) \xrightarrow{\text{a.s.}} 0$$

where

$$\bar{\gamma}_{kj}^{N(b)}(\sigma^2) = \frac{p_{kj}b_k}{\bar{c}_k - b_k\bar{b}_k}$$

with $b_k = b_k(\sigma^2)$ and $\bar{b}_k = \bar{b}_k(\sigma^2)$, given by Theorem 2.

Proof: The proof is provided in Appendix E. ■

Similar to the quasi-static channel scenario, we also have the following corollary.

Corollary 2: Under the conditions of Theorem 6, denote

$$R_N^{(b)}(\sigma^2) = \frac{1}{N}\sum_{k=1}^{K}\sum_{j=1}^{n_k} \log\left(1 + \gamma_{kj}^{N(b)}(\sigma^2)\right).$$

Then,

$$\mathbb{E}R_N^{(b)}(\sigma^2) - \bar{R}_N^{(b)}(\sigma^2) \to 0$$
$$R_N^{(b)}(\sigma^2) - \bar{R}_N^{(b)}(\sigma^2) \xrightarrow{\text{a.s.}} 0$$

where

$$\bar{R}_N^{(b)}(\sigma^2) = \frac{1}{N}\sum_{k=1}^{K}\sum_{j=1}^{n_k} \log\left(1 + \bar{\gamma}_{kj}^{N(b)}(\sigma^2)\right).$$

Proof: The proof is provided in Appendix E ■

Remark 5: Under scenario (ii-b), for the special case K=1, $\mathbf{P}_1 = \mathbf{I}_{n_1}$, $N_1 = n_1 = N$ and $\mathbf{R}_{1j} = \mathbf{I}_N$ for all j, the set of implicit equations in Theorem 2 reduces to:

$$\bar{b}(\sigma^2) = \frac{1}{1 - g(\sigma^2)\bar{b}(\sigma^2) + g(\sigma^2)}, \quad g(\sigma^2) = \frac{\zeta(\sigma^2)}{1 + \bar{b}(\sigma^2)\zeta(\sigma^2)}, \quad \zeta(\sigma^2) = \frac{1}{\frac{\bar{b}(\sigma^2)}{1+\bar{b}(\sigma^2)\zeta(\sigma^2)} + \sigma^2}.$$

Note that

$$0 = 1 - \frac{1 - g(\sigma^2)\bar{b}(\sigma^2) + g(\sigma^2)}{1 - g(\sigma^2)\bar{b}(\sigma^2) + g(\sigma^2)} = 1 - \left[1 - g(\sigma^2)\bar{b}(\sigma^2)\right]\bar{b}(\sigma^2) - g(\sigma^2)\bar{b}(\sigma^2) = \left[1 - g(\sigma^2)\bar{b}(\sigma^2)\right](1 - \bar{b}(\sigma^2))$$

which implies $\bar{b}(\sigma^2) = 1$ since $1 - g(\sigma^2)\bar{b}(\sigma^2) > 0$ by definition. Thus, the last equations further simplify to

$$g(\sigma^2) = \frac{\zeta(\sigma^2)}{1 + \zeta(\sigma^2)}, \qquad \zeta(\sigma^2) = \frac{1}{\frac{1}{1+\zeta(\sigma^2)} + \sigma^2}$$

which has a unique solution satisfying $\zeta(\sigma^2) \geq 0$ and that can be given in closed-form:

$$\zeta(\sigma^2) = \frac{-1 + \sqrt{1 + \frac{4}{\sigma^2}}}{2}.$$

Note that $\zeta(\sigma^2)$ is the Stieltjes transform of the Marčenko-Pastur law with scale parameter 1 [11, Equation (3.20)] evaluated on the negative real axis. This result is consistent with our expectations since $\mathbf{B}_N = \mathbf{Z}_1\mathbf{Z}_1^{\mathsf{H}}$, where $\mathbf{Z}_1 \in \mathbb{C}^{N\times N}$ has i.i.d. entries with zero mean and variance $1/N$. Moreover, the expression of the normalized asymptotic mutual information as given in Theorem 4 reduces to

$$\bar{I}_N^{(b)}(\sigma^2) = \bar{V}_N(\sigma^2) = \log\left(1 + \zeta(\sigma^2) + 1/\sigma^2\right) - \frac{\zeta(\sigma^2)}{1 + \zeta(\sigma^2)}$$

which is consistent with the asymptotic spectral efficiency of a Rayleigh-fading $N \times N$ MIMO channel [29, Equation (9)] (see also [11, Section 13.2.2]). Equivalently, the asymptotic SINR of the MMSE detector and the associated normalized sum-rate can be given as (cf. [29, Proposition VI.1]):

$$\bar{\gamma}_j^{N(b)} = \zeta(\sigma^2), \qquad \bar{R}_N^{(b)}(\sigma^2) = \log(1 + \zeta(\sigma^2)).$$

Remark 6: Technically, the results obtained for the quasi-static scenario unfold from the Stieltjes transform framework very similar to [4], [5]. However, some new tools are introduced which simplify the analysis made in these papers, such as the method of standard interference functions to prove existence and uniqueness of the derived deterministic equivalents. As for the results in the fading channel scenario, they unfold from the conjugation of the results obtained in the quasi-static scenario and the results obtained in [1] (recalled in Appendix G) for a channel model similar to (1) but without the presence of the $\mathbf{W}_k$ matrices. The central tool to allow this conjugation is the Tonelli (or Fubini) theorem, Lemma 9, on the product probability space engendering both the $\mathbf{W}_k$ and $\mathbf{H}_k$ matrices.

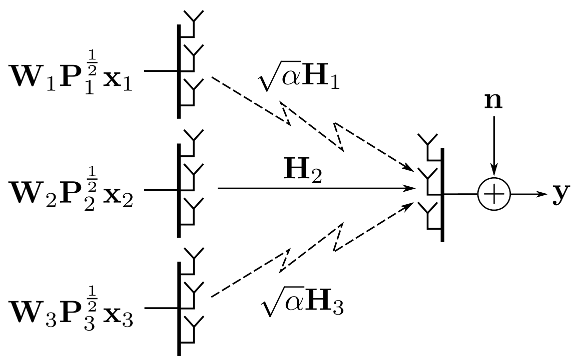

Fig. 1. Three-cell example: The BS in the center cell decodes the n streams from the UT in its own cell while treating the other signals as interference.

III. Numerical results

The results of Section II enable a simple characterization of different performance measures of isometric precoded multi-user systems with large dimensional quasi-static or fading channels, some of which were introduced in Section I. In the following, we apply these results to three practical examples.

A. Uplink orthogonal SDMA with inter-cell interference

In this first example, we apply the theoretical results of Section II under the quasi-static channel scenario (hypothesis (ii-a)) to the uplink channel of an orthogonal SDMA scheme with inter-cell interference. We consider a three cell system with one active user terminal (UT) per cell. The UT in cell k is equipped with N_k transmit antennas. We focus on the central cell, whose base station (BS) is equipped with N antennas, and assume that the the signals received from neighboring cells are treated as noise. This setup is schematically depicted in Figure 1. The received signal $\mathbf{y}$ at the BS reads

$$\mathbf{y} = \mathbf{H}_2\mathbf{W}_2\mathbf{P}_2^{\frac{1}{2}}\mathbf{x}_2 + \underbrace{\sqrt{\alpha}\mathbf{H}_1\mathbf{W}_1\mathbf{P}_1^{\frac{1}{2}}\mathbf{x}_1 + \sqrt{\alpha}\mathbf{H}_3\mathbf{W}_3\mathbf{P}_3^{\frac{1}{2}}\mathbf{x}_3 + \mathbf{n}}_{\triangleq \mathbf{z}}$$

with $\mathbf{H}_i \in \mathbb{C}^{N\times N_i}$ the channel matrix from UT i to the BS, $\mathbf{x}_i \sim \mathcal{CN}(0, \mathbf{I}_{n_i})$ the transmit symbol of UT i, $\mathbf{W}_i \in \mathbb{C}^{N_i\times n_i}$ the isometric precoding vectors composed of n_i orthogonal streams and $0 < \alpha < 1$ an inter-cell interference factor. The vector $\mathbf{z} \in \mathbb{C}^N$ combines the inter-cell interference and the thermal noise. The covariance matrix $\mathbf{Z} \in \mathbb{C}^{N\times N}$ of $\mathbf{z}$ is given as

$$\mathbf{Z} = \mathbb{E}\mathbf{z}\mathbf{z}^{\mathsf{H}} = \alpha\left[\mathbf{H}_1\mathbf{W}_1\mathbf{P}_1\mathbf{W}_1^{\mathsf{H}}\mathbf{H}_1^{\mathsf{H}} + \mathbf{H}_3\mathbf{W}_3\mathbf{P}_3\mathbf{W}_3^{\mathsf{H}}\mathbf{H}_i^{\mathsf{H}}\right] + \sigma^2\mathbf{I}_N.$$

We assume an SDMA system with channel matrices $\mathbf{H}_k \in \mathbb{C}^{N\times N_k}$ generated as realizations of a random standard Gaussian matrix with entries of zero mean and variance $1/N_k$. For simplicity, we further assume that each UT uses $n_k = n$ different transmit signatures to which it assigns equal unit power, i.e., $\mathbf{P}_k = \mathbf{I}_n$. Under these assumptions,

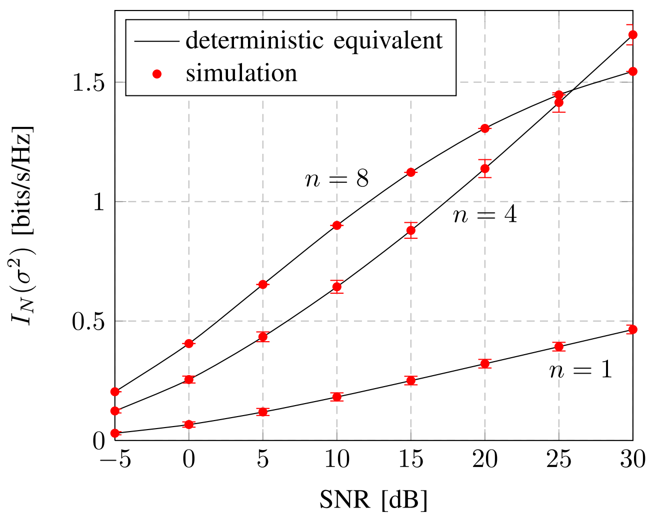

Fig. 2. Mutual information $I_N(\sigma^2)$ versus SNR for different numbers of transmit signatures n, $N = 16$, $N_i = 8$, $\mathbf{P}_i = \mathbf{I}_n$, $\alpha = 0.5$. Error bars represent one standard deviation on each side.

the mutual information $I_N(\sigma^2)$ of the central cell when the interference is treated as noise is given by

$$\begin{aligned} I_N(\sigma^2) &= \frac{1}{N}\log\det\left(\mathbf{I}_N + \mathbf{Z}^{-\frac{1}{2}}\mathbf{H}_2\mathbf{W}_2\mathbf{W}_2^{\mathsf{H}}\mathbf{H}_2\mathbf{Z}^{-\frac{1}{2}}\right) \\ &= \frac{1}{N}\log\det\left(\mathbf{I}_N + \frac{1}{\sigma^2}\sum_{k=1}^{3}\mathbf{H}_k\mathbf{W}_k\mathbf{W}_k^{\mathsf{H}}\mathbf{H}_k\right) - \frac{1}{N}\log\det\left(\mathbf{I}_N + \frac{1}{\sigma^2}\sum_{\substack{k=1\\k\neq 2}}^{3}\mathbf{H}_k\mathbf{W}_k\mathbf{W}_k^{\mathsf{H}}\mathbf{H}_k\right). \end{aligned}$$

According to [30], the spectral norm of $\mathbf{H}_k\mathbf{H}_k^{\mathsf{H}}$ is almost surely uniformly bounded. For such channel realizations, we are therefore in the conditions of Theorem 3. As a consequence, $I_N(\sigma^2) - \bar{I}_N(\sigma^2) \xrightarrow{\text{a.s.}} 0$, with $\bar{I}_N$ defined in Theorem 3 (termed $\bar{I}_N^{(a)}$). An approximation of the SINR at the output of the MMSE receiver for the jth entry of $\mathbf{x}_2$ can also be computed directly by Theorem 5. We assume $\alpha = 0.25$, $N = 16$, $N_1 = N_2 = N_3 = 8$ and define $\text{SNR} = 1/\sigma^2$. We consider a single random realization of the matrices $\mathbf{H}_k$, which is assumed to be static and therefore deterministically known.

Figure 2 depicts $I_N(\sigma^2)$ and the deterministic equivalent $\bar{I}_N(\sigma^2)$ versus SNR for different values of $n \in \{1, 4, 8\}$, scaled to bits/s/Hz instead of nats. We observe a very accurate fit between both results over the full range of SNR and n. This validates the deterministic approximation of the mutual information for systems of even small dimensions.

In Figure 3, we compare the per-receive antenna sum rate $R_N(\sigma^2)$ with single-stream MMSE-detection to the associated deterministic equivalent $\bar{R}_N(\sigma^2)$, for the same system conditions as in Figure 2. The sum rate $R_N(\sigma^2)$ is explicitly given by

$$R_N(\sigma^2) = \frac{1}{N}\sum_{k=1}^{n}\log\left(1 + \gamma_{2,k}^N(\sigma^2)\right)$$

with $\gamma_{ij}^N(\sigma^2)$ defined in (10) (termed $\gamma_{ij}^{N(a)}(\sigma^2)$). As for $\bar{R}_N(\sigma^2)$, from Theorem 5, it reads

$$\bar{R}_N(\sigma^2) = c_2\bar{c}_2\log\left(1 + \frac{a_2(\sigma^2)}{\bar{c}_2 - a_2(\sigma^2)\bar{a}_2(\sigma^2)}\right)$$

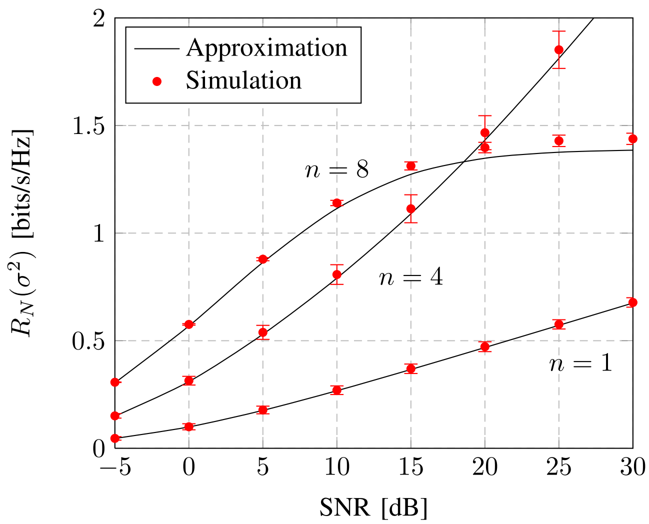

Fig. 3. Sum rate $R_N(\sigma^2)$ at the output of the MMSE decoder for user 2 versus SNR for different numbers of transmit signatures n, $N = 16$, $N_i = 8$, $\mathbf{P}_i = \mathbf{I}_n$, $\alpha = 0.5$. Error bars represent one standard deviation on each side.

with $a_2(\sigma^2)$ and $\bar{a}_2(\sigma^2)$ defined in Theorem 7. Similar to the previous observations, the deterministic equivalent provides an accurate approximation for all values of SNR and n, although the precision is slightly less than for the mutual information in Figure 2.

B. Multiple access channel

In this and the following example, we apply the theoretical results of Section II under the fading channel scenario (hypothesis (ii-b)). We consider a MAC from three transmitters to a single receiver as shown in Figure 4. The channel from each transmitter to the receiver is modeled by the Kronecker model (see Remark 1) with individual transmit and receive covariance matrices $\mathbf{T}_k$ and $\mathbf{R}_k$ and we assume additionally a different path loss $\alpha_k > 0$ on each link. The received signal vector $\mathbf{y}$ for this model reads

$$\mathbf{y} = \sum_{k=1}^{3} \sqrt{\alpha_k}\mathbf{R}_k^{\frac{1}{2}}\mathbf{Z}_k\mathbf{T}_k^{\frac{1}{2}}\mathbf{W}_k\mathbf{P}_k^{\frac{1}{2}}\mathbf{x}_k + \mathbf{n}$$

where $\mathbf{x}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{N_k})$ and $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I}_N)$. We create the correlation matrices according to a generalization of Jakes' model with non-isotropic signal transmission, see e.g. [31], [32], [33], where the elements of $\mathbf{T}_k$ and $\mathbf{R}_k$ are given as

$$[\mathbf{T}_k]_{ij} = \frac{1}{\theta_{\text{max}}^{t,k} - \theta_{\text{min}}^{t,k}} \int_{\theta_{\text{min}}^{t,k}}^{\theta_{\text{max}}^{t,k}} \exp\left(\frac{\mathbf{i}2\pi}{\lambda} d_{ij}^{t,k} \cos(\theta)\right) d\theta$$
$$[\mathbf{R}_k]_{ij} = \frac{1}{\theta_{\text{max}}^{r,k} - \theta_{\text{min}}^{r,k}} \int_{\theta_{\text{min}}^{r,k}}^{\theta_{\text{max}}^{r,k}} \exp\left(\frac{\mathbf{i}2\pi}{\lambda} d_{ij}^{r} \cos(\theta)\right) d\theta \qquad (11)$$

where $(\theta_{\text{min}}^{t,k}, \theta_{\text{max}}^{t,k})$ and $(\theta_{\text{min}}^{r,k}, \theta_{\text{max}}^{r,k})$ determine the azimuth angles over which useful signal power for the kth transmitter is radiated or received, $d_{ij}^{t,k}$ and d_{ij}^{r} are the distances between the antenna elements i and j at the

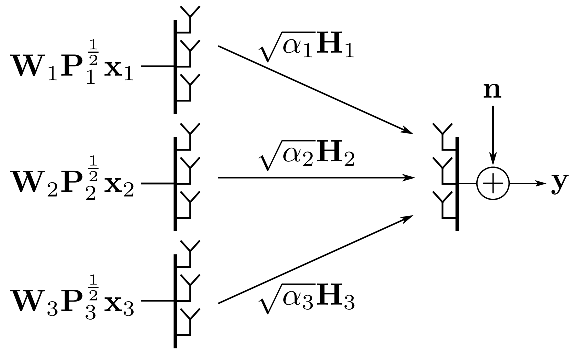

Fig. 4. MIMO MAC from three transmitters ($k = 1, 2, 3$) with N_k antennas to a receiver with N antennas. Each transmitter sends n_k streams with precoding matrix $\mathbf{W}_k$ and power allocation $\mathbf{P}_k$ over the channel $\sqrt{\alpha_k}\mathbf{H}_k$.

TABLE I

SIMULATION PARAMETERS FOR FIGURE 5: $N = 10$, $d_{ij}^r = 8\lambda(i-j)$

k	N_k	n_k	$\theta_{\text{min}}^{t,k}$	$\theta_{\text{max}}^{t,k}$	$\theta_{\text{min}}^{r,k}$	$\theta_{\text{max}}^{r,k}$	$d_{ij}^{t,k}$	α_k
1	10	8	0	$\pi/2$	$-\pi/4$	0	$4\lambda(i-j)$	1
2	5	4	$-\pi/4$	$\pi/4$	0	$\pi/3$	$4\lambda(i-j)$	$1/2$
3	5	4	$-\pi/2$	0	$-\pi/3$	$\pi/3$	$4\lambda(i-j)$	$1/2$

kth transmitter and receiver, respectively, and λ is the signal wavelength. We assume uniform power allocation for all k, and define SNR $= 1/\sigma^2$. All other parameters are summarized in Table I.

Figure 5 compares the normalized mutual information $I_N(\sigma^2)$ and the normalized rate with MMSE decoding $R_N(\sigma^2)$, averaged over $10,000$ different realizations of the matrices $\mathbf{H}_k$ and $\mathbf{W}_k$, against their deterministic approximations $\bar{I}_N(\sigma^2)$ and $\bar{R}_N(\sigma^2)$. Although we have chosen small dimensions for all matrices (see Table I), the match between both results is almost perfect. Also the fluctuations of $I_N(\sigma^2)$ and $R_N(\sigma^2)$ are rather small as can be seen from the error bars representing one standard deviation in each direction. The figure further illustrates the gains of optimal power allocation with a sum-power constraint (II), where we have chosen $P = \sum_{k=1}^{3} \frac{1}{n_k}\text{tr}\mathbf{I}_{n_k} = 3$.

C. Stream-control in interference channels

Our last example considers a MIMO interference channel consisting of two transmitter-receiver pairs as depicted in Figure 6. The received signal vectors $\mathbf{y}_1, \mathbf{y}_2 \in \mathbb{C}^N$ are respectively given as

$$\mathbf{y}_1 = \mathbf{H}_{11}\mathbf{W}_1\mathbf{P}_1^{\frac{1}{2}}\mathbf{x}_1 + \mathbf{H}_{12}\mathbf{W}_2\mathbf{P}_2^{\frac{1}{2}}\mathbf{x}_2 + \mathbf{n}_1$$
$$\mathbf{y}_2 = \mathbf{H}_{21}\mathbf{W}_1\mathbf{P}_1^{\frac{1}{2}}\mathbf{x}_1 + \mathbf{H}_{22}\mathbf{W}_2\mathbf{P}_2^{\frac{1}{2}}\mathbf{x}_2 + \mathbf{n}_2$$

where $\mathbf{H}_{qk} \in \mathbb{C}^{N\times N_k}$, $\mathbf{W}_k \in \mathbb{C}^{N_k\times N_k}$, $\mathbf{x}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{N_k})$, $\mathbf{P}_k \in \mathbb{R}_+^{N_k\times N_k}$ satisfying $\frac{1}{N_k}\text{tr}\mathbf{P}_k = 1$, and $\mathbf{n}_k \sim \mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I}_N)$, for $q, k \in \{1, 2\}$. Assuming that the receivers are aware of both precoding matrices and their respective

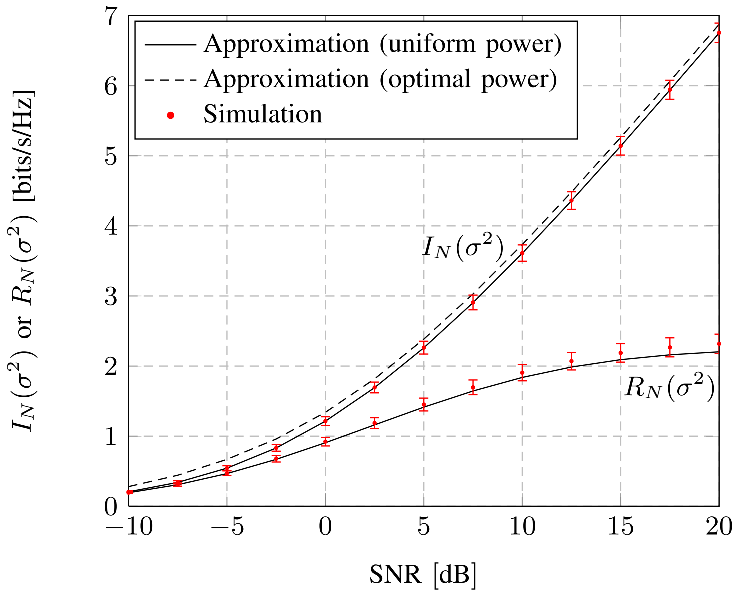

Fig. 5. Comparison of the average normalized mutual information $I_N(\sigma^2)$ and the normalized rate with MMSE decoding $R_N(\sigma^2)$ with their deterministic approximations $\bar{I}_N(\sigma^2)$ and $\bar{R}_N(\sigma^2)$. Error bars represent one standard deviation in each direction.

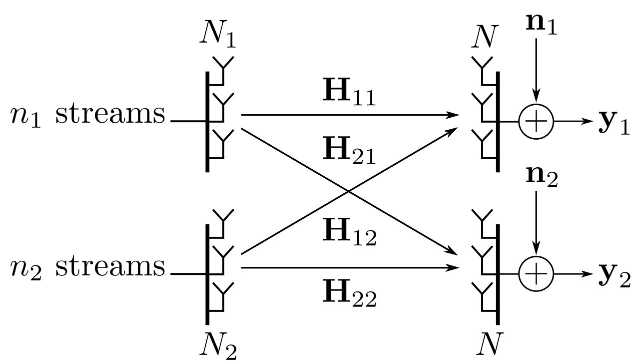

Fig. 6. Interference channel from two transmitters with N_k $(k = 1, 2)$ antennas, respectively, to two receivers with N antennas each. Each transmitter sends n_k independent data streams to its respective receiver.

channels but treat the interfering transmission as noise, the normalized mutual informations between $\mathbf{x}_1$ and $\mathbf{y}_1$, and $\mathbf{x}_2$ and $\mathbf{y}_2$, are respectively given as

$$I_1(\sigma^2) = \frac{1}{N}\log\det\left(\mathbf{I}_N + \frac{1}{\sigma^2}\sum_{k=1}^{2}\mathbf{H}_{1k}\mathbf{W}_k\mathbf{P}_k\mathbf{W}_k^{\mathsf{H}}\mathbf{H}_{1k}^{\mathsf{H}}\right) - \frac{1}{N}\log\det\left(\mathbf{I}_N + \frac{1}{\sigma^2}\mathbf{H}_{12}\mathbf{W}_2\mathbf{P}_2\mathbf{W}_2^{\mathsf{H}}\mathbf{H}_{12}^{\mathsf{H}}\right)$$

$$I_2(\sigma^2) = \frac{1}{N}\log\det\left(\mathbf{I}_N + \frac{1}{\sigma^2}\sum_{k=1}^{2}\mathbf{H}_{2k}\mathbf{W}_k\mathbf{P}_k\mathbf{W}_k^{\mathsf{H}}\mathbf{H}_{2k}^{\mathsf{H}}\right) - \frac{1}{N}\log\det\left(\mathbf{I}_N + \frac{1}{\sigma^2}\mathbf{H}_{21}\mathbf{W}_1\mathbf{P}_1\mathbf{W}_1^{\mathsf{H}}\mathbf{H}_{21}^{\mathsf{H}}\right).$$

We adopt the same channel model as in Section III-B, where the channel matrices $\mathbf{H}_{qk}$ are given as

$$\mathbf{H}_{qk} = \mathbf{R}_{qk}^{\frac{1}{2}}\mathbf{Z}_{qk}\mathbf{T}_k^{\frac{1}{2}}$$

where $\mathbf{Z}_{qk} \in \mathbb{C}^{N \times N_k}$ have independent $\mathcal{CN}(0, 1/N)$ entries and $\mathbf{T}_k$ and $\mathbf{R}_{qk}$ are calculated according to (11). We assume that no channel state information is available at the transmitters, so that the matrices $\mathbf{P}_k$ are simply used to determine the number of independently transmitted streams:

$$\mathbf{P}_k = \frac{N_k}{n_k}\text{diag}\left(\underbrace{1, \ldots, 1}_{n_k}, \underbrace{0, \ldots, 0}_{N_k - n_k}\right).$$

We will now apply the previously derived results to find the optimal number of streams $(n_1^\star, n_2^\star)$ maximizing the normalized ergodic sum-rate of the interference channel above. That is, we seek to find

$$\begin{aligned}(n_1^\star, n_2^\star) \ &= \ \max_{n_1, n_2} \mathbb{E}\left[I_1(\sigma^2) + I_2(\sigma^2)\right] \\ &\text{s.t. } 1 \leq n_1 \leq N_1, \ 1 \leq n_2 \leq N_2\end{aligned}$$

where the expectation is with respect to both channel and precoding matrices. Due to the complexity of the random matrix model, this optimization problem appears intractable by exact analysis. At the same time, any solution based on an exhaustive search in combination with Monte Carlo simulations becomes quickly prohibitive for large N_1, N_2, since $N_1 \times N_2$ possible combinations need to be tested. Relying on Theorem 4, we can calculate an approximation of $\mathbb{E}\left[I_1(\sigma^2) + I_2(\sigma^2)\right]$ to find an approximate solution which becomes asymptotically exact as N_1 and N_2 grow large. Thus, we determine $(\bar{n}_1^\star, \bar{n}_2^\star)$ as the solution to

$$\begin{aligned}(\bar{n}_1^\star, \bar{n}_2^\star) \ &= \ \max_{n_1, n_2} \bar{I}_1(\sigma^2) + \bar{I}_2(\sigma^2) \\ &\text{s.t. } 1 \leq n_1 \leq N_1, \ 1 \leq n_2 \leq N_2\end{aligned}$$

where $\bar{I}_1(\sigma^2), \bar{I}_2(\sigma^2)$ are calculated based on a direct application of Theorem 4 to each of the two log-det terms in $I_1(\sigma^2)$ and $I_2(\sigma^2)$, respectively. The optimal values $(\bar{n}_1^\star, \bar{n}_2^\star)$ are then found by an exhaustive search over all possible combinations. Although we still need to compute $N_1 \times N_2$ values, this is computationally much cheaper than Monte Carlo simulations.

Figure 7 and Figure 8 show the average normalized sum-rate $\mathbb{E}\left[I_1(\sigma^2) + I_2(\sigma^2)\right]$ and the deterministic approximation $\bar{I}_1(\sigma^2) + \bar{I}_2(\sigma^2)$, by Theorem 4, as a function of (n_1, n_2) for the simulation parameters as given in Table II. We have assumed SNR $= 0\,$dB and SNR $= 40\,$dB in Figure 7 and Figure 8, respectively. In both figures, the solid grid represents simulation results and the markers the deterministic approximations. We observe here again an almost perfect overlap between both sets of results for all values of (n_1, n_2). The optimal values $(n_1^\star, n_2^\star)$ and $(\bar{n}_1^\star, \bar{n}_2^\star)$ coincide for both values of SNR and are indicated by large crosses. At low SNR, both transmitters should send as many independent streams as transmit antennas, i.e., $n_1 = n_2 = 10$. At high SNR, one transmitter should use only a single stream ($n_2 = 1$) and the other transmitter $n_1 = N - 1 = 9$ streams. These results are in line with the observations of [19].

Obviously, the last result is highly unfair and better solutions can be achieved by using different objective functions, such as weighted sum-rate maximization. Also optimal stream-control with MMSE decoding could be carried out in a similar manner. Although we would still need to perform and exhaustive search over all possible

TABLE II

SIMULATION PARAMETERS FOR FIGURE 7 AND 8: $N = 10$, $d_{ij}^{r,k} = 4\lambda(i-j)$, $d_{ij}^{t,k} = 4\lambda(i-j)$

(q,k)	N_k	$\theta_{\min}^{t,k}$	$\theta_{\max}^{t,k}$	$\theta_{\min}^{r,q,k}$	$\theta_{\max}^{r,q,k}$
(1,1)	10	0	$\pi/2$	$-\pi/4$	0
(1,2)	10	$-\pi/2$	0	0	$\pi/4$
(2,1)	10	0	$\pi/2$	$-\pi/3$	0
(2,2)	10	$-\pi/2$	0	0	$\pi/3$

combinations of n_1, n_2, the computations based on deterministic equivalents are significantly faster than simulation-based approaches. The development of more intelligent algorithms to determine $(\bar{n}_1^\star, \bar{n}_2^\star)$ is outside the scope of this paper and left to future work. The extension to more than two transmitter-receiver pairs is straightforward.

IV. CONCLUSION

We have studied a class of wireless communication channels with random unitary signature or precoding matrices over quasi-static and fast fading channels and with multiple users or cells. We have provided deterministic approximations of the mutual information, the SINR at the output of the MMSE receiver and the associated sum-rate, which are asymptotically accurate as the system dimensions grow large. Simulations in the contexts of multi-cell SDMA, MIMO-MAC, and interference channels verify the accuracy of the approximations even for systems of small dimensions. This work also constitutes a novel contribution to the field of random matrix theory, which can be extended to more involved communication models featuring isometric precoders.

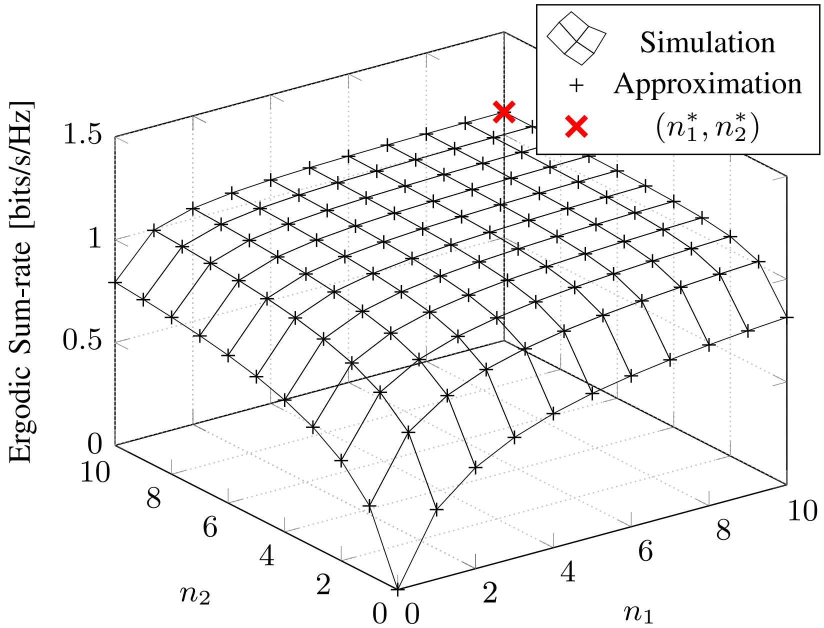

Fig. 7. Sum-rate versus number of transmitted data-streams (n_1, n_2) for SNR $= 0$ dB and all other parameters as provided in Table II. Solid lines correspond to simulation results, markers to the deterministic approximation by Theorem 4. As expected, both transmitters should send the maximum number of independent streams.

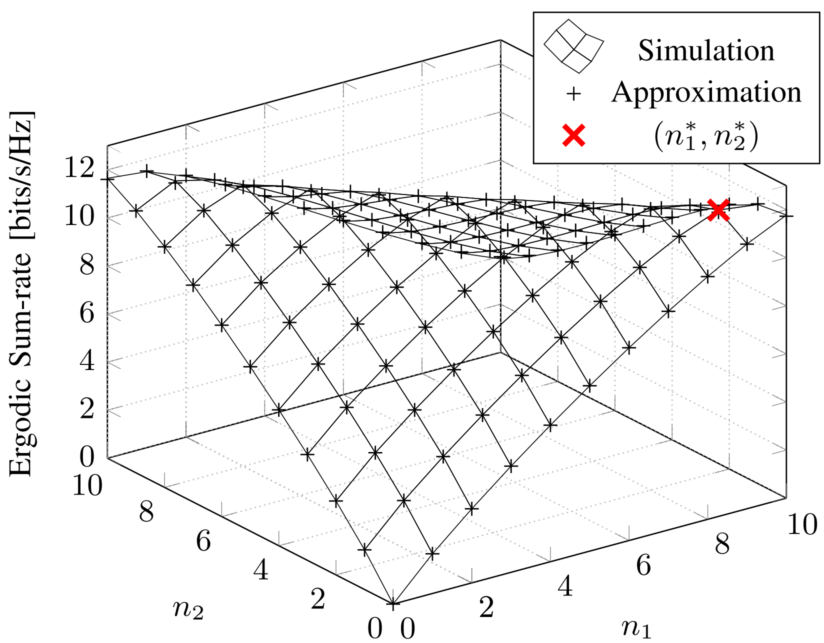

Fig. 8. Sum-rate versus number of transmitted data-streams (n_1, n_2) for SNR $= 40$ dB and all other parameters as provided in Table II. Solid lines correspond to simulation results, markers to the deterministic approximation by Theorem 4. As co-channel interference is dominant there is a clear gain of limiting the number of transmitted streams.

APPENDIX A
SPECTRAL APPROXIMATION OF $\mathbf{B}_N$ IN THE QUASI-STATIC MODEL

This section is dedicated to the proof of Theorem 7 introduced below. This theorem is the cornerstone result for all other results derived in this article. The proof is based on the Stieltjes transform method (documented extensively in [34], [11]).

We first remind some elementary notions needed in the following. For a Hermitian matrix $\mathbf{A} \in \mathbb{C}^{N\times N}$ with eigenvalues $\lambda_1 \leq \ldots \leq \lambda_N$, we denote by $F^{\mathbf{A}}$ the *empirical spectral distribution* (e.s.d.), defined as

$$F^{\mathbf{A}}(t) = \frac{1}{N}\sum_{i=1}^{N} \mathbf{1}_{\{\lambda_i \leq t\}}(t).$$

We now recall the definition of a Stieltjes transform.

Definition 1: Let F be the distribution function of a probability measure with support S. Then the Stieltjes transform of F, denoted m_F, is the function

$$\begin{aligned} m_F : \mathbb{C} \setminus S &\to \mathbb{C} \\ z &\mapsto \int \frac{1}{t-z} dF(t). \end{aligned}$$

In particular, for $F^{\mathbf{A}}$ the e.s.d. of a Hermitian matrix $\mathbf{A}$,

$$m_{F^{\mathbf{A}}}(z) = \frac{1}{N}\operatorname{tr}\left(\mathbf{A} - z\mathbf{I}_N\right)^{-1}$$

which will often be denoted $m_{\mathbf{A}}$.

In the course of the derivations, some defining properties of the Stieltjes transform will be needed. These are provided in Lemma 1 (Appendix F).

In the remainder of this section, we will prove the following result:

Theorem 7 (Convergence in distribution): For $i \in \{1,\ldots,K\}$, let $\mathbf{P}_i \in \mathbb{C}^{n_i\times n_i}$ be a Hermitian nonnegative matrix with spectral norm bounded uniformly along n_i and $\mathbf{W}_i \in \mathbb{C}^{N_i\times n_i}$ be $n_i \leq N_i$ columns of a unitary Haar distributed random matrix. Consider $\mathbf{H}_i \in \mathbb{C}^{N\times N_i}$ a random matrix such that $\mathbf{R}_i \triangleq \mathbf{H}_i\mathbf{H}_i^{\mathsf{H}} \in \mathbb{C}^{N\times N}$ has uniformly bounded spectral norm along N, almost surely. Define $c_i = \frac{n_i}{N_i}$ and $\bar{c}_i = \frac{N_i}{N}$ and denote

$$\mathbf{B}_N = \sum_{i=1}^{K} \mathbf{H}_i\mathbf{W}_i\mathbf{P}_i\mathbf{W}_i^{\mathsf{H}}\mathbf{H}_i^{\mathsf{H}}$$

and F_N the e.s.d. of $\mathbf{B}_N$.

Then, as $N\to\infty$, with $\bar{c}_i$ satisfying $0 < \liminf \bar{c}_i \leq \limsup \bar{c}_i < \infty$ and $0 \leq c_i \leq 1$ for all i, the following limit holds true almost surely

$$F_N - \bar{F}_N \Rightarrow 0$$

where $\bar{F}_N$ is the distribution function with Stieltjes transform $\bar{m}_N(z)$ defined by

$$\bar{m}_N(z) = \frac{1}{N}\operatorname{tr}\left(\sum_{i=1}^{K} \bar{e}_i(z)\mathbf{R}_i - z\mathbf{I}_N\right)^{-1} \tag{12}$$

where $(z \mapsto \bar{e}_1(z), \ldots, z \mapsto \bar{e}_K(z)) \in \mathcal{H}(\mathbb{C}\backslash\mathbb{R}^+, \mathbb{C})^K$ are the unique solutions of the following system of equations

$$\bar{e}_i(z) = \frac{1}{N} \operatorname{tr} \mathbf{P}_i \left(e_i(z)\mathbf{P}_i + [\bar{c}_i - e_i(z)\bar{e}_i(z)]\mathbf{I}_{n_i}\right)^{-1}$$

$$e_i(z) = \frac{1}{N} \operatorname{tr} \mathbf{R}_i \left(\sum_{j=1}^{K} \bar{e}_j(z)\mathbf{R}_j - z\mathbf{I}_N\right)^{-1} \tag{13}$$

which verify the conditions (i) $(z \mapsto e_1(z), \ldots, z \mapsto e_K(z)) \in \mathcal{S}(\mathbb{R}^+)^K$ and (ii) for z real negative and for all i, $0 \leq e_i(z) < c_i\bar{c}_i/\bar{e}_i(z)$.

Moreover, we have that for each real negative z,

$$\bar{e}_i(z) = \lim_{t\to\infty} \bar{e}_i^{(t)}(z), \quad \bar{e}_i^{(t)}(z) = \lim_{k\to\infty} \bar{e}_i^{(t,k)}(z)$$

and, for $k \geq 1$,

$$e_i^{(t)}(z) = \frac{1}{N} \operatorname{tr} \mathbf{R}_i \left(\sum_{j=1}^{K} \bar{e}_j^{(t-1)}(z)\mathbf{R}_j - z\mathbf{I}_N\right)^{-1}$$

$$\bar{e}_i^{(t,k)}(z) = \frac{1}{N} \operatorname{tr} \mathbf{P}_i \left(e_i^{(t)}(z)\mathbf{P}_i + [\bar{c}_i - e_i^{(t)}(z)\bar{e}_i^{(t,k-1)}(z)]\mathbf{I}_{n_i}\right)^{-1}$$

with the initial values $\bar{e}_i^{(t,0)}(z) \in [0, c_i\bar{c}_i/e_i^{(t)}(z))$ and $e_i^{(0)}(z) = 1$ for all i.

Remark 7: Denoting $a_i(\sigma^2) = e_i(-\sigma^2)$ for $\sigma^2 > 0$, we see immediately that Theorem 7 encompasses Theorem 1 as a special case.

We first provide an outline of the proof for better understanding. The full proof will be given in Appendix A-B.

A. Sketch of the proof

As a first step, we wish to prove that there exists a matrix $\mathbf{F}$ of the form $\mathbf{F} = \sum_{i=1}^{K} \bar{f}_i\mathbf{R}_i$, with $\bar{f}_i \in \mathbb{C}$, such that, for all nonnegative $\mathbf{A}$ with $\|\mathbf{A}\| < \infty$ uniformly on N,

$$\frac{1}{N} \operatorname{tr} \mathbf{A} \left(\mathbf{B}_N - z\mathbf{I}_N\right)^{-1} - \frac{1}{N} \operatorname{tr} \mathbf{A} \left(\mathbf{F} - z\mathbf{I}_N\right)^{-1} \xrightarrow{\text{a.s.}} 0.$$

Taking $\mathbf{A} = \mathbf{R}_i$ and denoting $f_i \triangleq \frac{1}{N} \operatorname{tr} \mathbf{R}_i \left(\mathbf{B}_N - z\mathbf{I}_N\right)^{-1}$, we will have in particular that

$$f_i - \frac{1}{N} \operatorname{tr} \mathbf{R}_i \left(\sum_{j=1}^{K} \bar{f}_j\mathbf{R}_j - z\mathbf{I}_N\right)^{-1} \xrightarrow{\text{a.s.}} 0.$$

Contrary to classical deterministic equivalent approaches for random matrices with i.i.d. entries, finding the approximation $\frac{1}{N} \operatorname{tr} \mathbf{A} \left(\mathbf{F} - z\mathbf{I}_N\right)^{-1}$ for $\frac{1}{N} \operatorname{tr} \mathbf{A} \left(\mathbf{B}_N - z\mathbf{I}_N\right)^{-1}$ is not straightforward. The reason is that, during the derivation, terms such as $\frac{1}{N_i - n_i} \operatorname{tr} \left(\mathbf{I}_{N_i} - \mathbf{W}_i\mathbf{W}_i^{\mathsf{H}}\right) \mathbf{H}_i^{\mathsf{H}} \left(\mathbf{B}_N - z\mathbf{I}_N\right)^{-1} \mathbf{H}_i$ with the $\left(\mathbf{I}_{N_i} - \mathbf{W}_i\mathbf{W}_i^{\mathsf{H}}\right)$ prefix will naturally appear and will be required to be controlled. We proceed as follows.

- We first denote for all i, $\delta_i \triangleq \frac{1}{N_i - n_i} \operatorname{tr} \left(\mathbf{I}_{N_i} - \mathbf{W}_i\mathbf{W}_i^{\mathsf{H}}\right) \mathbf{H}_i^{\mathsf{H}} \left(\mathbf{B}_N - z\mathbf{I}_N\right)^{-1} \mathbf{H}_i$ some auxiliary variable. Then we prove

$$f_i - \frac{1}{N} \operatorname{tr} \mathbf{R}_i \left(\mathbf{G} - z\mathbf{I}_N\right)^{-1} \xrightarrow{\text{a.s.}} 0,$$

with $\mathbf{G} = \sum_{j=1}^{K} \bar{g}_j \mathbf{R}_j$ and

$$\bar{g}_i = \frac{1}{(1-c_i)\bar{c}_i + \frac{1}{N}\sum_{l=1}^{n_i}\frac{1}{1+p_{il}\delta_i}} \frac{1}{N}\sum_{l=1}^{n_i}\frac{p_{il}}{1+p_{il}\delta_i},$$

where p_{il} denotes the lth eigenvalue of $\mathbf{P}_i$, and δ_i is linked to f_i through

$$f_i - \left((1-c_i)\bar{c}_i\delta_i + \frac{1}{N}\sum_{l=1}^{n_i}\frac{\delta_i}{1+p_{il}\delta_i}\right) \xrightarrow{\text{a.s.}} 0.$$

- This expression of $\bar{g}_i$, which is not convenient under this form, is then shown to satisfy

$$\bar{g}_i - \frac{1}{N}\sum_{l=1}^{n_i}\frac{p_{il}}{\bar{c}_i + p_{il}f_i - f_i\bar{g}_i} = \bar{g}_i - \frac{1}{N}\operatorname{tr}\mathbf{P}_i\left(f_i\mathbf{P}_i + [\bar{c}_i - f_i\bar{g}_i]\mathbf{I}_{n_i}\right)^{-1} \xrightarrow{\text{a.s.}} 0,$$

  which induces the $2K$-equation system

$$f_i - \frac{1}{N}\operatorname{tr}\mathbf{R}_i\left(\sum_{j=1}^{K}\bar{g}_j\mathbf{R}_j - z\mathbf{I}_N\right)^{-1} \xrightarrow{\text{a.s.}} 0$$

$$\bar{g}_i - \frac{1}{N}\operatorname{tr}\mathbf{P}_i\left(\bar{g}_i\mathbf{P}_i + [\bar{c}_i - f_i\bar{g}_i]\mathbf{I}_{n_i}\right)^{-1} \xrightarrow{\text{a.s.}} 0.$$

- These relations are sufficient to infer the deterministic equivalent, but will be made more attractive for further considerations by introducing $\mathbf{F} = \sum_{i=1}^{K}\bar{f}_i\mathbf{R}_i$, and proving that

$$f_i - \frac{1}{N}\operatorname{tr}\mathbf{R}_i\left(\sum_{j=1}^{K}\bar{f}_j\mathbf{R}_j - z\mathbf{I}_N\right)^{-1} \xrightarrow{\text{a.s.}} 0$$

$$\bar{f}_i - \frac{1}{N}\operatorname{tr}\mathbf{P}_i\left(\bar{f}_i\mathbf{P}_i + [\bar{c}_i - f_i\bar{f}_i]\mathbf{I}_{n_i}\right)^{-1} = 0,$$

  where, for $z < 0$, $\bar{f}_i$ lies in $[0, c_i\bar{c}_i/f_i)$ and is now uniquely determined by f_i.

This is the very technical part of the proof. We then prove in a second step the existence and uniqueness of a solution to the fixed-point equation

$$e_i - \frac{1}{N}\operatorname{tr}\mathbf{R}_i\left(\sum_{j=1}^{K}\bar{e}_j\mathbf{R}_j - z\mathbf{I}_N\right)^{-1} = 0$$

$$\bar{e}_i - \frac{1}{N}\operatorname{tr}\mathbf{P}_i\left(\bar{e}_i\mathbf{P}_i + [\bar{c}_i - e_i\bar{e}_i]\mathbf{I}_{n_i}\right)^{-1} = 0,$$

for all finite N, $z < 0$ and for $\bar{e}_i \in [0, c_i\bar{c}_i/f_i)$. This unfolds from a property of so-called *standard functions*. We will show precisely that the vector application $\mathbf{h} = (h_1, \ldots, h_K)$ defined for $z < 0$ by

$$h_i : (x_1, \ldots, x_K) \mapsto \frac{1}{N}\operatorname{tr}\mathbf{R}_i\left(\sum_{j=1}^{K}\bar{x}_j\mathbf{R}_j - z\mathbf{I}_N\right)^{-1}$$

$\bar{x}_i$ being the unique solution to

$$\bar{x}_i = \frac{1}{N}\operatorname{tr}\mathbf{P}_i\left(\bar{x}_i\mathbf{P}_i + [\bar{c}_i - x_i\bar{x}_i]\mathbf{I}_{n_i}\right)^{-1}$$

lying in $[0, c_i\bar{c}_i/x_i)$, is a standard function. It will unfold, from [35, Theorem 2], that the fixed-point equation in $(e_1, \ldots, e_K)$ has a unique solution with positive entries and that this solution can be determined as the limiting iteration of a classical fixed point algorithm.

The last step proves that the unique solution $(e_1,\dots,e_N)$ is such that

$$e_i - f_i \xrightarrow{\text{a.s.}} 0,$$

which is solved by standard arguments. This will entail immediately by classical complex analysis arguments that $m_N(z) - \bar{m}_N(z) \xrightarrow{\text{a.s.}} 0$ for all $z \in \mathbb{C} \setminus \mathbb{R}^+$, form which the almost sure convergence $F_N - \bar{F}_N \Rightarrow 0$ unfolds.

B. Complete proof

We will prove Theorem 7 by assuming first that, as N grows, the ratios $c_i = \frac{n_i}{N_i}$ for $i = \{1,\dots,K\}$ satisfy

$$\limsup_N c_i < 1.$$

We also assume for the time being that for all i, $\|\mathbf{R}_i\|$ is uniformly bounded surely. The cases where $\limsup_N c_i = 1$ for a certain i, or where $\|\mathbf{R}_i\|$ is uniformly bounded only in the almost sure sense, will be treated subsequently.

1) Case $\max_i \limsup_N c_i < 1$*:*

Step 1: Convergence

In this section, we take $z < 0$, until further notice. Let us first introduce the following parameters. We will denote $P = \max_i\{\limsup \|\mathbf{P}_i\|\}$, $R = \max_i\{\limsup \|\mathbf{R}_i\|\}$, $c_+ = \max_i\{\limsup c_i\}$, $\bar{c}_- = \min_i\{\liminf \bar{c}_i\}$ and $\bar{c}_+ = \max_i\{\limsup \bar{c}_i\}$.

We start with classical deterministic equivalent techniques.

Let $\mathbf{A} \in \mathbb{C}^{N\times N}$ be a Hermitian nonnegative definite matrix with spectral norm uniformly bounded by A. Recall the definition $\mathbf{R}_i = \mathbf{H}_i\mathbf{H}_i^{\mathsf{H}}$. Taking $\mathbf{G} = \sum_{j=1}^K \bar{g}_j \mathbf{R}_j$, with $\bar{g}_1,\dots,\bar{g}_K$ scalars left undefined for the moment, we have

$$\begin{aligned}
&\frac{1}{N}\operatorname{tr}\mathbf{A}(\mathbf{B}_N - z\mathbf{I}_N)^{-1} - \frac{1}{N}\operatorname{tr}\mathbf{A}(\mathbf{G} - z\mathbf{I}_N)^{-1} \\
&\overset{(a)}{=} \frac{1}{N}\operatorname{tr}\left[\mathbf{A}(\mathbf{B}_N - z\mathbf{I}_N)^{-1}\sum_{i=1}^K \mathbf{H}_i\left(-\mathbf{W}_i\mathbf{P}_i\mathbf{W}_i^{\mathsf{H}} + \bar{g}_i\mathbf{I}_{N_i}\right)\mathbf{H}_i^{\mathsf{H}}(\mathbf{G} - z\mathbf{I}_N)^{-1}\right] \\
&\overset{(b)}{=} \sum_{i=1}^K \bar{g}_i \frac{1}{N}\operatorname{tr}\mathbf{A}(\mathbf{B}_N - z\mathbf{I}_N)^{-1}\mathbf{R}_i(\mathbf{G} - z\mathbf{I}_N)^{-1} - \frac{1}{N}\sum_{i=1}^K\sum_{l=1}^{n_i} p_{il}\mathbf{w}_{il}^{\mathsf{H}}\mathbf{H}_i^{\mathsf{H}}(\mathbf{G} - z\mathbf{I}_N)^{-1}\mathbf{A}(\mathbf{B}_N - z\mathbf{I}_N)^{-1}\mathbf{H}_i\mathbf{w}_{il} \\
&\overset{(c)}{=} \sum_{i=1}^K \bar{g}_i \frac{1}{N}\operatorname{tr}\mathbf{A}(\mathbf{B}_N - z\mathbf{I}_N)^{-1}\mathbf{R}_i(\mathbf{G} - z\mathbf{I}_N)^{-1} - \frac{1}{N}\sum_{i=1}^K\sum_{l=1}^{n_i} \frac{p_{il}\mathbf{w}_{il}^{\mathsf{H}}\mathbf{H}_i^{\mathsf{H}}(\mathbf{G} - z\mathbf{I}_N)^{-1}\mathbf{A}(\mathbf{B}_{(i,l)} - z\mathbf{I}_N)^{-1}\mathbf{H}_i\mathbf{w}_{il}}{1 + p_{il}\mathbf{w}_{il}^{\mathsf{H}}\mathbf{H}_i^{\mathsf{H}}(\mathbf{B}_{(i,l)} - z\mathbf{I}_N)^{-1}\mathbf{H}_i\mathbf{w}_{il}}
\end{aligned} \tag{14}$$

with $\mathbf{w}_{il} \in \mathbb{C}^{N_i}$ the lth column of $\mathbf{W}_i$, $p_{i1},\dots,p_{in_i}$ the eigenvalues of $\mathbf{P}_i$ and $\mathbf{B}_{(i,l)} = \mathbf{B}_N - p_{il}\mathbf{H}_i\mathbf{w}_{il}\mathbf{w}_{il}^{\mathsf{H}}\mathbf{H}_i^{\mathsf{H}}$. The equality (a) follows from Lemma 2, (b) follows from the decomposition $\mathbf{W}_i\mathbf{P}_i\mathbf{W}_i^{\mathsf{H}} = \sum_{l=1}^{n_i} p_{il}\mathbf{w}_{il}\mathbf{w}_{il}^{\mathsf{H}}$, while the equality (c) follows from Lemma 3.

The idea now is to infer the values of the $\bar{g}_i$ such that the differences in (14) go to zero almost surely as N grows large. We will therefore proceed by studying the quantities $\mathbf{w}_{il}^{\mathsf{H}}\mathbf{H}_i^{\mathsf{H}}(\mathbf{B}_{(i,l)} - z\mathbf{I}_N)^{-1}\mathbf{H}_i\mathbf{w}_{il}$ and $\mathbf{w}_{il}^{\mathsf{H}}\mathbf{H}_i^{\mathsf{H}}(\mathbf{G} - z\mathbf{I}_N)^{-1}\mathbf{A}(\mathbf{B}_{(i,l)} - z\mathbf{I}_N)^{-1}\mathbf{H}_i\mathbf{w}_{il}$ in the denominator and numerator of the second term in (14).

For every $i \in \{1, \ldots, K\}$, denote

$$\delta_i \triangleq \frac{1}{N_i - n_i} \operatorname{tr} \left(\mathbf{I}_{N_i} - \mathbf{W}_i \mathbf{W}_i^{\mathsf{H}} \right) \mathbf{H}_i^{\mathsf{H}} \left(\mathbf{B}_N - z \mathbf{I}_N \right)^{-1} \mathbf{H}_i \ . \tag{15}$$

Introducing the additional term $(\mathbf{G} - z\mathbf{I}_N)^{-1}\mathbf{A}$ in the argument of the trace in δ_i, we denote

$$\beta_i \triangleq \frac{1}{N_i - n_i} \operatorname{tr} \left(\mathbf{I}_{N_i} - \mathbf{W}_i \mathbf{W}_i^{\mathsf{H}} \right) \mathbf{H}_i \left(\mathbf{G} - z \mathbf{I}_N \right)^{-1} \mathbf{A} \left(\mathbf{B}_N - z \mathbf{I}_N \right)^{-1} \mathbf{H}_i \ .$$

Under these notations, according to Lemma 5, the quantity $\mathbf{w}_{il}^{\mathsf{H}} \mathbf{H}_i^{\mathsf{H}} (\mathbf{B}_{(i,l)} - z\mathbf{I}_N)^{-1} \mathbf{H}_i \mathbf{w}_{il}$ is asymptotically close to δ_i, and, if $\mathbf{G}$ is independent of $\mathbf{w}_{il}$, the quantity $\mathbf{w}_{il}^{\mathsf{H}} \mathbf{H}_i^{\mathsf{H}} (\mathbf{G} - z\mathbf{I}_N)^{-1} \mathbf{A} (\mathbf{B}_{(i,l)} - z\mathbf{I}_N)^{-1} \mathbf{H}_i \mathbf{w}_{il}$ is asymptotically close to β_i.

We also define

$$f_i \triangleq \frac{1}{N} \operatorname{tr} \mathbf{R}_i \left(\mathbf{B}_N - z \mathbf{I}_N \right)^{-1} \geq 0$$

for any $z < 0$. Remark first, from standard matrix inequalities and the fact that $\mathbf{w}^{\mathsf{H}} \mathbf{A} \mathbf{w} \leq \|\mathbf{A}\|$ for any Hermitian matrix $\mathbf{A}$ and any unitary vector $\mathbf{w}$, that we have the following bounds on δ_i, β_i and f_i,

$$\delta_i \leq \frac{R}{|z|}, \quad \beta_i \leq \frac{RA}{|z|^2}, \quad f_i \leq \frac{R}{|z|} \ .$$

From Lemma 3, we have that

$$\begin{aligned}
(1 - c_i)\bar{c}_i \delta_i &= f_i - \frac{1}{N} \sum_{l=1}^{n_i} \mathbf{w}_{il}^{\mathsf{H}} \mathbf{H}_i^{\mathsf{H}} \left(\mathbf{B}_N - z \mathbf{I}_N \right)^{-1} \mathbf{H}_i \mathbf{w}_{il} \\
&= f_i - \frac{1}{N} \sum_{l=1}^{n_i} \frac{\mathbf{w}_{il}^{\mathsf{H}} \mathbf{H}_i^{\mathsf{H}} \left(\mathbf{B}_{(i,l)} - z \mathbf{I}_N \right)^{-1} \mathbf{H}_i \mathbf{w}_{il}}{1 + p_{il} \mathbf{w}_{il}^{\mathsf{H}} \mathbf{H}_i^{\mathsf{H}} \left(\mathbf{B}_{(i,l)} - z \mathbf{I}_N \right)^{-1} \mathbf{H}_i \mathbf{w}_{il}} \ .
\end{aligned} \tag{16}$$

Since $z < 0$, $\delta_i \geq 0$, so $\frac{1}{1 + p_{il}\delta_i}$ is well defined. By adding the term $\frac{1}{N} \sum_{l=1}^{n_i} \frac{\delta_i}{1 + p_{il}\delta_i}$ on both sides, (16) can be re-written as

$$\begin{aligned}
&(1 - c_i)\bar{c}_i \delta_i - f_i + \frac{1}{N} \sum_{l=1}^{n_i} \frac{\delta_i}{1 + p_{il}\delta_i} \\
&= \frac{1}{N} \sum_{l=1}^{n_i} \left[\frac{\delta_i}{1 + p_{il}\delta_i} - \frac{\mathbf{w}_{il}^{\mathsf{H}} \mathbf{H}_i^{\mathsf{H}} \left(\mathbf{B}_{(i,l)} - z \mathbf{I}_N \right)^{-1} \mathbf{H}_i \mathbf{w}_{il}}{1 + p_{il} \mathbf{w}_{il}^{\mathsf{H}} \mathbf{H}_i^{\mathsf{H}} \left(\mathbf{B}_{(i,l)} - z \mathbf{I}_N \right)^{-1} \mathbf{H}_i \mathbf{w}_{il}} \right] \\
&= \frac{1}{N} \sum_{l=1}^{n_i} \left[\frac{\delta_i - \mathbf{w}_{il}^{\mathsf{H}} \mathbf{H}_i^{\mathsf{H}} \left(\mathbf{B}_{(i,l)} - z \mathbf{I}_N \right)^{-1} \mathbf{H}_i \mathbf{w}_{il}}{(1 + p_{il}\delta_i) \left(1 + p_{il} \mathbf{w}_{il}^{\mathsf{H}} \mathbf{H}_i^{\mathsf{H}} \left(\mathbf{B}_{(i,l)} - z \mathbf{I}_N \right)^{-1} \mathbf{H}_i \mathbf{w}_{il} \right)} \right] \ .
\end{aligned}$$

We now apply Lemma 5 and Lemma 7, which together with $\delta_i \leq R|z|^{-1}$ ensures that

$$\mathbb{E} \left[\left| (1 - c_i)\bar{c}_i \delta_i - f_i + \frac{1}{N} \sum_{l=1}^{n_i} \frac{\delta_i}{1 + p_{il}\delta_i} \right|^4 \right] \leq 8 \frac{C}{N^2} \tag{17}$$

for some constant $C > 0$. This determines the asymptotic behaviour of δ_i and, thus, the asymptotic behaviour of the quantity $\mathbf{w}_{il}^{\mathsf{H}} \mathbf{H}_i^{\mathsf{H}} (\mathbf{B}_{(i,l)} - z\mathbf{I}_N)^{-1} \mathbf{H}_i \mathbf{w}_{il}$ in the denominator of (14).

We now proceed similarly with β_i as with δ_i. Assuming first that $\mathbf{G}$ is independent of $\mathbf{w}_{il}$, we first obtain

$$\begin{aligned}\beta_i &= \frac{1}{N_i-n_i}\operatorname{tr}\mathbf{H}_i^{\mathsf{H}}\left(\mathbf{G}-z\mathbf{I}_N\right)^{-1}\mathbf{A}\left(\mathbf{B}_N-z\mathbf{I}_N\right)^{-1}\mathbf{H}_i \\ &\quad -\frac{1}{N_i-n_i}\sum_{l=1}^{n_i}\frac{\mathbf{w}_{il}^{\mathsf{H}}\mathbf{H}_i^{\mathsf{H}}\left(\mathbf{G}-z\mathbf{I}_N\right)^{-1}\mathbf{A}\left(\mathbf{B}_{(i,l)}-z\mathbf{I}_N\right)^{-1}\mathbf{H}_i\mathbf{w}_{il}}{1+p_{il}\mathbf{w}_{il}^{\mathsf{H}}\mathbf{H}_i^{\mathsf{H}}\left(\mathbf{B}_{(i,l)}-z\mathbf{I}_N\right)^{-1}\mathbf{H}_i\mathbf{w}_{il}}\end{aligned}$$

from which we have

$$\begin{aligned}&\frac{1}{N_i-n_i}\operatorname{tr}\mathbf{H}_i^{\mathsf{H}}\left(\mathbf{G}-z\mathbf{I}_N\right)^{-1}\mathbf{A}\left(\mathbf{B}_N-z\mathbf{I}_N\right)^{-1}\mathbf{H}_i-\frac{1}{N_i-n_i}\sum_{l=1}^{n_i}\frac{\beta_i}{1+p_{il}\delta_i}-\beta_i \\ &=\frac{1}{N_i-n_i}\sum_{l=1}^{n_i}\left[\frac{\mathbf{w}_{il}^{\mathsf{H}}\mathbf{H}_i^{\mathsf{H}}\left(\mathbf{G}-z\mathbf{I}_N\right)^{-1}\mathbf{A}\left(\mathbf{B}_{(i,l)}-z\mathbf{I}_N\right)^{-1}\mathbf{H}_i\mathbf{w}_{il}}{1+p_{il}\mathbf{w}_{il}^{\mathsf{H}}\mathbf{H}_i^{\mathsf{H}}\left(\mathbf{B}_{(i,l)}-z\mathbf{I}_N\right)^{-1}\mathbf{H}_i\mathbf{w}_{il}}-\frac{\beta_i}{1+p_{il}\delta_i}\right]\,. \qquad (18)\end{aligned}$$

With the same inequalities as above, and with

$$\mathbf{w}_{il}^{\mathsf{H}}\mathbf{H}_i^{\mathsf{H}}\left(\mathbf{G}-z\mathbf{I}_N\right)^{-1}\mathbf{A}\left(\mathbf{B}_{(i,l)}-z\mathbf{I}_N\right)^{-1}\mathbf{H}_i\mathbf{w}_{il}\leq\frac{RA}{|z|^2}$$

we have that

$$\begin{aligned}&\mathbb{E}\left[\left|\frac{\mathbf{w}_{il}^{\mathsf{H}}\mathbf{H}_i^{\mathsf{H}}\left(\mathbf{G}-z\mathbf{I}_N\right)^{-1}\mathbf{A}\left(\mathbf{B}_{(i,l)}-z\mathbf{I}_N\right)^{-1}\mathbf{H}_i\mathbf{w}_{il}}{1+p_{il}\mathbf{w}_{il}^{\mathsf{H}}\mathbf{H}_i^{\mathsf{H}}\left(\mathbf{B}_{(i,l)}-z\mathbf{I}_N\right)^{-1}\mathbf{H}_i\mathbf{w}_{il}}-\frac{\beta_i}{1+p_{il}\delta_i}\right|^4\right] \\ &=\mathbb{E}\left[\left|\frac{\mathbf{w}_{il}^{\mathsf{H}}\mathbf{H}_i^{\mathsf{H}}\left(\mathbf{G}-z\mathbf{I}_N\right)^{-1}\mathbf{A}\left(\mathbf{B}_{(i,l)}-z\mathbf{I}_N\right)^{-1}\mathbf{H}_i\mathbf{w}_{il}-\beta_i}{(1+p_{il}\delta_i)(1+p_{il}\mathbf{w}_{il}^{\mathsf{H}}\mathbf{H}_i^{\mathsf{H}}\left(\mathbf{B}_{(i,l)}-z\mathbf{I}_N\right)^{-1}\mathbf{H}_i\mathbf{w}_{il})}\right.\right. \\ &\qquad+\frac{p_{il}\delta_i\left[\mathbf{w}_{il}^{\mathsf{H}}\mathbf{H}_i^{\mathsf{H}}\left(\mathbf{G}-z\mathbf{I}_N\right)^{-1}\mathbf{A}\left(\mathbf{B}_{(i,l)}-z\mathbf{I}_N\right)^{-1}\mathbf{H}_i\mathbf{w}_{il}-\beta_i\right]}{(1+p_{il}\delta_i)(1+p_{il}\mathbf{w}_{il}^{\mathsf{H}}\mathbf{H}_i^{\mathsf{H}}\left(\mathbf{B}_{(i,l)}-z\mathbf{I}_N\right)^{-1}\mathbf{H}_i\mathbf{w}_{il})} \\ &\qquad\left.\left.+\frac{p_{il}\beta_i\left[\delta_i-\mathbf{w}_{il}^{\mathsf{H}}\mathbf{H}_i^{\mathsf{H}}\left(\mathbf{B}_{(i,l)}-z\mathbf{I}_N\right)^{-1}\mathbf{H}_i\mathbf{w}_{il}\right]}{(1+p_{il}\delta_i)(1+p_{il}\mathbf{w}_{il}^{\mathsf{H}}\mathbf{H}_i^{\mathsf{H}}\left(\mathbf{B}_{(i,l)}-z\mathbf{I}_N\right)^{-1}\mathbf{H}_i\mathbf{w}_{il})}\right|^4\right] \\ &\leq 8\frac{C'}{N^2}\left(1+\frac{P^4R^4}{|z|^4}\left(1+\frac{A^4}{|z|^4}\right)\right) \qquad (19)\end{aligned}$$

for some $C'>C$. Multiplying (18) by $\frac{N_i-n_i}{N}$, we obtain

$$\begin{aligned}&\mathbb{E}\left[\left|\frac{1}{N}\operatorname{tr}\mathbf{H}_i^{\mathsf{H}}\left(\mathbf{G}-z\mathbf{I}_N\right)^{-1}\mathbf{A}\left(\mathbf{B}_N-z\mathbf{I}_N\right)^{-1}\mathbf{H}_i-\beta_i\left((1-c_i)\bar{c}_i+\frac{1}{N}\sum_{l=1}^{n_i}\frac{1}{1+p_{il}\delta_i}\right)\right|^4\right] \\ &\leq 8\frac{C'}{N^2}\left(1+\frac{P^4R^4}{|z|^4}\left(1+\frac{A^4}{|z|^4}\right)\right). \qquad (20)\end{aligned}$$

This now provides us with the asymptotic behaviour of β_i or equivalently of the quantity $\mathbf{w}_{il}^{\mathsf{H}}\mathbf{H}_i^{\mathsf{H}}(\mathbf{G}-z\mathbf{I}_N)^{-1}\mathbf{A}(\mathbf{B}_{(i,l)}-z\mathbf{I}_N)^{-1}\mathbf{H}_i\mathbf{w}_{il}$ in the numerator of (14).

We are now in position to infer the $\bar{g}_i$ such that $\frac{1}{N}\operatorname{tr}\mathbf{A}(\mathbf{B}_N-z\mathbf{I}_N)^{-1}-\frac{1}{N}\operatorname{tr}\mathbf{A}(\mathbf{G}-z\mathbf{I}_N)^{-1}$ is asymptotically small. For the previous derivations to hold, the scalars $\bar{g}_k$, $k\in\{1,\ldots,K\}$, were assumed independent of $\mathbf{w}_{il}$. It is however easy to see that these derivations still hold true (up to the choice of larger constants C, C') if $\bar{g}_k=\bar{g}_k^{(il)}+\varepsilon_N^{(ij)}$ with $\bar{g}_k^{(il)}$ independent of $\mathbf{w}_{il}$ and $\varepsilon_N^{(ij)}=\mathcal{O}(1/N)$.

We choose

$$\bar{g}_k = \frac{1}{(1-c_k)\bar{c}_k + \frac{1}{N}\sum_{m=1}^{n_k}\frac{1}{1+p_{km}\delta_k}}\frac{1}{N}\sum_{m=1}^{n_k}\frac{p_{km}}{1+p_{km}\delta_k} \tag{21}$$

and remark that $\bar{g}_k - \bar{g}_k^{(il)} = \mathcal{O}(1/N)$ with $\bar{g}_k^{(il)}$ defined similarly to $\bar{g}_k$ (15), with column $\mathbf{w}_{il}$ removed from the expression of δ_k.

Summing the previous results over i and l, we then have

$$\begin{aligned}
&\frac{1}{N}\operatorname{tr}\mathbf{A}(\mathbf{B}_N - z\mathbf{I}_N)^{-1} - \frac{1}{N}\operatorname{tr}\mathbf{A}(\mathbf{G}-z\mathbf{I}_N)^{-1}\\
&= \sum_{i=1}^K \frac{1}{(1-c_i)\bar{c}_i + \frac{1}{N}\sum_{l=1}^{n_i}\frac{1}{1+p_{il}\delta_i}}\frac{1}{N}\sum_{l=1}^{n_i}\frac{p_{il}}{1+p_{il}\delta_i}\frac{1}{N}\operatorname{tr}\mathbf{H}_i^{\mathsf{H}}\left(\mathbf{G}-z\mathbf{I}_N\right)^{-1}\mathbf{A}\left(\mathbf{B}_N - z\mathbf{I}_N\right)^{-1}\mathbf{H}_i\\
&\quad - \frac{1}{N}\sum_{i=1}^K\sum_{l=1}^{n_i}\frac{p_{il}\mathbf{w}_{il}^{\mathsf{H}}\mathbf{H}_i^{\mathsf{H}}(\mathbf{G}-z\mathbf{I}_N)^{-1}\mathbf{A}(\mathbf{B}_{(i,l)}-z\mathbf{I}_N)^{-1}\mathbf{H}_i\mathbf{w}_{il}}{1+p_{il}\mathbf{w}_{il}^{\mathsf{H}}\mathbf{H}_i^{\mathsf{H}}(\mathbf{B}_{(i,l)}-z\mathbf{I}_N)^{-1}\mathbf{H}_i\mathbf{w}_{il}}\\
&= \sum_{i=1}^K\frac{1}{N}\sum_{l=1}^{n_i}p_{il}\left[\frac{\frac{1}{N}\operatorname{tr}\mathbf{H}_i^{\mathsf{H}}\left(\mathbf{G}-z\mathbf{I}_N\right)^{-1}\mathbf{A}\left(\mathbf{B}_N-z\mathbf{I}_N\right)^{-1}\mathbf{H}_i}{((1-c_i)\bar{c}_i + \frac{1}{N}\sum_{l'=1}^{n_i}\frac{1}{1+p_{il'}\delta_i})(1+p_{il}\delta_i)} - \frac{\mathbf{w}_{il}^{\mathsf{H}}\mathbf{H}_i^{\mathsf{H}}(\mathbf{G}-z\mathbf{I}_N)^{-1}\mathbf{A}(\mathbf{B}_{(i,l)}-z\mathbf{I}_N)^{-1}\mathbf{H}_i\mathbf{w}_{il}}{1+p_{il}\mathbf{w}_{il}^{\mathsf{H}}\mathbf{H}_i^{\mathsf{H}}(\mathbf{B}_{(i,l)}-z\mathbf{I}_N)^{-1}\mathbf{H}_i\mathbf{w}_{il}}\right].
\end{aligned}$$

Notice now that $1+p_{il}\delta_i \geq 1$ and

$$(1-c_i)\bar{c}_i < (1-c_i)\bar{c}_i + \frac{1}{N}\sum_{l=1}^{n_i}\frac{1}{1+p_{il}\delta_i} \leq \bar{c}_i$$

which ensure that we can divide the term in the expectation of the left-hand side of (20) by $1+p_{il}\delta_i$ and $(1-c_i)\bar{c}_i + \frac{1}{N}\sum_{l=1}^{n_i}\frac{1}{1+p_{il}\delta_i}$ without taking the risk of the denominator getting close to 0. This leads to

$$\mathbb{E}\left[\left|\frac{\beta_i}{1+p_{il}\delta_i} - \frac{\frac{1}{N}\operatorname{tr}\mathbf{H}_i^{\mathsf{H}}\left(\mathbf{G}-z\mathbf{I}_N\right)^{-1}\mathbf{A}\left(\mathbf{B}_N-z\mathbf{I}_N\right)^{-1}\mathbf{H}_i}{\left((1-c_i)\bar{c}_i + \frac{1}{N}\sum_{l=1}^{n_i}\frac{1}{1+p_{il}\delta_i}\right)(1+p_{il}\delta_i)}\right|^4\right] \leq 8\frac{C'}{N^2(1-c_i)^4\bar{c}_i^4}\left(1+\frac{P^4R^4}{|z|^4}\left(1+\frac{A^4}{|z|^4}\right)\right). \tag{22}$$

From (19) and (22), we therefore have that

$$\begin{aligned}
&\mathbb{E}\left[\left|\frac{\frac{1}{N}\operatorname{tr}\mathbf{H}_i^{\mathsf{H}}\left(\mathbf{G}-z\mathbf{I}_N\right)^{-1}\mathbf{A}\left(\mathbf{B}_N-z\mathbf{I}_N\right)^{-1}\mathbf{H}_i}{\left((1-c_i)\bar{c}_i + \frac{1}{N}\sum_{l=1}^{n_i}\frac{1}{1+p_{il}\delta_i}\right)(1+p_{il}\delta_i)} - \frac{\mathbf{w}_{il}^{\mathsf{H}}\mathbf{H}_i^{\mathsf{H}}\left(\mathbf{G}-z\mathbf{I}_N\right)^{-1}\mathbf{A}\left(\mathbf{B}_{(i,l)}-z\mathbf{I}_N\right)^{-1}\mathbf{H}_i\mathbf{w}_{il}}{1+p_{il}\mathbf{w}_{il}^{\mathsf{H}}\mathbf{H}_i^{\mathsf{H}}\left(\mathbf{B}_{(i,l)}-z\mathbf{I}_N\right)^{-1}\mathbf{H}_i\mathbf{w}_{il}}\right|^4\right]\\
&\leq 128\frac{C'}{N^2(1-c_i)^4\bar{c}_i^4}\left(1+\frac{P^4R^4}{|z|^4}\left(1+\frac{A^4}{|z|^4}\right)\right).
\end{aligned}$$

We finally obtain

$$\mathbb{E}\left[\left|\frac{1}{N}\operatorname{tr}\mathbf{A}(\mathbf{B}_N-z\mathbf{I}_N)^{-1} - \frac{1}{N}\operatorname{tr}\mathbf{A}(\mathbf{G}-z\mathbf{I}_N)^{-1}\right|^4\right] \leq 128K^4\frac{C'}{N^2(1-c_+)^4\bar{c}_-^4}\left(1+\frac{P^4R^4}{|z|^4}\left(1+\frac{A^4}{|z|^4}\right)\right). \tag{23}$$

This provides a first convergence result as a function of the parameters δ_i, from which a deterministic equivalent can be determined. Nonetheless, the expression of $\bar{g}_i$ is rather impractical as it stands and we need to go further.

Observe in particular that $\bar{g}_i$ can be written under the form

$$\bar{g}_i = \frac{1}{N}\sum_{l=1}^{n_i}\frac{p_{il}}{((1-c_i)\bar{c}_i + \frac{1}{N}\sum_{l'=1}^{n_i}\frac{1}{1+p_{il'}\delta_i}) + p_{il}\delta_i((1-c_i)\bar{c}_i + \frac{1}{N}\sum_{l'=1}^{n_i}\frac{1}{1+p_{il'}\delta_i})}\ .$$

We will study the denominator of the above expression and show that it can be synthesized into a much more attractive form.

From (17), we first have

$$\mathbb{E}\left[\left|f_i - \delta_i\left((1-c_i)\bar{c}_i + \frac{1}{N}\sum_{l=1}^{n_i}\frac{1}{1+p_{il}\delta_i}\right)\right|^4\right] \leq \frac{8C}{N^2}\ . \tag{24}$$

Multiplying (21) by $-\delta_i\left((1-c_i)\bar{c}_i + \frac{1}{N}\sum_{l=1}^{n_i}\frac{1}{1+p_{il}\delta_i}\right)$ and adding $\bar{c}_i$ to both sides yields

$$\bar{c}_i - \bar{g}_i\delta_i\left((1-c_i)\bar{c}_i + \frac{1}{N}\sum_{l=1}^{n_i}\frac{1}{1+p_{il}\delta_i}\right) = (1-c_i)\bar{c}_i + \frac{1}{N}\sum_{l=1}^{n_i}\frac{1}{1+p_{il}\delta_i}\ .$$

By definition, $\bar{g}_i \leq \frac{P}{(1-c_i)\bar{c}_i}$, and we therefore also have

$$\mathbb{E}\left[\left|(\bar{c}_i - f_i\bar{g}_i) - \left((1-c_i)\bar{c}_i + \frac{1}{N}\sum_{l=1}^{n_i}\frac{1}{1+p_{il}\delta_i}\right)\right|^4\right] \leq 8\frac{C}{N^2}\frac{P^4}{(1-c_+)^4\bar{c}_-^4}\ . \tag{25}$$

The equations (24) and (25) can now be used to approximate the denominator of $\bar{g}_i$ as follows

$$\begin{aligned}&\mathbb{E}\left[\left|\bar{g}_i - \frac{1}{N}\sum_{l=1}^{n_i}\frac{p_{il}}{\bar{c}_i - f_i\bar{g}_i + p_{il}f_i}\right|^4\right]\\ &= \mathbb{E}\left[\left|\frac{1}{N}\sum_{l=1}^{n_i}p_{il}\frac{p_{il}\left[f_i - \delta_i((1-c_i)\bar{c}_i + \frac{1}{N}\sum_{l'=1}^{n_i}\frac{1}{1+p_{il'}\delta_i})\right] + \left[\bar{c}_i - f_i\bar{g}_i - ((1-c_i)\bar{c}_i + \frac{1}{N}\sum_{l'=1}^{n_i}\frac{1}{1+p_{il'}\delta_i})\right]}{\left[(1+p_{il}\delta_i)((1-c_i)\bar{c}_i + \frac{1}{N}\sum_{l'=1}^{n_i}\frac{1}{1+p_{il'}\delta_i})\right]\left[\bar{c}_i - f_i\bar{g}_i + p_{il}f_i\right]}\right|^4\right].\end{aligned} \tag{26}$$

Before to provide a useful bound, we need to ensure here that the term $\bar{c}_i - f_i\bar{g}_i + p_{il}f_i$ is uniformly away from zero, for all random f_i and for all N. For this, we recall the bounds $0 \leq f_i \leq \frac{R}{|z|}$ and $0 \leq \bar{g}_i \leq \frac{P}{(1-c_i)\bar{c}_i}$.

Let us consider $0 < \varepsilon < 1$ and take from now on $z < -\frac{RP}{(1-c_+)\bar{c}_-(\bar{c}_- - \varepsilon)}$, so that $\bar{c}_i - f_i\bar{g}_i > \varepsilon$ for all i. From (24), (25) and (26), we have

$$\mathbb{E}\left[\left|\bar{g}_i - \frac{1}{N}\sum_{l=1}^{n_i}\frac{p_{il}}{\bar{c}_i - f_i\bar{g}_i + p_{il}f_i}\right|^4\right] \leq 64\frac{C}{N^2}\frac{P^8}{(1-c_i)^4\bar{c}_i^4\varepsilon^4}\left(1 + \frac{1}{(1-c_i)^4\bar{c}_i^4}\right)$$

which is of order $\mathcal{O}(1/N^2)$ since we assumed $\limsup_N c_i < 1$.

We are now ready to introduce the matrix $\mathbf{F}$. Consider

$$\mathbf{F} = \sum_{i=1}^{K}\bar{f}_i\mathbf{R}_i,$$

with $\bar{f}_i$ defined as the unique solution to the equation in x

$$x = \frac{1}{N}\sum_{l=1}^{n_i}\frac{p_{il}}{\bar{c}_i - f_i x + f_i p_{il}}$$

within the interval $0 \leq x < c_i\bar{c}_i/f_i$. To prove the uniqueness of the solution within this interval, note simply that

$$\begin{aligned}\frac{c_i\bar{c}_i}{f_i} &> \frac{1}{N}\sum_{l=1}^{n_i}\frac{p_{il}}{\bar{c}_i - f_i(c_i\bar{c}_i/f_i) + f_i p_{il}}\\ 0 &\leq \frac{1}{N}\sum_{l=1}^{n_i}\frac{p_{il}}{\bar{c}_i - f_i\cdot 0 + f_i p_{il}}\end{aligned}$$

and that the function $x \mapsto \frac{1}{N}\sum_{l=1}^{n_i}\frac{p_{il}}{\bar{c}_i - f_i x + f_i p_{il}}$ is convex for $x \in [0, c_i\bar{c}_i/f_i)$. Hence the uniqueness of the solution in $[0, c_i\bar{c}_i/f_i)$. We also show that this solution is an attractor of the fixed-point algorithm, when correctly initialized. Indeed, let $x_0, x_1, \ldots$ be defined by

$$x_{n+1} = \frac{1}{N}\sum_{l=1}^{n_i}\frac{p_{il}}{\bar{c}_i - f_i x_n + f_i p_{il}},$$

with $x_0 \in [0, c_i\bar{c}_i/f_i)$. Then, $x_n \in [0, c_i\bar{c}_i/f_i)$ implies $\bar{c}_i - f_i x_n + f_i p_{il} > (1-c_i)\bar{c}_i + f_i p_{il} \geq f_i p_{il}$ and therefore $f_i x_{n+1} \leq c_i\bar{c}_i$, so $x_0, x_1, \ldots$ is contained in $[0, c_i\bar{c}_i/f_i)$. Now observe that

$$x_{n+1} - x_n = \frac{1}{N}\sum_{l=1}^{n_i}\frac{p_{il}f_i(x_n - x_{n-1})}{(\bar{c}_i + p_{il}f_i - f_i x_n)(\bar{c}_i + p_{il}f_i - f_i x_{n-1})}$$

with all terms being nonnegative in the sum, so that the differences $x_{n+1} - x_n$ and $x_n - x_{n-1}$ have the same sign (we also have from the above remarks that $x_{n+1} - x_n \leq c_i(x_n - x_{n-1})$). The sequence $x_0, x_1, \ldots$ is therefore monotonic and bounded: it converges. Calling x_∞ this limit, we have that

$$x_\infty = \frac{1}{N}\sum_{l=1}^{n_i}\frac{p_{il}}{\bar{c}_i + p_{il}f_i - f_i x_\infty}$$

as required.

To be able to finally prove that $\frac{1}{N}\operatorname{tr}\mathbf{A}(\mathbf{B}_N - z\mathbf{I}_N)^{-1} - \frac{1}{N}\operatorname{tr}\mathbf{A}(\mathbf{F} - z\mathbf{I}_N)^{-1} \xrightarrow{\text{a.s.}} 0$, we want now to show that $\bar{g}_i - \bar{f}_i$ tends to zero at a sufficiently fast rate. For this, we write

$$\begin{aligned}
&\mathbb{E}\left[\left|\bar{g}_i - \bar{f}_i\right|^4\right] \\
&\leq 8\left(\mathbb{E}\left[\left|\bar{g}_i - \frac{1}{N}\sum_{l=1}^{n_i}\frac{p_{il}}{\bar{c}_i - f_i\bar{g}_i + p_{il}f_i}\right|^4\right] + \mathbb{E}\left[\left|\frac{1}{N}\sum_{l=1}^{n_i}\frac{p_{il}}{\bar{c}_i - f_i\bar{g}_i + p_{il}f_i} - \frac{1}{N}\sum_{l=1}^{n_i}\frac{p_{il}}{\bar{c}_i - f_i\bar{f}_i + p_{il}f_i}\right|^4\right]\right) \\
&= 8\left(\mathbb{E}\left[\left|\bar{g}_i - \frac{1}{N}\sum_{l=1}^{n_i}\frac{p_{il}}{\bar{c}_i - f_i\bar{g}_i + p_{il}f_i}\right|^4\right] + \mathbb{E}\left[\left|\bar{g}_i - \bar{f}_i\right|^4\left|\frac{1}{N}\sum_{l=1}^{n_i}\frac{p_{il}f_i}{(\bar{c}_i - f_i\bar{f}_i + p_{il}f_i)(\bar{c}_i - f_i\bar{g}_i + p_{il}f_i)}\right|^4\right]\right)
\end{aligned} \tag{27}$$

where we have simply written $\bar{g}_i - \bar{f}_i = (\bar{g}_i - \frac{1}{N}\sum_{l=1}^{n_i}\frac{p_{il}}{\bar{c}_i - f_i\bar{g}_i + p_{il}f_i}) + (\frac{1}{N}\sum_{l=1}^{n_i}\frac{p_{il}}{\bar{c}_i - f_i\bar{g}_i + p_{il}f_i} - \bar{f}_i)$ and used the triangular inequality on the fourth power of each term.

We only need to ensure now that the coefficient multiplying $\left|\bar{g}_i - \bar{f}_i\right|$ in the right-hand side term is uniformly smaller than 1. For this, observe that, as $z \to -\infty$, $|p_{il}f_i| \leq \frac{PR}{|z|} \to 0$ in the numerator. In the denominator, we already know that $\bar{c}_i - f_i\bar{f}_i + p_{il}f_i \geq (1-c_i)\bar{c}_i$ and we also have that $\bar{c}_i - f_i\bar{g}_i + p_{il}f_i \geq \bar{c}_i - \frac{RP}{(1-c_i)|z|}$, which is greater than some $\eta > 0$ for $|z|$ taken large.

Take $\eta > 0$ and smaller than 1, and choose z to be such that, for all i,

$$\left|\frac{1}{N}\sum_{l=1}^{n_i}\frac{p_{il}f_i}{(\bar{c}_i - f_i\bar{f}_i + p_{il}f_i)(\bar{c}_i - f_i\bar{g}_i + p_{il}f_i)}\right| \leq \frac{PR}{|z|(1-c_i)\bar{c}_i\eta} < \frac{1-\eta}{8}$$

That is, from now on, we take $z < \min\left(-\frac{8PR}{\eta(1-\eta)(1-c_+)\bar{c}_-}, -\frac{RP}{(1-c_+)\bar{c}_-(1-\varepsilon)}\right)$.

From the inequality (27), gathering the terms in $\mathbb{E}\left[\left|\bar{g}_i - \bar{f}_i\right|^4\right]$ on the left side, we finally have

$$\mathbb{E}\left[\left|\bar{g}_i - \bar{f}_i\right|^4\right] \leq \frac{512}{\eta^4}\frac{C}{N^2}\frac{P^8}{(1-c_i)^4\bar{c}_i^4\varepsilon^4}\left(1+\frac{1}{(1-c_i)^4\bar{c}_i^4}\right) . \tag{28}$$

We can now proceed to prove the deterministic equivalent relations:

$$\begin{aligned}
&\frac{1}{N}\operatorname{tr}\mathbf{A}\left(\mathbf{G}-z\mathbf{I}_N\right)^{-1} - \frac{1}{N}\operatorname{tr}\mathbf{A}\left(\mathbf{F}-z\mathbf{I}_N\right)^{-1} \\
&= \sum_{i=1}^{K}\frac{1}{N}\sum_{l=1}^{n_i} p_{il}\left[\frac{\frac{1}{N}\operatorname{tr}\mathbf{H}_i^{\mathsf{H}}\mathbf{A}\left(\mathbf{G}-z\mathbf{I}_N\right)^{-1}\left(\mathbf{F}-z\mathbf{I}_N\right)^{-1}\mathbf{H}_i}{((1-c_i)\bar{c}_i+\frac{1}{N}\sum_{l'=1}^{n_i}\frac{1}{1+p_{i,l'}\delta_i})(1+p_{il}\delta_i)} - \frac{\frac{1}{N}\operatorname{tr}\mathbf{H}_i^{\mathsf{H}}\mathbf{A}\left(\mathbf{G}-z\mathbf{I}_N\right)^{-1}\left(\mathbf{F}-z\mathbf{I}_N\right)^{-1}\mathbf{H}_i}{\bar{c}_i - f_i\bar{f}_i + p_{il}f_i}\right] \\
&= \sum_{i=1}^{K}\frac{1}{N}\sum_{l=1}^{n_i} p_{il}\left[\left(\frac{\frac{1}{N}\operatorname{tr}\mathbf{H}_i^{\mathsf{H}}\mathbf{A}\left(\mathbf{G}-z\mathbf{I}_N\right)^{-1}\left(\mathbf{F}-z\mathbf{I}_N\right)^{-1}\mathbf{H}_i}{((1-c_i)\bar{c}_i+\frac{1}{N}\sum_{l'=1}^{n_i}\frac{1}{1+p_{i,l'}\delta_i})(1+p_{il}\delta_i)} - \frac{\frac{1}{N}\operatorname{tr}\mathbf{H}_i^{\mathsf{H}}\mathbf{A}\left(\mathbf{G}-z\mathbf{I}_N\right)^{-1}\left(\mathbf{F}-z\mathbf{I}_N\right)^{-1}\mathbf{H}_i}{\bar{c}_i - f_i\bar{g}_i + p_{il}f_i}\right)\right. \\
&\qquad \left. + \left(\frac{\frac{1}{N}\operatorname{tr}\mathbf{H}_i^{\mathsf{H}}\mathbf{A}\left(\mathbf{G}-z\mathbf{I}_N\right)^{-1}\left(\mathbf{F}-z\mathbf{I}_N\right)^{-1}\mathbf{H}_i}{\bar{c}_i - f_i\bar{g}_i + p_{il}f_i} - \frac{\frac{1}{N}\operatorname{tr}\mathbf{H}_i^{\mathsf{H}}\left(\mathbf{G}-z\mathbf{I}_N\right)^{-1}\left(\mathbf{F}-z\mathbf{I}_N\right)^{-1}\mathbf{H}_i}{\bar{c}_i - f_i\bar{f}_i + p_{il}f_i}\right)\right] \\
&= \sum_{i=1}^{K}\frac{1}{N}\operatorname{tr}\mathbf{H}_i^{\mathsf{H}}\mathbf{A}\left(\mathbf{G}-z\mathbf{I}_N\right)^{-1}\left(\mathbf{F}-z\mathbf{I}_N\right)^{-1}\mathbf{H}_i\frac{1}{N}\sum_{l=1}^{n_i}p_{il}\left[\frac{f_i(\bar{g}_i-\bar{f}_i)}{(\bar{c}_i - f_i\bar{f}_i + p_{il}f_i)(\bar{c}_i - f_i\bar{g}_i + p_{il}f_i)}\right. \\
&\qquad \left. + \frac{\left((\bar{c}_i - f_i\bar{g}_i) - ((1-c_i)\bar{c}_i+\frac{1}{N}\sum_{l'=1}^{n_i}\frac{1}{1+p_{i,l'}\delta_i})\right) + p_{il}\left(f_i - \delta_i((1-c_i)\bar{c}_i+\frac{1}{N}\sum_{l'=1}^{n_i}\frac{1}{1+p_{i,l'}\delta_i})\right)}{((1-c_i)\bar{c}_i+\frac{1}{N}\sum_{l'=1}^{n_i}\frac{1}{1+p_{i,l'}\delta_i})(1+p_{il}\delta_i)(\bar{c}_i - f_i\bar{f}_i + p_{il}f_i)}\right]
\end{aligned}$$

Therefore, from (24), (25) and (28),

$$\mathbb{E}\left[\left|\frac{1}{N}\operatorname{tr}\mathbf{A}\left(\mathbf{G}-z\mathbf{I}_N\right)^{-1} - \frac{1}{N}\operatorname{tr}\mathbf{A}\left(\mathbf{F}-z\mathbf{I}_N\right)^{-1}\right|^4\right] \leq \frac{64R^4P^4A^4K}{|z|^8(1-c_+)^8\bar{c}_-^8}\frac{C}{N^2}\left(1+\frac{1}{(1-c_+)^4\bar{c}_-^4}\right)^4\left[1+\frac{64R^4P^4}{|z|^4\eta^4\varepsilon^4}\right]$$

which is of order $\mathcal{O}(1/N^2)$.

Together with (23), applying Markov inequality, (5.31) of [36], and the Borel Cantelli lemma, Theorem 4.3 of [36], we finally have

$$\frac{1}{N}\operatorname{tr}\mathbf{A}\left(\mathbf{B}_N-z\mathbf{I}_N\right)^{-1} - \frac{1}{N}\operatorname{tr}\mathbf{A}\left(\mathbf{F}-z\mathbf{I}_N\right)^{-1} \xrightarrow{\text{a.s.}} 0, \tag{29}$$

as N grows large for realizations of $\{\mathbf{W}_1,\ldots,\mathbf{W}_K\}$ taken from a set $A_z \subset \Omega$ of probability one. This therefore holds true for countably many z (smaller than the established bound) with a cluster point in $\mathbb{R}^-$, on a set $A \subset \Omega$ of probability one. From Vitali convergence theorem, the analyticity of the functions under study and the fact that $\frac{1}{N}\operatorname{tr}\mathbf{A}\left(\mathbf{B}_N-z\mathbf{I}_N\right)^{-1}$ and $\frac{1}{N}\operatorname{tr}\mathbf{A}\left(\mathbf{F}-z\mathbf{I}_N\right)^{-1}$ are uniformly bounded on all closed subsets of $z \in \mathbb{C}\setminus\mathbb{R}^+$, we have that (23) holds true for all $z \in \mathbb{C}\setminus\mathbb{R}^+$ and the convergence (23) is uniform on all closed subsets of $\mathbb{C}\setminus\mathbb{R}^+$.

Applying the result for $\mathbf{A} = \mathbf{R}_j$, this is in particular

$$f_j - \frac{1}{N}\operatorname{tr}\mathbf{R}_j\left(\sum_{i=1}^{K}\bar{f}_i\mathbf{R}_i - z\mathbf{I}_N\right)^{-1} \xrightarrow{\text{a.s.}} 0$$

where we remind that $\bar{f}_i$ is the unique solution to

$$x = \frac{1}{N}\sum_{i=1}^{n_i}\frac{p_{il}}{\bar{c}_i - f_i x + p_{il}f_i}$$

within the set $[0, c_i\bar{c}_i/f_i)$. For $\mathbf{A} = \mathbf{I}_N$, this implies

$$m_N(z) - \frac{1}{N}\operatorname{tr}\left(\sum_{i=1}^{K}\bar{f}_i\mathbf{R}_i - z\mathbf{I}_N\right)^{-1} \xrightarrow{\text{a.s.}} 0$$

which proves the convergence.

Step 2: Existence and Uniqueness

For existence, it suffices to consider the matrices $\mathbf{P}_{[p],i} \in \mathbb{C}^{n_i p}$ and $\mathbf{H}_{[p],i} \in \mathbb{C}^{Np\times N_i p}$ for all i defined as the Kronecker products $\mathbf{P}_{[p],i} \triangleq \mathbf{P}_i \otimes \mathbf{I}_p$, $\mathbf{H}_{[p],i} \triangleq \mathbf{H}_i \otimes \mathbf{I}_p$, such that $\mathbf{P}_{[p],i}$ and $\mathbf{R}_{[p],i} = \mathbf{H}_{[p],i}\mathbf{H}_{[p],i}^{\mathsf{H}}$ have respectively the distribution function $F^{\mathbf{P}_i}$ and $F^{\mathbf{R}_i}$ for all p. It is easy to see that e_i is unchanged by substituting the $\mathbf{P}_{[p],i}$ and $\mathbf{R}_{[p],i}$ to the $\mathbf{P}_i$ and $\mathbf{R}_i$, respectively. Denoting similarly $f_{[p],i}$ the f_i adapted to $\mathbf{P}_{[p],i}$ and $\mathbf{H}_{[p],i}$, from the convergence result of Step 1, we can choose $f_{[1],i}, f_{[2],i}, \ldots$ a sequence of the set of probability one where convergence is ensured as p grows large (N and the n_i are kept fixed). This sequence is uniformly bounded (by $R/|z|$) in $\mathbb{C} \setminus \mathbb{R}^+$, and therefore we can extract a converging subsequence $f_{[\phi(p)],i}$ out of it. The limit of this subsequence satisfies the fixed-point equation, which therefore proves existence. Call $e_i(z)$ this limit.

We wish to prove that e_i, seen as a function of z, is the *Stieltjes transform* of a distribution function. For this, we prove the defining properties of a Stieltjes transform, provided in Lemma 1. The fact that e_i is analytic on $\mathbb{C}^+$ comes as an immediate application of Vitali's convergence theorem [37], as e_i is the almost sure limit of a series of analytic functions, bounded on every compact of $\mathbb{C} \setminus \mathbb{R}^+$. It is clear that for $z \in \mathbb{C}^+$, $\operatorname{Im}[f_{[p],i}(z)] > 0$, $\operatorname{Im}[zf_{[p],i}(z)] > 0$ and $\lim_{y\to\infty} -\mathbf{i}yf_{[p],i}(\mathbf{i}y) \leq R$. This implies that for $z \in \mathbb{C}^+$, $\operatorname{Im}[e_i(z)] \geq 0$, $z\operatorname{Im}[e_i(z)] \geq 0$ and $\lim_{y\to\infty} -\mathbf{i}ye_i(\mathbf{i}y) \leq R$. In addition, note that, for $z \in \mathbb{C}^+$,

$$\operatorname{Im}[f_{[p],i}] \geq \frac{1}{N}\frac{r}{(RP+|z|)^2}\operatorname{Im}[z] > 0$$

and

$$\operatorname{Im}[zf_{[p],i}] \geq \frac{1}{N}\frac{Kr^2t}{(RP+|z|)^2}\operatorname{Im}[z] > 0$$

with r a lower bound on the smallest non-zero eigenvalues of $\mathbf{R}_1, \ldots, \mathbf{R}_K$ (we naturally assume all $\mathbf{R}_k$ non-zero) and t a lower bound on the smallest non-zero eigenvalues of $\mathbf{T}_1, \ldots, \mathbf{T}_K$ (again, none assumed identically zero). Take $z \in \mathbb{C}^+$ and $\varepsilon < \frac{1}{2}\min(\frac{1}{N}\frac{r}{(RP+|z|)^2}\operatorname{Im}[z], \frac{1}{N}\frac{Kr^2t}{(RP+|z|)^2}\operatorname{Im}[z])$. There now exists p_0 such that $p \geq p_0$ implies $|\operatorname{Im}[f_{[\phi(p)],i}] - \operatorname{Im}[e_i]| < \varepsilon/2$ and $|\operatorname{Im}[zf_{[\phi(p)],i}] - \operatorname{Im}[ze_i]| < \varepsilon/2$, and therefore $\operatorname{Im}[e_i] > \varepsilon/2$ and $z\operatorname{Im}[e_i(z)] > \varepsilon/2$ so that $e_i(z)$ is the Stieltjes transform of a finite measure on $\mathbb{R}^+$.

We will prove uniqueness of positive solutions $e_1(z), \ldots, e_K(z)$ for $z < 0$ and the convergence of the classical fixed point algorithm to these values. We first introduce some notations and useful identities. Notice that, similar to the auxilliary variables δ_i in Step 1, we can define, for any pair of variables x_i and $\bar{x}_i$, with $\bar{x}_i$ defined as the solution y to $y = \frac{1}{N}\sum_{l=1}^{n_i}\frac{p_{il}}{\bar{c}_j - x_j y + x_j p_{il}}$ such that $0 \leq y < c_j\bar{c}_j/x_j$, the auxiliary variables $\Delta_1, \ldots, \Delta_K$, with the properties

$$x_i = \Delta_i \left((1-c_i)\bar{c}_i + \frac{1}{N}\sum_{l=1}^{n_i} \frac{1}{1+p_{il}\Delta_i} \right)$$
$$= \Delta_i \left(\bar{c}_i - \frac{1}{N}\sum_{l=1}^{n_i} \frac{p_{il}\Delta_i}{1+p_{il}\Delta_i} \right)$$

and

$$\bar{c}_i - x_i\bar{x}_i = (1-c_i)\bar{c}_i + \frac{1}{N}\sum_{l=1}^{n_i} \frac{1}{1+p_{il}\Delta_i}$$
$$= \bar{c}_i - \frac{1}{N}\sum_{l=1}^{n_i} \frac{p_{il}\Delta_i}{1+p_{il}\Delta_i} \ . \tag{30}$$

Indeed, firstly, there exists a unique mapping between x_i and Δ_i. This unfolds from noticing that

$$\frac{d\,x_i}{d\Delta_i} = \frac{d}{d\Delta_i}\left[\Delta_i \left((1-c_i)\bar{c}_i + \frac{1}{N}\sum_{l=1}^{n_i} \frac{1}{1+p_{il}\Delta_i} \right) \right] = (1-c_i)\bar{c}_i + \frac{1}{N}\sum_{l=1}^{n_i} \frac{1}{(1+p_{il}\Delta_i)^2} > 0$$

and therefore x_i and Δ_i are one-to-one. Additionally, x_i is a strictly growing function of Δ_i with $\Delta_i = 0$ for $x_i = 0$. This ensures that $\Delta_i > 0$ if and only if $x_i > 0$.

Secondly, from the definition of $\bar{x}_i$, we have

$$\bar{c}_i - x_i\bar{x}_i = \bar{c}_i - x_i \frac{1}{N}\sum_{l=1}^{n_i} \frac{p_{il}}{(\bar{c}_i - x_i\bar{x}_i) + p_{il}x_i}$$
$$= \bar{c}_i - \Delta_i \left(\bar{c}_i - \frac{1}{N}\sum_{l=1}^{n_i} \frac{p_{il}\Delta_i}{1+p_{il}\Delta_i} \right) \frac{1}{N}\sum_{l=1}^{n_i} \frac{p_{il}}{\bar{c}_i - x_i\bar{x}_i + p_{il}\Delta_i \left(\bar{c}_i - \frac{1}{N}\sum_{l=1'}^{n_i} \frac{p_{il'}\Delta_i}{1+p_{il'}\Delta_i} \right)} \ .$$

Note in particular that for $x_i\bar{x}_i = \frac{1}{N}\sum_{l=1'}^{n_i} \frac{p_{il'}\Delta_i}{1+p_{il'}\Delta_i}$, the above equation simplifies to

$$\bar{c}_i - \Delta_i \left(\bar{c}_i - \frac{1}{N}\sum_{l=1}^{n_i} \frac{p_{il}\Delta_i}{1+p_{il}\Delta_i} \right) \frac{1}{N}\sum_{l=1}^{n_1} \frac{p_{il}}{\left(\bar{c}_i - \frac{1}{N}\sum_{l=1'}^{n_i} \frac{p_{il'}\Delta_i}{1+p_{il'}\Delta_i} \right) + p_{il}\Delta_i \left(\bar{c}_i - \frac{1}{N}\sum_{l'=1}^{n_i} \frac{p_{il'}\Delta_i}{1+p_{il'}\Delta_i} \right)}$$
$$= \bar{c}_i - \frac{1}{N}\sum_{l=1}^{n_i} \frac{p_{il}\Delta_i}{1+p_{il}\Delta_i}$$

and therefore $\bar{c}_i - \frac{1}{N}\sum_{l=1}^{n_i} \frac{p_{il}\Delta_i}{1+p_{il}\Delta_i}$ is one of the solution of the implicit equation in u,

$$u = \bar{c}_i - x_i \frac{1}{N}\sum_{l=1}^{n_i} \frac{p_{il}}{u + p_{il}x_i} \ .$$

Equivalently, writing $u = \bar{c}_i - x_i y$, it follows that $\frac{1}{x_i}\frac{1}{N}\sum_{l=1}^{n_i} \frac{p_{il}\Delta_i}{1+p_{il}\Delta_i}$ is one of the solutions of the equation in y

$$y = \frac{1}{N}\sum_{l=1}^{n_i} \frac{p_{il}}{\bar{c}_i - x_i y + p_{il}x_i} \ .$$

Since

$$x_i \left(\frac{1}{x_i}\frac{1}{N}\sum_{l=1}^{n_i} \frac{p_{il}\Delta_i}{1+p_{il}\Delta_i} \right) < c_i\bar{c}_i$$

this solution lies in $[0, c_i\bar{c}_i/x_i)$ and is exactly equal to $\bar{x}_i$. This proves that the equations in $(x_i, \bar{x}_i)$ can be written under the form of the equations in $(\Delta_i, \bar{x}_i)$, as presented above.

We take the opportunity of the above definitions to notice that, for $x_i > x_i'$ and $\bar{x}_i'$, Δ_i' defined similarly as $\bar{x}_i$ and Δ_i,

$$x_i\bar{x}_i - x_i'\bar{x}_i' = \frac{1}{N}\sum_{l=1}^{n_i}\frac{p_{il}(\Delta_i - \Delta_i')}{(1+p_{il}\Delta_i)(1+p_{il}\Delta_i')} > 0 \tag{31}$$

whenever $\mathbf{P}_i \neq 0$. Therefore $x_i\bar{x}_i$ is a growing function of x_i (or equivalently of Δ_i). This will turn out to be a useful remark later.

We are now in position to prove the step of uniqueness. Define for $i \in \{1,\ldots,K\}$, the functions

$$h_i : (x_1,\ldots,x_K) \mapsto \frac{1}{N}\operatorname{tr}\mathbf{R}_i\left(\sum_{j=1}^{K}\bar{x}_j\mathbf{R}_j - z\mathbf{I}_N\right)^{-1}$$

with $\bar{x}_j$ the unique solution of the equation in y

$$y = \frac{1}{N}\sum_{l=1}^{n_j}\frac{p_{jl}}{\bar{c}_j + x_jp_{jl} - x_jy} \tag{32}$$

such that $0 \leq \bar{x}_j < c_j\bar{c}_j/x_j$.

We will prove in the following that the multivariate function $\mathbf{h} = (h_1,\ldots,h_K)$ is a *standard function* (or *interference standard function*), defined in [35], as follows:

Definition 2: A function $\mathbf{h}(x_1,\ldots,x_K) \in \mathbb{R}^K$ 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, $\alpha h_j(x_1,\ldots,x_K) > h_j(\alpha x_1,\ldots,\alpha x_K)$.

The important result regarding standard functions, [35, Theorem 2], is given as follows:

Theorem 8: 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

$$x_j = h_j(x_1,\ldots,x_K)$$

with $j \in \{1,\ldots,K\}$.

Since we have proved the existence of a solution of the fixed-point equation, there exists $(x_1,\ldots,x_K)$ such that for all j, $x_j = h_j(x_1,\ldots,x_K)$. Therefore, by showing that $\mathbf{h} \triangleq (h_1,\ldots,h_K)$ is standard, we will prove that the classical fixed point algorithm converges to the unique set of positive solutions $e_1,\ldots,e_K$, when $z < 0$.

The positivity condition is straightforward as $\bar{x}_i$ is positive for x_i positive and therefore $h_j(x_1,\ldots,x_K)$ is always positive whenever $x_1,\ldots,x_K$ are nonnegative.

The scalability is also rather direct. Let $\alpha > 1$, then

$$\alpha h_j(x_1,\dots,x_K) - h_j(\alpha x_1,\dots,\alpha x_K)$$
$$= \frac{1}{N}\operatorname{tr}\mathbf{R}_j\left(\sum_{k=1}^K \frac{\bar{x}_k}{\alpha}\mathbf{R}_k - \frac{z}{\alpha}\mathbf{I}_N\right)^{-1} - \frac{1}{N}\operatorname{tr}\mathbf{R}_j\left(\sum_{k=1}^K \bar{x}_k^{(\alpha)}\mathbf{R}_k - z\mathbf{I}_N\right)^{-1}$$

where we denoted $\bar{x}_j^{(\alpha)}$ the unique solution to (32) within $[0, c_j\bar{c}_j/(\alpha x_j))$ with x_j replaced by αx_j. From Lemma 6, it suffices to show that

$$\sum_{k=1}^K \left[\bar{x}_k^{(\alpha)} - \frac{\bar{x}_k}{\alpha}\right]\mathbf{R}_k + \left[z - \frac{z}{\alpha}\right]\mathbf{I}_N$$

is positive definite. Since $\alpha x_i > x_i$, we have from the property (31) that

$$\alpha x_k \bar{x}_k^{(\alpha)} - x_k\bar{x}_k > 0$$

or equivalently

$$\bar{x}_k^{(\alpha)} - \frac{\bar{x}_k}{\alpha} > 0.$$

Along with $1 - 1/\alpha > 0$ and $z < 0$, this ensures that $\alpha h_j(x_1,\dots,x_K) > h_j(\alpha x_1,\dots,\alpha x_K)$.

The monotonicity requires some more calculus. This unfolds from considering $\bar{x}_i$ as a function of Δ_i, by verifying that $\frac{d}{d\Delta_i}\bar{x}_i$ is negative.

$$\frac{d}{d\Delta_i}\bar{x}_i = \frac{1}{\Delta_i^2}\left(1 - \frac{\bar{c}_i}{\bar{c}_i - \frac{1}{N}\sum_{l=1}^{n_i}\frac{p_{il}\Delta_i}{1+p_{il}\Delta_i}}\right) + \frac{\bar{c}_i}{\Delta_i^2}\left(\frac{\frac{1}{N}\sum_{l=1}^{n_i}\frac{p_{il}\Delta_i}{(1+p_{il}\Delta_i)^2}}{\left(\bar{c}_i - \frac{1}{N}\sum_{l=1}^{n_i}\frac{p_{il}\Delta_i}{1+p_{il}\Delta_i}\right)^2}\right)$$
$$= \frac{1}{\Delta_i^2\left(\bar{c}_i - \frac{1}{N}\sum_{l=1}^{n_i}\frac{p_{il}\Delta_i}{1+p_{il}\Delta_i}\right)^2}\left[-\frac{1}{N}\left(\sum_{l=1}^{n_i}\frac{p_{il}\Delta_i}{1+p_{il}\Delta_i}\right)\left(\bar{c}_i - \frac{1}{N}\sum_{l=1}^{n_i}\frac{p_{il}\Delta_i}{1+p_{il}\Delta_i}\right) + \frac{\bar{c}_i}{N}\sum_{l=1}^{n_i}\frac{p_{il}\Delta_i}{(1+p_{il}\Delta_i)^2}\right]$$
$$= \frac{1}{\Delta_i^2\left(\bar{c}_i - \frac{1}{N}\sum_{l=1}^{n_i}\frac{p_{il}\Delta_i}{1+p_{il}\Delta_i}\right)^2}\left[\left(\frac{1}{N}\sum_{l=1}^{n_i}\frac{p_{il}\Delta_i}{1+p_{il}\Delta_i}\right)^2 - \frac{\bar{c}_i}{N}\sum_{l=1}^{n_i}\frac{p_{il}\Delta_i}{1+p_{il}\Delta_i} + \frac{\bar{c}_i}{N}\sum_{l=1}^{n_i}\frac{p_{il}\Delta_i}{(1+p_{il}\Delta_i)^2}\right]$$
$$= \frac{1}{\Delta_i^2\left(\bar{c}_i - \frac{1}{N}\sum_{l=1}^{n_i}\frac{p_{il}\Delta_i}{1+p_{il}\Delta_i}\right)^2}\left[\left(\frac{1}{N}\sum_{l=1}^{n_i}\frac{p_{il}\Delta_i}{1+p_{il}\Delta_i}\right)^2 - \frac{\bar{c}_i}{N}\sum_{l=1}^{n_i}\frac{(p_{il}\Delta_i)^2}{(1+p_{il}\Delta_i)^2}\right].$$

From the Cauchy-Schwarz inequality, we have

$$\left(\sum_{l=1}^{n_i}\frac{1}{N}\frac{p_{il}\Delta_i}{1+p_{il}\Delta_i}\right)^2 \leq \sum_{k=1}^{n_i}\frac{1}{N^2}\sum_{l=1}^{n_i}\frac{(p_{il}\Delta_i)^2}{(1+p_{il}\Delta_i)^2} = c_i\bar{c}_i\frac{1}{N}\sum_{l=1}^{n_i}\frac{(p_{il}\Delta_i)^2}{(1+p_{il}\Delta_i)^2} < \frac{\bar{c}_i}{N}\sum_{l=1}^{n_i}\frac{(p_{il}\Delta_i)^2}{(1+p_{il}\Delta_i)^2} \tag{33}$$

which is sufficient to conclude that $\frac{d}{d\Delta_i}\bar{x}_i < 0$. Since Δ_i is an increasing function of x_i, we have that $\bar{x}_i$ is a decreasing function of x_i, i.e., $\frac{d}{dx_i}\bar{x}_i < 0$. Therefore, for two sets $x_1,\dots,x_K$ and $x'_1,\dots,x'_K$ of positive values such that $x_j > x'_j$, defining $\bar{x}'_j$ equivalently as $\bar{x}_j$ for the terms x'_j, we have $\bar{x}'_k > \bar{x}_k$. Therefore, from Lemma 6, we finally have

$$h_j(x_1,\dots,x_K) - h_j(x'_1,\dots,x'_K) = \frac{1}{N}\operatorname{tr}\mathbf{R}_j\left(\sum_{k=1}^K \bar{x}_k\mathbf{R}_k - z\mathbf{I}_N\right)^{-1} - \frac{1}{N}\operatorname{tr}\mathbf{R}_j\left(\sum_{k=1}^K \bar{x}'_k\mathbf{R}_k - z\mathbf{I}_N\right)^{-1} > 0. \tag{34}$$

This proves the monotonicity condition and, finally, that $\mathbf{h} = (h_1, \ldots, h_K)$ is a standard function.

It follows from Theorem 8 that $(e_1, \ldots, e_K)$ is uniquely defined and that the classical fixed-point algorithm converges to this solution from any initialisation point (remember that, at each step of the algorithm, the set $\bar{e}_1, \ldots, \bar{e}_K$ must be evaluated, possibly thanks to a further fixed-point algorithm).

Consider now two sets of Stieltjes transforms $(e_1(z), \ldots, e_K(z))$ and $(e'_1(z), \ldots, e'_K(z))$, $z \in \mathbb{C} \setminus \mathbb{R}^+$, solutions of the fixed-point equation. Since $\sup_i |e_i(z) - e'_i(z)| = 0$ for all $z < 0$, and $e_i(z) - e'_i(z)$ is holomorphic on $\mathbb{C} \setminus \mathbb{R}^+$ as the difference of Stieltjes transforms, $e_i(z) - e'_i(z) = 0$ over $\mathbb{C} \setminus \mathbb{R}^+$ [38]. This therefore proves, in addition to *point-wise* uniqueness on the negative half-line, the uniqueness of the *Stieltjes transform* solution of the functional implicit equation such that, for $z < 0$, $0 \leq \bar{e}_i < c_i \bar{c}_i / e_i$ for all i.

This terminates the proof of Theorem 1.

Step 3: Convergence of $e_i - f_i$

For this step, we follow the same approach as in [5]. Denote

$$\varepsilon_{N,i} \triangleq f_i - \frac{1}{N} \operatorname{tr} \mathbf{R}_i \left(\sum_{k=1}^{K} \bar{f}_k \mathbf{R}_k - z \mathbf{I}_N \right)^{-1}$$

and recall the definitions of f_i, e_i, $\bar{f}_i$ and $\bar{e}_i$:

$$\begin{aligned}
f_i &= \frac{1}{N} \operatorname{tr} \mathbf{R}_i \left(\mathbf{B}_N - z \mathbf{I}_N \right)^{-1} \\
e_i &= \frac{1}{N} \operatorname{tr} \mathbf{R}_i \left(\sum_{j-1}^{K} \bar{e}_j \mathbf{R}_j - z \mathbf{I}_N \right)^{-1} \\
\bar{f}_i &= \frac{1}{N} \sum_{l=1}^{n_i} \frac{p_{il}}{\bar{c}_i - f_i \bar{f}_i + p_{il} f_i}, \qquad \bar{f}_i \in [0, c_i \bar{c}_i / f_i) \\
\bar{e}_i &= \frac{1}{N} \sum_{l=1}^{n_i} \frac{p_{il}}{\bar{c}_i - e_i \bar{e}_i + p_{il} e_i}, \qquad \bar{e}_i \in [0, c_i \bar{c}_i / e_i) \ .
\end{aligned}$$

From the definitions above, we have the following set of inequalities

$$f_i \leq \frac{R}{|z|}, \quad e_i \leq \frac{R}{|z|}, \quad \bar{f}_i \leq \frac{P}{(1 - c_i) \bar{c}_i}, \quad \bar{e}_i \leq \frac{P}{(1 - c_i) \bar{c}_i} \ . \tag{35}$$

We will show in the sequel that

$$e_i - f_i \xrightarrow{\text{a.s.}} 0, \tag{36}$$

for all $i \in \{1, \ldots, N\}$. Write the following differences

$$\begin{aligned}
f_i - e_i &= \sum_{j=1}^{K} (\bar{e}_j - \bar{f}_j) \frac{1}{N} \operatorname{tr} \mathbf{R}_i \left(\sum_{k=1}^{K} \bar{e}_k \mathbf{R}_k - z \mathbf{I}_N \right)^{-1} \mathbf{R}_j \left(\sum_{k=1}^{K} \bar{f}_k \mathbf{R}_k - z \mathbf{I}_N \right)^{-1} + \varepsilon_{N,i} \\
\bar{e}_i - \bar{f}_i &= \frac{1}{N} \sum_{l=1}^{n_i} \frac{p_{il}^2 (f_i - e_i) - p_{il} \left[f_i \bar{f}_i - e_i \bar{e}_i \right]}{(\bar{c}_i - \bar{e}_i e_i + p_{il} e_i)(\bar{c}_i - \bar{f}_i f_i + p_{il} f_i)}
\end{aligned}$$

$$f_i \bar{f}_i - e_i \bar{e}_i = \bar{f}_i (f_i - e_i) + e_i (\bar{f}_i - \bar{e}_i) \ .$$

For notational convenience, we define the following values

$$\alpha \triangleq \sup_i \mathbb{E}\left[|f_i - e_i|^4\right]$$
$$\bar{\alpha} \triangleq \sup_i \mathbb{E}\left[|\bar{f}_i - \bar{e}_i|^4\right] \ .$$

It is thus sufficient to show that α is summable to prove (36). By applying (35) to the absolute of the first difference, we obtain

$$|f_i - e_i| \leq \frac{KR^2}{|z|^2} \sup_i |\bar{f}_i - \bar{e}_i| + \sup_i |\varepsilon_{N,i}|$$

and hence

$$\alpha \leq \frac{8K^4R^8}{|z|^8}\bar{\alpha} + \frac{8C}{N^2} \tag{37}$$

for some $C > 0$ such that $\mathbb{E}[\sup_i |\varepsilon_{N,i}|^4] \leq 8K \sup_i \mathbb{E}[|\varepsilon_{N,i}|^4] \leq C/N^2$. Similarly, we have for the third difference

$$\begin{aligned} |f_i\bar{f}_i - e_i\bar{e}_i| &\leq |\bar{f}_i||f_i - e_i| + |e_i||\bar{f}_i - \bar{e}_i| \\ &\leq \frac{P}{(1-c_+)\bar{c}_-} \sup_i |f_i - e_i| + \frac{R}{|z|} \sup_i |\bar{f}_i - \bar{e}_i| \ . \end{aligned}$$

This result can be used to upperbound the second difference term, which writes

$$\begin{aligned} |\bar{f}_i - \bar{e}_i| &\leq \frac{1}{(1-c_+)^2\bar{c}_-^2}\left(P^2 \sup_i |f_i - e_i| + P|f_i\bar{f}_i - e_i\bar{e}_i|\right) \\ &\leq \frac{1}{(1-c_+)^2\bar{c}_-^2}\left(P^2 \sup_i |f_i - e_i| + P\left[\frac{P}{(1-c_+)\bar{c}_-} \sup_i |f_i - e_i| + \frac{R}{|z|} \sup_i |\bar{f}_i - \bar{e}_i|\right]\right) \\ &\leq \frac{P^2(\bar{c}_- + 1)}{(1-c_+)^3\bar{c}_-^3} \sup_i |f_i - e_i| + \frac{RP}{|z|(1-c_+)^2\bar{c}_-^2} \sup_i |\bar{f}_i - \bar{e}_i| \ . \end{aligned}$$

Hence

$$\bar{\alpha} \leq \frac{8P^8(\bar{c}_- + 1)^4}{(1-c_+)^{12}\bar{c}_-^{12}}\alpha + \frac{8R^4P^4}{|z|^4(1-c_+)^8\bar{c}_-^8}\bar{\alpha} \ . \tag{38}$$

For any z satisfying $|z| > \frac{2RP}{(1-c_+)^2}$, we have $\frac{8R^4P^4}{|z|^4(1-c_+)^8} < 1/2$ and thus

$$\bar{\alpha} < \frac{16P^8(\bar{c}_- + 1)^4}{(1-c_+)^{12}\bar{c}_-^{12}}\alpha \ .$$

Plugging this result into (37) yields

$$\alpha \leq \frac{128K^4R^8P^8(2-c)^4}{|z|^8(1-c_+)^{12}}\alpha + \frac{8C}{N^2} \ .$$

Take $0 < \varepsilon < 1$. It is easy to check that for $|z| > \frac{128^{1/8}RP\sqrt{K(\bar{c}_-+1)}}{(1-c_+)^{3/2}\bar{c}_-^{3/2}(1-\varepsilon)^{1/8}}$, $\frac{128K^4R^8P^8(\bar{c}_-+1)^4}{|z|^8(1-c_+)^{12}\bar{c}_-^{12}} < 1 - \varepsilon$ and thus

$$\alpha < \frac{8C}{\varepsilon N^2} \ . \tag{39}$$

Since C does not depend on N, α is clearly summable which, along with Markov inequality and the Borel Cantelli lemma, concludes the proof.

Finally, taking the same steps as previously, we also have

$$\mathbb{E}\left[|m_N(z) - \bar{m}_N(z)|^4\right] \leq \frac{8C}{\varepsilon N^2}$$

for some $|z|$ large enough. For these z, the same conclusion holds: $m_N(z) - \bar{m}_N(z) \xrightarrow{\text{a.s.}} 0$. From Vitali convergence theorem, since f_i and e_i are uniformly bounded on all closed sets of $\mathbb{C} \setminus \mathbb{R}^+$, we finally have that the convergence is true for all $z \in \mathbb{C} \setminus \mathbb{R}^+$. The almost sure convergence of the Stieltjes transform implies the almost sure weak convergence of $F_N - \bar{F}_N$ to 0, uniformly over every compact set of $\mathbb{R}^+$, which is our final result.

This concludes the proof of Theorem 7 for $\limsup_i c_i < 1$ and surely bounded $\mathbf{R}_i$.

2) Case $\max_i \limsup_i c_i = 1$*:*

We now need to extend the previous result to the case $\limsup_i c_i = 1$ for some i. The previous approach no longer holds as Lemma 5 is no longer valid. We will assume here without loss of generality that $c_1 = \ldots = c_K = 1$. Since $\mathbf{P}_1, \ldots, \mathbf{P}_K$ are allowed to have null eigenvalues, this assumption also covers the case presented in the previous section. Observe in particular that this assumption does not alter the fundamental equations in $(e_1, \ldots, e_K, \bar{e}_1, \ldots, \bar{e}_K)$ that do not depend on the c_i.

For a given matrix $\mathbf{B}_N$, we now define the matrix $\mathbf{B}_N^{(n)}$ as

$$\mathbf{B}_N^{(n)} = \mathbf{B}_N - \sum_{i=1}^{K} \sum_{l_i=n+1}^{N_i} p_{il_i} \mathbf{H}_i \mathbf{w}_{il_i} \mathbf{w}_{il_i}^{\mathsf{H}} \mathbf{H}_i^{\mathsf{H}}.$$

That is, $\mathbf{B}_N^{(n)}$ corresponds to $\mathbf{B}_N$ with all columns of $\mathbf{H}_i \mathbf{W}_i \mathbf{P}_i \mathbf{W}_i^{\mathsf{H}} \mathbf{H}_i^{\mathsf{H}}$ of index superior to n discarded. We will further define $c = \lim n/N$.

Similarly, we shall denote $\mathbf{P}_i^{(n)} = \text{diag}(p_{i,1}, \ldots, p_{i,n})$, $e_i^{(n)}$ and $\bar{e}_i^{(n)}$ the unique solutions to

$$\bar{e}_i^{(n)} = \frac{1}{N} \operatorname{tr} \mathbf{P}_i^{(n)} \left(e_i^{(n)} \mathbf{P}_i^{(n)} + [\bar{c}_i - \bar{e}_i^{(n)} e_i^{(n)}] \mathbf{I}_n \right)^{-1}$$

$$e_i^{(n)} = \frac{1}{N} \operatorname{tr} \mathbf{R}_i \left(\sum_{j=1}^{K} \bar{e}_j^{(n)} \mathbf{R}_j - z \mathbf{I}_N \right)^{-1}$$

such that $\bar{e}_i^{(n)} \in [0, c\bar{c}_i / e_i^{(n)})$ and

$$\bar{m}_N^{(n)} = \frac{1}{N} \operatorname{tr} \left(\sum_{j=1}^{K} \bar{e}_j^{(n)} \mathbf{R}_j - z \mathbf{I}_N \right)^{-1}.$$

We will prove that the Stieltjes transforms and its deterministic equivalent for $\mathbf{B}_N$ and $\mathbf{B}_N^{(n)}$ are within any $\varepsilon > 0$ for n chosen such that $n/N \to c$ for some $c < 1$. This will ensure that for all large N, the Stieltjes transform of $\mathbf{B}_N$ and its deterministic equivalent are within 2ε in the large N limit, ε being arbitrary. This will complete the proof.

We start by proving the uniqueness of the solution to the fundamental equations for $\mathbf{B}_N$. The only step that needs to be modified compared to the proof for $\limsup c_i < 1$ lies in (33) where the strict inequality (due to $c_i < 1$) becomes a loose inequality. We then see that (34) becomes an equality if and only if the equality (33) is established for all i. This requires, from the statement of the Cauchy-Schwarz theorem, that, for each given i, all p_{il} be equal. But this means that $\mathbf{P}_i = t_i \mathbf{I}_{N_i}$ for all i and for some $t_i \geq 0$ and therefore $\mathbf{B}_N$ becomes $\sum_{k=1}^K t_i \mathbf{R}_i$, which is deterministic. Since this case is trivial, we discard it and assume that for at least one i, $\mathbf{P}_i$ is not proportional to the identity matrix. This implies that the difference (34) is positive and therefore $\mathbf{h}$ is still a standard function. By noticing that $e_i \leq \frac{R}{|z|}$, we necessarily have that

$$\frac{R}{|z|} \geq h_i\left(\frac{R}{|z|}, \dots, \frac{R}{|z|}\right)$$

and therefore Theorem 8 can be applied to $\mathbf{h}$.

We can therefore uniquely define $e_1, \dots, e_K, \bar{e}_1, \dots, \bar{e}_K$ the solution to

$$\begin{aligned}
\bar{e}_i &= \frac{1}{N} \operatorname{tr} \mathbf{P}_i \left(e_i \mathbf{P}_i + [\bar{c}_i - \bar{e}_i e_i] \mathbf{I}_N\right)^{-1} \\
e_i &= \frac{1}{N} \operatorname{tr} \mathbf{R}_i \left(\sum_{j=1}^K \bar{e}_j \mathbf{R}_j - z \mathbf{I}_N\right)^{-1}
\end{aligned}$$

such that $\bar{e}_i \in [0, \bar{c}_i / e_i)$.[2] We can therefore uniquely define $\bar{m}_N(z)$ the holomorphic function equal to

$$\bar{m}_N(z) = \frac{1}{N} \operatorname{tr} \left(\sum_{j=1}^K \bar{e}_j \mathbf{R}_j - z \mathbf{I}_N\right)^{-1} .$$

One of the major problems we will face here is that the former inequality $\bar{e}_i \leq \frac{P}{(1-c_i)\bar{c}_i}$ is no longer useful when $c_i = 1$. We need to refine this inequality with the following remark.

Note that we have from the definitions above

$$\begin{aligned}
\bar{c}_i &= \frac{1}{N} \sum_{l=1}^{N_i} \frac{\bar{c}_i - e_i \bar{e}_i + e_i p_{il}}{\bar{c}_i - e_i \bar{e}_i + e_i p_{il}} \\
&= (\bar{c}_i - e_i \bar{e}_i) \frac{1}{N} \sum_{l=1}^{N_i} \frac{1}{\bar{c}_i - e_i \bar{e}_i + e_i p_{il}} + \frac{1}{N} \sum_{l=1}^{N_i} \frac{e_i p_{il}}{\bar{c}_i - e_i \bar{e}_i + e_i p_{il}} \\
&= (\bar{c}_i - e_i \bar{e}_i) \frac{1}{N} \sum_{l=1}^{N_i} \frac{1}{\bar{c}_i - e_i \bar{e}_i + e_i p_{il}} + e_i \bar{e}_i
\end{aligned}$$

from which follows that

$$(\bar{c}_i - e_i \bar{e}_i) \left(1 - \frac{1}{N} \sum_{l=1}^{N_i} \frac{1}{\bar{c}_i - e_i \bar{e}_i + e_i p_{il}}\right) = 0 .$$

But we also know that $0 \leq \bar{e}_i < \bar{c}_i / e_i$ and therefore $\bar{c}_i - e_i \bar{e}_i > 0$. This entails

$$\frac{1}{N} \sum_{l=1}^{N_i} \frac{1}{\bar{c}_i - e_i \bar{e}_i + e_i p_{il}} = 1 . \tag{40}$$

Since this sums to 1, necessarily $\max_l (\bar{c}_i - e_i \bar{e}_i + e_i p_{il}) \geq \bar{c}_i$ or equivalently $e_i \bar{e}_i \leq e_i \max_l p_{il}$. Since $e_i > 0$ whenever one of the $\mathbf{R}_i$ is non identically zero, this entails $\bar{e}_i \leq \max_l (p_{il})$. Hence, we can state the refined

[2]Note that, if all p_{il} are non zero, $\bar{c}_i / e_i$ is the second solution of the implicit equation in $\bar{e}_i$, which has to be excluded from the interval. Hence the importance of opening the right edge of the interval.

inequality

$$\bar{e}_i \leq P.$$

We are now in position to complete the proof. Following the approach pursued in Step 3, we have the following differences

$$\begin{aligned}
e_i^{(n)} - e_i &= \sum_{j=1}^{K} (\bar{e}_j - \bar{e}_j^{(n)}) \frac{1}{N} \operatorname{tr} \mathbf{R}_i \left(\sum_{k=1}^{K} \bar{e}_k \mathbf{R}_k - z\mathbf{I}_N \right)^{-1} \mathbf{R}_j \left(\sum_{k=1}^{K} \bar{e}_k^{(n)} \mathbf{R}_k - z\mathbf{I}_N \right)^{-1} \\
\bar{e}_i - \bar{e}_i^{(n)} &= \frac{1}{N} \sum_{l=1}^{n} \frac{p_{il}^2 (e_i^{(n)} - e_i) - p_{il} \left[e_i^{(n)} \bar{e}_i^{(n)} - e_i \bar{e}_i \right]}{(\bar{c}_i + p_{il} e_i - \bar{e}_i e_i)(\bar{c}_i + p_{il} e_i^{(n)} - \bar{e}_i^{(n)} e_i^{(n)})} + \frac{1}{N} \sum_{l=n+1}^{N_i} \frac{p_{il}}{\bar{c}_i + p_{il} e_i - e_i \bar{e}_i} \\
e_i^{(n)} \bar{e}_i^{(n)} - e_i \bar{e}_i &= \bar{e}_i^{(n)} (e_i^{(n)} - e_i) + e_i (\bar{e}_i^{(n)} - \bar{e}_i).
\end{aligned}$$

Remembering that $e_i^{(n)} \to 0$ whenever $z \to -\infty$ (irrespective of N or c), and noticing, due to $\bar{e}_i \leq P$, that we also have $e_i \to 0$ whenever $z \to -\infty$, we can set $z < z_0$ for some $z_0 < 0$ to be such that

$$\min \left(\bar{c}_i - e_i \bar{e}_i + e_i p_{il}, \bar{c}_i - e_i^{(n)} \bar{e}_i^{(n)} + e_i^{(n)} p_{il} \right) \geq \eta$$

for some $\eta > 0$. For these z, we therefore have

$$\begin{aligned}
|e_i^{(n)} - e_i| &\leq \frac{KR^2}{|z|^2} \sup_i |\bar{e}_i - \bar{e}_i^{(n)}| \qquad (41) \\
|\bar{e}_i - \bar{e}_i^{(n)}| &\leq \frac{P^2}{\eta^2} \sup_i |e_i^{(n)} - e_i| + \frac{P}{\eta^2} \sup_i |e_i^{(n)} \bar{e}_i^{(n)} - e_i \bar{e}_i| + \bar{c}_+ (1-c) \frac{P}{\eta} \\
|e_i^{(n)} \bar{e}_i^{(n)} - e_i \bar{e}_i| &\leq P \sup_i |e_i^{(n)} - e_i| + \frac{R}{|z|} \sup_i |\bar{e}_i - \bar{e}_i^{(n)}|.
\end{aligned}$$

Denoting $\beta \triangleq \sup_i |\bar{e}_i - \bar{e}_i^{(n)}|$, together this implies

$$\beta \leq \beta \left[\frac{2KP^2R^2}{|z|^2\eta^2} + \frac{PR}{|z|\eta^2} \right] + \bar{c}_+ (1-c) \frac{P}{\eta}.$$

We now take $z < z_0$ to be such that $\frac{2KP^2R^2}{|z|^2\eta^2} + \frac{PR}{|z|\eta^2} \leq 1 - \kappa$ for some κ such that $0 < \kappa < 1$ (note that κ is chosen independently of N or c). Therefore, for these z,

$$\beta \leq \bar{c}_+ (1-c) \frac{P}{\kappa\eta}. \qquad (42)$$

The same reasoning holds for $\bar{m}_N^{(n)}$ in the sense that, there exists $z_1 < 0$, such that for $z < z_1$,

$$|\bar{m}_N(z) - \bar{m}_N^{(n)}(z)| \leq \bar{c}_+ (1-c) \frac{P}{\kappa\eta}.$$

Now, for any matrix $\mathbf{A} \in \mathbb{C}^{N \times N}$ with spectral norm bounded by A, we also have from $K(N-n)$ iterations of the rank-1 perturbation (Lemma 7), that

$$\left| \frac{1}{N} \operatorname{tr} \mathbf{A} \left(\mathbf{B}_N - z\mathbf{I}_N \right)^{-1} - \frac{1}{N} \operatorname{tr} \mathbf{A} \left(\mathbf{B}_N^{(n)} - z\mathbf{I}_N \right)^{-1} \right| \leq \frac{1}{N} \sum_{i=1}^{K} \sum_{l_i=n+1}^{N_i} \frac{A}{|z|} \leq \bar{c}_+ (1-c) \frac{KA}{|z|}.$$

Take $\varepsilon > 0$. With $z < z_1$, one can now choose $c < 1$ sufficiently close to 1 such that

$$\max \left(\bar{c}_+ (1-c) \frac{P}{\kappa\eta}, \bar{c}_+ (1-c) \frac{KA}{|z|} \right) < \varepsilon.$$

Letting n and N grow large, with $n/N \leq c < (n+1)/N$, we have

$$\left|\frac{1}{N}\operatorname{tr}\mathbf{R}_i\left(\mathbf{B}_N - z\mathbf{I}_N\right)^{-1} - e_i(z)\right| \leq \left|\frac{1}{N}\operatorname{tr}\mathbf{R}_i\left(\mathbf{B}_N - z\mathbf{I}_N\right)^{-1} - \frac{1}{N}\operatorname{tr}\mathbf{R}_i\left(\mathbf{B}_N^{(n)} - z\mathbf{I}_N\right)^{-1}\right| + \left|\frac{1}{N}\operatorname{tr}\mathbf{R}_i\left(\mathbf{B}_N^{(n)} - z\mathbf{I}_N\right)^{-1} - e_i^{(n)}(z)\right| + \left|e_i^{(n)}(z) - e_i(z)\right|$$

for $\mathbf{A} = \mathbf{R}_i$ and

$$|m_N(z) - \bar{m}_N(z)| \leq \left|m_N(z) - m_N^{(n)}(z)\right| + \left|m_N^{(n)}(z) - \bar{m}_N^{(n)}(z)\right| + \left|\bar{m}_N^{(n)}(z) - \bar{m}_N(z)\right|$$

for $\mathbf{A} = \mathbf{I}_N$.

Taking the limit superior for N on both sides, we finally have

$$\limsup_N \left|\frac{1}{N}\operatorname{tr}\mathbf{R}_i\left(\mathbf{B}_N - z\mathbf{I}_N\right)^{-1} - e_i(z)\right| < 2\varepsilon$$

almost surely and

$$\limsup_N |m_N(z) - \bar{m}_N(z)| < 2\varepsilon$$

almost surely, since we have proved in the previous section that $\frac{1}{N}\operatorname{tr}\mathbf{R}_i\left(\mathbf{B}_N^{(n)} - z\mathbf{I}_N\right)^{-1} - e_i^{(n)}(z) \xrightarrow{\text{a.s.}} 0$ and $m_{\mathbf{B}_N^{(n)}}(z) - \bar{m}_N^{(n)}(z) \xrightarrow{\text{a.s.}} 0$.

Since ε was arbitrary, this means that

$$\frac{1}{N}\operatorname{tr}\mathbf{R}_i\left(\mathbf{B}_N - z\mathbf{I}_N\right)^{-1} - e_i(z) \xrightarrow{\text{a.s.}} 0$$

and

$$m_N(z) - \bar{m}_N(z) \xrightarrow{\text{a.s.}} 0.$$

Since $m_N(z)$ and $\bar{m}_N(z)$ are uniformly bounded on all compact sets of $\mathbb{C}\backslash\mathbb{R}^+$, from Vitali convergence theorem, we finally have $m_N(z) - \bar{m}_N(z) \xrightarrow{\text{a.s.}} 0$ for all $z < 0$. The uniqueness of holomorphic functions defined on a set with a cluster point then ensures the uniqueness of the Stieltjes transform $\bar{m}_N(z)$ for $z \in \mathbb{C} \setminus \mathbb{R}^+$.

We complete the proof with the relaxation of the constraint $\|\mathbf{R}_i\| \leq R$ surely to $\|\mathbf{R}_i\| \leq R$ almost surely.

3) Almost sure boundedness of $\|\mathbf{R}_i\|$:

To extend Theorem 7 to the case where $\|\mathbf{R}_i\|$ is only almost surely bounded, we merely apply the Tonelli theorem (Lemma 9).

Call $(\Omega_R, \mathcal{F}_R, P_R)$ the probability space that generates the sequences of matrices of growing sizes $\{\mathbf{R}_i, 1 \leq i \leq K, N_i \in \mathbb{N}\}$ and $(\Omega_W, \mathcal{F}_W, P_W)$ the probability space that generates the sequences of matrices of growing sizes $\{\mathbf{W}_i, 1 \leq i \leq K, N_i \in \mathbb{N}\}$ and $(\Omega_R \times \Omega_W, \mathcal{F}_R \times \mathcal{F}_W, Q)$. Denote A the subspace of $\mathcal{F}_R \times \mathcal{F}_W$ for which $F_N - \bar{F}_N \to 0$. Then, from Tonelli theorem,

$$Q(A) = \int_{\Omega_R \times \Omega_W} 1_A(r,w) Q(d(r,w)) = \int_{\Omega_R}\int_{\Omega_W} 1_A(r,w) P_W(dw) P_R(dr).$$

Take r such that the $\|\mathbf{R}_i\|$ are all uniformly bounded with growing N. Then, from Theorem 7, for this r, $\int_{\Omega_W} 1_A(r,w)P_W(dw) = 1$. But these $r \in \Omega_R$ belong to a space of probability one, as the intersection of K spaces of probability one, and finally $Q(A) = 1$.

APPENDIX B
PROOF OF THEOREM 3

It is easy to see (e.g. [11, Definition 3.2]) that, for F a probability distribution function with support in $\mathbb{R}^+$

$$\int_0^\infty \log\left(1+\frac{t}{x}\right)dF(t) = \int_x^\infty \left(-\frac{1}{t}+m_F(-t)\right)dF(t)$$

where $m_F(z)$ is the Stieltjes transform of F (this is sometimes called the Shannon-transform in $1/x$). In particular,

$$I_N^{(a)}(\sigma^2) = \frac{1}{N}\log\det\left(\mathbf{I}_N+\frac{1}{\sigma^2}\mathbf{B}_N\right) = \int_{\sigma^2}^\infty \left(-\frac{1}{t}+m_N(-t)\right)dF_N(t).$$

We will first show that the expression $\bar{I}_N^{(a)}(\sigma^2)$ given in Theorem 3 satisfies the same property with $\bar{F}_N$.

For notational simplicity, we will write $e_i = e_i(-\sigma^2)$ and $\bar{e}_i = \bar{e}_i(-\sigma^2)$.

We take here $c_i \leq 1$ from the beginning. First note that the system of equations (13) is unchanged if we extend the $\mathbf{P}_i$ matrices into $N_i \times N_i$ diagonal matrices filled with $N_i - n_i$ zero eigenvalues. Therefore, we can assume that all $\mathbf{P}_i$ have size $N_i \times N_i$ although we restrict the $F^{\mathbf{P}_i}$ to have a mass $1-c_i$ in zero. Since this does not alter the equations (13), we have in particular $\bar{e}_i < \bar{c}_i/e_i$ for $\sigma^2 > 0$.

This being said, $\bar{I}_N^{(a)}$ is given by

$$\bar{I}_N^{(a)}(\sigma^2) = \frac{1}{N}\log\det\left(\mathbf{I}_N+\frac{1}{\sigma^2}\sum_{i=1}^K \bar{e}_i\mathbf{R}_i\right) + \sum_{i=1}^K\left[\frac{1}{N}\log\det\left([\bar{c}_i - e_i\bar{e}_i]\mathbf{I}_N + e_i\mathbf{P}_i\right) - \bar{c}_i\log(\bar{c}_i)\right].$$

Calling $\bar{I}$ the function

$$\begin{aligned}
&\bar{I} : (x_1,\ldots,x_K,\bar{x}_1,\ldots,\bar{x}_K,\sigma^2) \\
&\mapsto \frac{1}{N}\log\det\left(\mathbf{I}_N+\frac{1}{\sigma^2}\sum_{i=1}^K \bar{x}_i\mathbf{R}_i\right) + \sum_{i=1}^K\left[\frac{1}{N}\log\det\left([\bar{c}_i - x_i\bar{x}_i]\mathbf{I}_N + x_i\mathbf{P}_i\right) - \bar{c}_i\log(\bar{c}_i)\right],
\end{aligned}$$

we have

$$\begin{aligned}
\frac{\partial \bar{I}}{\partial x_i}(e_1,\ldots,e_K,\bar{e}_1,\ldots,\bar{e}_K,\sigma^2) &= \bar{e}_i - \bar{e}_i\frac{1}{N}\sum_{l=1}^{N_i}\frac{1}{\bar{c}_i - e_i\bar{e}_i + e_i p_{il}} \\
\frac{\partial \bar{I}}{\partial \bar{x}_i}(e_1,\ldots,e_K,\bar{e}_1,\ldots,\bar{e}_K,\sigma^2) &= e_i - e_i\frac{1}{N}\sum_{l=1}^{N_i}\frac{1}{\bar{c}_i - e_i\bar{e}_i + e_i p_{il}}.
\end{aligned}$$

From (40), we conclude that

$$\begin{aligned}
\frac{\partial \bar{I}}{\partial x_i}(e_1,\ldots,e_K,\bar{e}_1,\ldots,\bar{e}_K,\sigma^2) &= 0 \\
\frac{\partial \bar{I}}{\partial \bar{x}_i}(e_1,\ldots,e_K,\bar{e}_1,\ldots,\bar{e}_K,\sigma^2) &= 0.
\end{aligned}$$

We therefore have, from the differentiation chain rule,

$$
\begin{aligned}
\frac{d}{d\sigma^2}\bar{I}_N^{(a)}(\sigma^2) &= \sum_{i=1}^{K}\left[\frac{\partial \bar{I}}{\partial e_i}\frac{\partial e_i}{\partial \sigma^2}+\frac{\partial \bar{I}}{\partial \bar{e}_i}\frac{\partial \bar{e}_i}{\partial \sigma^2}\right]+\frac{\partial \bar{I}}{\partial \sigma^2} \\
&= \frac{\partial \bar{I}}{\partial \sigma^2} \\
&= -\frac{1}{\sigma^4}\sum_{i=1}^{K}\bar{e}_i\frac{1}{N}\operatorname{tr}\mathbf{R}_i\left(\mathbf{I}_N+\frac{1}{\sigma^2}\sum_{j=1}^{K}\bar{e}_j\mathbf{R}_j\right)^{-1} \\
&= -\frac{1}{\sigma^2}\frac{1}{N}\operatorname{tr}\left[\left(\sum_{i=1}^{K}\frac{1}{\sigma^2}\bar{e}_i\mathbf{R}_i+\mathbf{I}_N-\mathbf{I}_N\right)\left(\mathbf{I}_N+\frac{1}{\sigma^2}\sum_{j=1}^{K}\bar{e}_j\mathbf{R}_j\right)^{-1}\right] \\
&= -\frac{1}{\sigma^2}+\frac{1}{N}\operatorname{tr}\left(\sigma^2\mathbf{I}_N+\sum_{j=1}^{K}\bar{e}_j\mathbf{R}_j\right)^{-1}
\end{aligned}
$$

Recognizing the Stieltjes transform of $\bar{F}_N$, we therefore have, along with the fact that $\bar{I}_N^{(a)}(\infty)=0$,

$$
\bar{I}_N^{(a)}(\sigma^2)=\int_{\sigma^2}^{\infty}\left(\frac{1}{t}-\frac{1}{t^2}\bar{m}_N\left(-\frac{1}{t}\right)\right)dt
$$

and therefore

$$
\bar{I}_N^{(a)}(\sigma^2)=\int_0^{\infty}\log\left(1+\frac{t}{\sigma^2}\right)d\bar{F}_N(t).
$$

In order to prove the almost sure convergence $I_N^{(a)}(\sigma^2)-\bar{I}_N^{(a)}(\sigma^2)\xrightarrow{\text{a.s.}}0$, we simply need to remark that the support of the eigenvalues of $\mathbf{B}_N$ is bounded. Indeed, the non-zero eigenvalues of $\mathbf{W}_i\mathbf{W}_i^{\mathsf{H}}$ have unit modulus and therefore $\|\mathbf{B}_N\|\leq KPR$. Similarly, the support of $\bar{F}_N$ is the support of the eigenvalues of $\sum_{i=1}^{K}\bar{e}_i\mathbf{R}_i$, which are bounded by KPR as well.

As a consequence, for $\mathbf{B}_1,\mathbf{B}_2,\ldots$ a realization for which $F_N-\bar{F}_N\Rightarrow 0$ (these lie in a space of probability one), we have, from the dominated convergence theorem

$$
\int_0^{\infty}\log\left(1+\frac{t}{\sigma^2}\right)d[F_N-\bar{F}_N](t)\to 0
$$

Hence the almost sure convergence of the instantaneous mutual information.

Because of sure boundedness of $\|\mathbf{B}_N\|$, an immediate application of the dominated convergence theorem on the probability space Ω that engenders the sequences of matrices $\mathbf{B}_1(\omega),\mathbf{B}_2(\omega),\ldots$, $\omega\in\Omega$, entails convergence in the first mean as well.

APPENDIX C
PROOF OF THEOREM 5

To prove Theorem 5, we will pursue a similar approach as for the proof of Theorem 7, but we can now take advantage of all results derived so far.

First denote d_i the unique positive solution, for $e_i>0$, to

$$
e_i=d_i\left(\bar{c}_i-\frac{1}{N}\sum_{l=1}^{n_i}\frac{p_{il}d_i}{1+p_{il}d_i}\right).
$$

This solution exists and is unique due to the arguments given in the introduction of Step 2 of the proof of Theorem 7.

Whatever the value of c_i, we will proceed as previously by extending the matrix $\mathbf{P}_i$ to an N_i-dimensional matrix with the last $N_i - n_i$ diagonal entries filled with zeros. This way, we can write

$$e_i = d_i \left(\frac{1}{N} \sum_{l=1}^{N_i} 1 - \frac{p_{il} d_i}{1 + p_{il} d_i} \right) = \frac{1}{N} \sum_{l=1}^{N_i} \frac{d_i}{1 + p_{il} d_i}.$$

Since d_i is a continuous mapping of e_i and $e_i \leq \frac{P}{|z|}$, it follows that d_i is bounded from above.

Remember now that for $\limsup c_i < 1$ for all i and, for some $z_0 < 0$, we have that $z < z_0$ implies

$$\mathbb{E}[|f_i - e_i|^4] = \mathbb{E}\left[\left| f_i - \frac{1}{N} \sum_{l=1}^{N_i} \frac{d_i}{1 + p_{il} d_i} \right|^4 \right] \leq \frac{C}{N^2}$$

for some constant $C > 0$. Also, from (17),

$$\mathbb{E}\left[\left| f_i - \frac{1}{N} \sum_{l=1}^{N_i} \frac{\delta_i}{1 + p_{il} \delta_i} \right|^4 \right] \leq \frac{C_1}{N^2}$$

for some $C_1 > C$. From these two inequalities, we have

$$\mathbb{E}\left[\left| \frac{1}{N} \sum_{l=1}^{N_i} \frac{\delta_i}{1 + p_{il} \delta_i} - \frac{1}{N} \sum_{l=1}^{N_i} \frac{d_i}{1 + p_{il} d_i} \right|^4 \right] \leq \frac{16 C_1}{N^2}.$$

Also, from an immediate application of the trace lemma, Lemma 5, we remind that

$$\mathbb{E}\left[\left| \mathbf{w}_{il}^{\mathsf{H}} \mathbf{H}_i^{\mathsf{H}} \left(\mathbf{B}_{(i,l)} - z \mathbf{I}_N \right)^{-1} \mathbf{H}_i \mathbf{w}_{il} - \delta_i \right|^4 \right] \leq \frac{C_2}{N^2}$$

for some $C_2 > C_1$.

Together, this implies that for z small enough and for any $k \in \{1, \ldots, n_k\}$,

$$\begin{aligned}
& \mathbb{E}\left[\left| \frac{1}{N} \sum_{l=1}^{N_i} \frac{d_i}{1 + p_{il} d_i} - \frac{1}{N} \sum_{l=1}^{N_i} \frac{\mathbf{w}_{ik}^{\mathsf{H}} \mathbf{H}_i^{\mathsf{H}} \left(\mathbf{B}_{(i,k)} - z \mathbf{I}_N \right)^{-1} \mathbf{H}_i \mathbf{w}_{ik}}{1 + p_{il} \mathbf{w}_{ik}^{\mathsf{H}} \mathbf{H}_i^{\mathsf{H}} \left(\mathbf{B}_{(i,k)} - z \mathbf{I}_N \right)^{-1} \mathbf{H}_i \mathbf{w}_{ik}} \right|^4 \right] \\
\leq \quad & 8 \left[\mathbb{E}\left[\left| \frac{1}{N} \sum_{l=1}^{N_i} \frac{d_i}{1 + p_{il} d_i} - \frac{1}{N} \sum_{l=1}^{N_i} \frac{\delta_i}{1 + p_{il} \delta_i} \right|^4 \right] \right. \\
& \left. + \mathbb{E}\left[\left| \frac{1}{N} \sum_{l=1}^{N_i} \frac{\delta_i}{1 + p_{il} \delta_i} - \frac{1}{N} \sum_{l=1}^{N_i} \frac{\mathbf{w}_{ik}^{\mathsf{H}} \mathbf{H}_i^{\mathsf{H}} \left(\mathbf{B}_{(i,k)} - z \mathbf{I}_N \right)^{-1} \mathbf{H}_i \mathbf{w}_{ik}}{1 + p_{il} \mathbf{w}_{ik}^{\mathsf{H}} \mathbf{H}_i^{\mathsf{H}} \left(\mathbf{B}_{(i,k)} - z \mathbf{I}_N \right)^{-1} \mathbf{H}_i \mathbf{w}_{ik}} \right|^4 \right] \right] \\
= \quad & 8 \left[\mathbb{E}\left[\left| \frac{1}{N} \sum_{l=1}^{N_i} \frac{d_i}{1 + p_{il} d_i} - \frac{1}{N} \sum_{l=1}^{N_i} \frac{\delta_i}{1 + p_{il} \delta_i} \right|^4 \right] \right. \\
& \left. + \mathbb{E}\left[\left| \frac{1}{N} \sum_{l=1}^{N_i} \frac{\delta_i - \mathbf{w}_{ik}^{\mathsf{H}} \mathbf{H}_i^{\mathsf{H}} \left(\mathbf{B}_{(i,k)} - z \mathbf{I}_N \right)^{-1} \mathbf{H}_i \mathbf{w}_{ik}}{(1 + p_{il} \delta_i)(1 + p_{il} \mathbf{w}_{ik}^{\mathsf{H}} \mathbf{H}_i^{\mathsf{H}} \left(\mathbf{B}_{(i,k)} - z \mathbf{I}_N \right)^{-1} \mathbf{H}_i \mathbf{w}_{ik})} \right|^4 \right] \right] \\
\leq \quad & \frac{136 C_2}{N^2} \ .
\end{aligned}$$

This ensures that for $z < z_0$,

$$\frac{1}{N} \sum_{l=1}^{N_i} \frac{d_i}{1 + p_{il} d_i} - \frac{1}{N} \sum_{l=1}^{N_i} \frac{\mathbf{w}_{ik}^{\mathsf{H}} \mathbf{H}_i^{\mathsf{H}} \left(\mathbf{B}_{(i,k)} - z \mathbf{I}_N \right)^{-1} \mathbf{H}_i \mathbf{w}_{ik}}{1 + p_{il} \mathbf{w}_{ik}^{\mathsf{H}} \mathbf{H}_i^{\mathsf{H}} \left(\mathbf{B}_{(i,k)} - z \mathbf{I}_N \right)^{-1} \mathbf{H}_i \mathbf{w}_{ik}} \xrightarrow{\text{a.s.}} 0 \tag{43}$$

irrespectively of the choice of k.

Since the function $f : x \mapsto \frac{1}{N}\sum_{l=1}^{N_i} \frac{x}{1+p_{il}x}$ is continuous and has positive derivative, it is a one-to-one continuous function. Therefore, for $\mathbf{B}_1, \mathbf{B}_2, \ldots$ a realization such that the convergence of (43) is ensured, we also have by continuity $d_i - \mathbf{w}_{ik}^{\mathsf{H}}\mathbf{H}_i^{\mathsf{H}}\left(\mathbf{B}_{(i,k)} - z\mathbf{I}_N\right)^{-1}\mathbf{H}_i\mathbf{w}_{ik} \to 0$. Finally,

$$d_i - \mathbf{w}_{ik}^{\mathsf{H}}\mathbf{H}_i^{\mathsf{H}}\left(\mathbf{B}_{(i,k)} - z\mathbf{I}_N\right)^{-1}\mathbf{H}_i\mathbf{w}_{ik} \xrightarrow{\text{a.s.}} 0\ .$$

Noticing from (30) that $d_i = \frac{e_i}{\bar{c}_i - e_i\bar{e}_i}$, we have proved the convergence for $z < z_0$. The Vitali theorem then ensures that the convergence holds true for all $z \in \mathbb{C} \setminus \mathbb{R}^+$. This is however only valid to this point for $\limsup_N c_i < 1$ for all i.

To extend the result to the case where $c_i = 1$ for some i, we proceed as in the proof of Theorem 7. Take $n < N$ and let $c = n/N$. We first have that, for some $k \leq n$ (this does not restrict the generality up to a change in the ordering of the eigenvalues),

$$\begin{aligned}
&\left|\mathbf{w}_{ik}^{\mathsf{H}}\mathbf{H}_i^{\mathsf{H}}\left(\mathbf{B}_{(i,k)} - z\mathbf{I}_N\right)^{-1}\mathbf{H}_i\mathbf{w}_{ik} - \mathbf{w}_{ik}^{\mathsf{H}}\mathbf{H}_i^{\mathsf{H}}\left(\mathbf{B}_{(i,k)}^{(n)} - z\mathbf{I}_N\right)^{-1}\mathbf{H}_i\mathbf{w}_{ik}\right| \\
&= \left|\mathbf{w}_{ik}^{\mathsf{H}}\mathbf{H}_i^{\mathsf{H}}\left(\mathbf{B}_{(i,k)} - z\mathbf{I}_N\right)^{-1}\sum_{j=1}^{K}\sum_{l_j=n+1}^{N_i} p_{j,l_j}\mathbf{H}_j\mathbf{w}_{jl_j}\mathbf{w}_{jl_j}^{\mathsf{H}}\mathbf{H}_j^{\mathsf{H}}\left(\mathbf{B}_{(i,k)}^{(n)} - z\mathbf{I}_N\right)^{-1}\mathbf{H}_i\mathbf{w}_{ik}\right| \\
&\leq (1-c)\bar{c}_+\frac{KR^2P}{|z|^2}
\end{aligned}$$

with $\mathbf{B}_{(i,k)}^{(n)}$ the matrix $\mathbf{B}_{(i,k)}$ with entries p_{il}, $l > n$, set to 0.

Note now that d_i introduced above is well defined if $c_i = 1$ for some i. Denoting $d_i^{(n)}$ the term d_i with $\mathbf{B}_N$ replaced by $\mathbf{B}_N^{(n)}$, we have shown previously that d_i is a continuous mapping of e_i and $d_i^{(n)}$ is a continuous mapping of $e_i^{(n)}$. Also, from (41) and (42), for some $z_1 < z_0$, we also have that

$$\sup_i \left|e_i - e_i^{(n)}\right| \leq (1-c)C_3$$

for some further constant C_3. This implies by continuity that

$$\sup_i \left|d_i - d_i^{(n)}\right| \leq (1-c)C_4$$

for another constant C_4.

As a consequence, for some realization of $\mathbf{B}_1, \mathbf{B}_2, \ldots$ for which

$$\mathbf{w}_{ik}^{\mathsf{H}}\mathbf{H}_i^{\mathsf{H}}\left(\mathbf{B}_{(i,k)}^{(n)} - z\mathbf{I}_N\right)^{-1}\mathbf{H}_i\mathbf{w}_{ik} - d_i^{(n)} \to 0$$

we have

$$\begin{aligned}
&\sup_i \left|d_i - \mathbf{w}_{ik}^{\mathsf{H}}\mathbf{H}_i^{\mathsf{H}}\left(\mathbf{B}_{(i,k)} - z\mathbf{I}_N\right)^{-1}\mathbf{H}_i\mathbf{w}_{ik}\right| \\
\leq\ &\sup_i \Bigg[\left|d_i - d_i^{(n)}\right| + \left|d_i^{(n)} - \mathbf{w}_{ik}^{\mathsf{H}}\mathbf{H}_i^{\mathsf{H}}\left(\mathbf{B}_{(i,k)}^{(n)} - z\mathbf{I}_N\right)^{-1}\mathbf{H}_i\mathbf{w}_{ik}\right| \\
&\quad + \left|\mathbf{w}_{ik}^{\mathsf{H}}\mathbf{H}_i^{\mathsf{H}}\left(\mathbf{B}_{(i,k)}^{(n)} - z\mathbf{I}_N\right)^{-1}\mathbf{H}_i\mathbf{w}_{ik} - \mathbf{w}_{ik}^{\mathsf{H}}\mathbf{H}_i^{\mathsf{H}}\left(\mathbf{B}_{(i,k)} - z\mathbf{I}_N\right)^{-1}\mathbf{H}_i\mathbf{w}_{ik}\right|\Bigg]
\end{aligned}$$

whose superior limit is less than $(1-c)\left(\frac{KR^2P}{|z|^2}+C_4\right)$. For this realization, and for some $\varepsilon>0$, we can therefore choose c such that

$$\limsup_n \max_i \left| d_i - \mathbf{w}_{ik}^{\mathsf{H}}\mathbf{H}_i^{\mathsf{H}}\left(\mathbf{B}_{(i,k)}-z\mathbf{I}_N\right)^{-1}\mathbf{H}_i\mathbf{w}_{ik}\right| < \varepsilon.$$

Since ε is arbitrary, we finally have

$$\max_i \left| d_i - \mathbf{w}_{ik}^{\mathsf{H}}\mathbf{H}_i^{\mathsf{H}}\left(\mathbf{B}_{(i,k)}-z\mathbf{I}_N\right)^{-1}\mathbf{H}_i\mathbf{w}_{ik}\right| \to 0.$$

The realization of $\mathbf{B}_1,\mathbf{B}_2,\ldots$ being taken from a set of probability one, we finally have, for all i and k, and for $z<z_1$,

$$d_i - \mathbf{w}_{ik}^{\mathsf{H}}\mathbf{H}_i^{\mathsf{H}}\left(\mathbf{B}_{(i,k)}-z\mathbf{I}_N\right)^{-1}\mathbf{H}_i\mathbf{w}_{ik} \xrightarrow{\text{a.s.}} 0. \tag{44}$$

Again, we complete the proof of the almost sure convergence by invoking the Vitali theorem to ensure that this holds for all $z\in\mathbb{C}\setminus\mathbb{R}^+$.

Since the quantities d_i and $\mathbf{w}_{ik}^{\mathsf{H}}\mathbf{H}_i^{\mathsf{H}}\left(\mathbf{B}_{(i,k)}-z\mathbf{I}_N\right)^{-1}\mathbf{H}_i\mathbf{w}_{ik}$ are uniformly bounded for all N (a result that holds surely since we assumed the $\mathbf{H}_i$ deterministic), the dominated convergence theorem also ensures that the convergence holds in the first mean.

In order to prove Corollary 1 in the almost sure form, we simply invoke the continuous mapping theorem [39, Theorem 2.3] for the function $\phi : x \mapsto \frac{1}{N}\sum_{k=1}^{K}\sum_{i=1}^{n_k}\log(1+p_{ik}x)$ on the convergence (44). The convergence in the mean sense is obtained using the boundedness of d_i and $\mathbf{w}_{ik}^{\mathsf{H}}\mathbf{H}_i^{\mathsf{H}}\left(\mathbf{B}_{(i,k)}-z\mathbf{I}_N\right)^{-1}\mathbf{H}_i\mathbf{w}_{ik}$ uniformly on N and hence the boundedness of their image by ϕ. The dominated convergence theorem then gives the result.

APPENDIX D
PROOF OF THEOREM 2

It was shown in (32) that, for any fixed $b_k(\sigma^2)\geq 0$, the following equation in $\bar{b}_k(\sigma^2)$:

$$\bar{b}_k(\sigma^2) = \frac{1}{N}\mathrm{tr}\mathbf{P}_k\Big(b_k(\sigma^2)\mathbf{P}_k + \left[\bar{c}_k - b_k(\sigma^2)\bar{b}_k(\sigma^2)\right]\mathbf{I}_{n_k}\Big)^{-1}$$

has a unique solution, satisfying $0\leq \bar{b}_k(\sigma^2) < c_k\bar{c}_k/b_k(\sigma^2)$. Thus, $\bar{b}_k(\sigma^2)$ is uniquely determined by $b_k(\sigma^2)$. Consider now the following functions for $k\in\{1,\ldots,K\}$ and $\sigma^2>0$:

$$h_k(x_1,\ldots,x_K)\mapsto \frac{1}{N}\sum_{j=1}^{N_k}\frac{\zeta_{kj}(\sigma^2)}{1+\bar{b}_k\zeta_{kj}(\sigma^2)}$$

where $\bar{b}_k\in[0,c_k\bar{c}_k/x_k)$ and $\zeta_{kj}(\sigma^2)\geq 0$ are the unique solutions to the following fixed-point equations:

$$\bar{b}_k = \frac{1}{N}\mathrm{tr}\mathbf{P}_k\Big(x_k\mathbf{P}_k + \left[\bar{c}_k - x_k\bar{b}_k\right]\mathbf{I}_{n_k}\Big)^{-1} \tag{45}$$

$$\zeta_{kj}(\sigma^2) = \frac{1}{N}\mathrm{tr}\mathbf{R}_{kj}\left(\frac{1}{N}\sum_{k=1}^{K}\sum_{j=1}^{N_k}\frac{\bar{b}_k\mathbf{R}_{k,j}}{1+\bar{b}_k\zeta_{kj}(\sigma^2)}+\sigma^2\mathbf{I}_N\right)^{-1}. \tag{46}$$

Similar to the proof of Theorem 1, it is now sufficient to prove that the K-variate function $\mathbf{h} : (x_1, \ldots, x_K) \mapsto (h_1, \ldots, h_K)$ is a standard function and to apply Theorem 8 to conclude on the existence and uniqueness of a solution to $x_k = h_k(x_1, \ldots, x_K)$ for all k. The associated fixed-point algorithm follows the recursive equations

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

for $t \geq 0$ and for any set of initial values $x_1^{(0)}, \ldots, x_K^{(0)} > 0$, which then converge, as $t \to \infty$, to the fixed-point.

Showing positivity is straightforward: For $\sigma^2 > 0$, we have $\zeta_{kj}(\sigma^2) > 0$ by Theorem 9 in Appendix G and $\bar{b}_k \geq 0$ by its definition. Thus, $h_k(x_1, \ldots, x_K) > 0$ for all $x_1, \ldots, x_K > 0$.

To prove monotonicity of $h_k(x_1, \ldots, x_K)$, we first recall the following result from (31). Let $x_k > x'_k$, and consider $\bar{b}_k$ and $\bar{b}'_k$ the corresponding solutions to (45). Then,

$$\text{(i)} \;\; \bar{b}_k < \bar{b}'_k \qquad\qquad \text{(ii)} \;\; x_k \bar{b}_k > x'_k \bar{b}'_k. \tag{47}$$

We now prove a further result. Let $\sigma^2 > 0$ and assume $\bar{b}_k > \bar{b}'_k$. Consider $\zeta_{kj}(\sigma^2)$ and $\zeta'_{kj}(\sigma^2)$ as the unique solutions to (46) for $\bar{b}_k$ and $\bar{b}'_k$, respectively. Then,

$$\text{(i)} \;\; \zeta_{kj}(\sigma^2) \leq \zeta'_{kj}(\sigma^2) \qquad\qquad \text{(ii)} \;\; \bar{b}_k \zeta_{kj}(\sigma^2) > \bar{b}'_k \zeta'_{kj}(\sigma^2). \tag{48}$$

Proof: The proof is based on the consideration of an extended version of the random matrix model assumed in Theorem 9. Let us consider the following random matrices $\mathbf{H}_k^L \in \mathbb{C}^{LN \times LN_k}$, given as

$$\mathbf{H}_k^L = \frac{1}{\sqrt{LN}} \left[\left(\mathbf{R}_{k1}^L\right)^{\frac{1}{2}} \mathbf{Z}_{k1}^L, \ldots, \left(\mathbf{R}_{kN_k}^L\right)^{\frac{1}{2}} \mathbf{Z}_{kN_k}^L \right]$$

where $\mathbf{R}_{kj}^L = \mathrm{diag}(\mathbf{R}_{kj}, \ldots, \mathbf{R}_{kj}) \in \mathbb{C}^{LN \times LN}$ are block-diagonal matrices consisting of L copies of the matrix $\mathbf{R}_{kj}$ and $\mathbf{Z}_{kj}^L \in \mathbb{C}^{LN \times L}$ are random matrices composed of i.i.d. entries with zero mean, unit variance and finite moment of order $4 + \epsilon$, for some $\epsilon > 0$. We define the following matrices which will be of repeated use:

$$\tilde{\mathbf{B}}^L = \sum_{k=1}^{K} \bar{b}_k \mathbf{H}_k^L \left(\mathbf{H}_k^L\right)^{\mathsf{H}}, \qquad \tilde{\mathbf{B}'}^L = \bar{b}'_k \mathbf{H}_k^L \left(\mathbf{H}_k^L\right)^{\mathsf{H}} + \sum_{l=1, l \neq k}^{K} \bar{b}_l \mathbf{H}_l^L \left(\mathbf{H}_l^L\right)^{\mathsf{H}}$$

$$\mathbf{Q} = \left(\tilde{\mathbf{B}}^L + \sigma^2 \mathbf{I}_{NL}\right)^{-1}, \qquad \mathbf{Q}' = \left(\tilde{\mathbf{B}'}^L + \sigma^2 \mathbf{I}_{NL}\right)^{-1}.$$

One can verify from Theorem 9 that for any fixed $N, N_1, \ldots, N_K$, the following limit holds:

$$\frac{1}{LN} \mathrm{tr} \mathbf{R}_{kj}^L \left(\tilde{\mathbf{B}}^L + \sigma^2 \mathbf{I}_{NK}\right)^{-1} \xrightarrow[L \to \infty]{\text{a.s}} \zeta_{kj}(\sigma^2).$$

Thus, any properties of the random quantities on the left-hand side of the previous equation also hold for the deterministic quantities $\zeta_{kj}(\sigma^2)$. We will exploit this fact for the termination of the proof. The matrices $\tilde{\mathbf{B}}^L$ and $\tilde{\mathbf{B}'}^L$ differ only by $\bar{b}_k$. This assumption will be sufficient for the proof since the case $\bar{b}_l > \bar{b}'_l$ for $l \in \{1, \ldots, K\}$ follows by simple iteration of the case $\bar{b}_l = \bar{b}'_l$ for $l \neq k$ and $\bar{b}_k > \bar{b}'_k$.

To prove (i), it is now sufficient to show that, for any L,

$$\frac{1}{N} \mathrm{tr} \mathbf{R}_{k,j}^L \left(\mathbf{Q} - \mathbf{Q}'\right) < 0.$$

By Lemma 6, this is equivalent to proving $(\mathbf{Q})^{-1} - (\mathbf{Q}')^{-1} \succ 0$, which is straightforward since

$$(\mathbf{Q})^{-1} - (\mathbf{Q}')^{-1} = \tilde{\mathbf{B}}^L - \tilde{\mathbf{B}'}^L = (\bar{b}_k - \bar{b}'_k)\mathbf{H}_k^L \left(\mathbf{H}_k^L\right)^{\mathsf{H}} \succ 0.$$

Thus,

$$\frac{1}{NL}\mathrm{tr}\mathbf{R}_{k,j}^L \left(\mathbf{Q} - \mathbf{Q}'\right) \xrightarrow[L\to\infty]{\text{a.s}} \zeta_{kj}(\sigma^2) - \zeta'_{kj}(\sigma^2) \leq 0$$

since $\zeta_{kj}(\sigma^2)$ and $\zeta'_{kj}(\sigma^2)$ do not depend on L.

For (ii), we need to show that

$$\bar{b}_k \frac{1}{LN}\mathrm{tr}\mathbf{R}_{kj}^L\mathbf{Q} - \bar{b}'_k \frac{1}{LN}\mathrm{tr}\mathbf{R}_{kj}^L\mathbf{Q}' > 0.$$

Similarly to the previous part of the proof, it is sufficient to show that $\left(\bar{b}_k\mathbf{Q}\right)^{-1} - \left(\bar{b}'_k\mathbf{Q}'\right)^{-1} \prec 0$. Hence,

$$\begin{aligned}
\left(\bar{b}_k\mathbf{Q}\right)^{-1} - \left(\bar{b}'_k\mathbf{Q}'\right)^{-1} &= \frac{1}{\bar{b}_k}\left(\tilde{\mathbf{B}}^L + \sigma^2\mathbf{I}_{NL}\right) - \frac{1}{\bar{b}'_k}\left(\tilde{\mathbf{B}'}^L + \sigma^2\mathbf{I}_{NL}\right) \\
&= \sigma^2\left(\frac{1}{\bar{b}_k} - \frac{1}{\bar{b}'_k}\right)\mathbf{I}_{NL} + \left(\frac{1}{\bar{b}_k} - \frac{1}{\bar{b}'_k}\right)\sum_{l=1,l\neq k}^{K} \bar{b}_l\mathbf{H}_l^L\left(\mathbf{H}_l^L\right)^{\mathsf{H}} \\
&\prec 0
\end{aligned}$$

since $\sigma^2 > 0$, $\bar{b}_k > \bar{b}'_k$ and $\bar{b}_l \geq 0$ for all l. ■

Consider now $(x_1, \ldots, x_K)$ and $(x'_1, \ldots, x'_K)$, such that $x_k > x'_k$ $\forall k$, and denote by $(\bar{b}_1, \ldots, \bar{b}_K)$ and $(\bar{b}'_1, \ldots, \bar{b}'_K)$ the corresponding solutions to (45). Denote by $\zeta_{kj}(\sigma^2)$ and $\zeta'_{kj}(\sigma^2)$ the unique solutions to (46) for $(\bar{b}_1, \ldots, \bar{b}_K)$ and $(\bar{b}'_1, \ldots, \bar{b}'_K)$, respectively. It follows from (47), that $\bar{b}_k < \bar{b}'_k$ $\forall k$. Equation (48) now implies that $\zeta_{kj}(\sigma^2) \geq \zeta'_{kj}(\sigma^2)$ and $\bar{b}_k\zeta_{kj}(\sigma^2) < \bar{b}'_k\zeta'_{kj}(\sigma^2)$. Combining these results yields

$$h_k(x_1, \ldots, x_K) = \frac{1}{N}\sum_{j=1}^{N_k}\frac{\zeta_{kj}(\sigma^2)}{1 + \bar{b}_k\zeta_{kj}(\sigma^2)} > \frac{1}{N}\sum_{j=1}^{N_k}\frac{\zeta'_{kj}(\sigma^2)}{1 + \bar{b}'_k\zeta'_{kj}(\sigma^2)} = h_k(x'_1, \ldots, x'_K)$$

which proves monotonicity.

To prove scalability, let $\alpha > 1$, and consider the following difference:

$$\begin{aligned}
\alpha h_k(x_1, \ldots, x_K) - h_k(\alpha x_1, \ldots, \alpha x_K) &= \frac{1}{N}\sum_{j=1}^{N_k}\frac{\alpha\zeta_{kj}(\sigma^2)}{1 + \bar{b}_k\zeta_{kj}(\sigma^2)} - \frac{\zeta_{kj}^{(\alpha)}(\sigma^2)}{1 + \bar{b}_k^{(\alpha)}\zeta_{kj}^{(\alpha)}(\sigma^2)} \\
&= \frac{1}{N}\sum_{i=1}^{N_k}\frac{\left[\alpha\zeta_{kj}(\sigma^2) - \zeta_{kj}^{(\alpha)}(\sigma^2)\right] + \zeta_{kj}(\sigma^2)\zeta_{kj}^{(\alpha)}(\sigma^2)\left[\alpha\bar{b}_k^{(\alpha)} - \bar{b}_k\right]}{\left[1 + \bar{b}_k\zeta_{kj}(\sigma^2)\right]\left[1 + \bar{b}_k^{(\alpha)}\zeta_{kj}^{(\alpha)}(\sigma^2)\right]}
\end{aligned}$$

where we have denoted by $\bar{b}_k^{(\alpha)}$ the solution to (45) with x_k replaced by αx_k and by $\zeta_{kj}^{(\alpha)}(\sigma^2)$ the solution to (46) for $\bar{b}_k^{(\alpha)}$. We have from (47)-(i) that $\bar{b}_k^{(\alpha)} < \bar{b}_k$ and from (47)-(ii) that

$$\alpha x_k\bar{b}_k^{(\alpha)} > x_k\bar{b}_k \iff \alpha\bar{b}_k^{(\alpha)} - \bar{b}_k > 0. \tag{49}$$

It remains now to show that also $\alpha\zeta_{kj}(\sigma^2) - \zeta_{kj}^{(\alpha)}(\sigma^2) > 0$. To this end, consider the following difference:

$$\alpha\zeta_{kj}(\sigma^2) - \zeta_{kj}^{(\alpha)}(\sigma^2) = \frac{1}{N}\mathrm{tr}\mathbf{R}_{kj}\left(\alpha\mathbf{T}(\sigma^2) - \mathbf{T}^{(\alpha)}(\sigma^2)\right)$$

where

$$\mathbf{T}(\sigma^2) = \left(\frac{1}{N}\sum_{k=1}^{K}\sum_{j=1}^{N_k}\frac{\bar{b}_k\mathbf{R}_{k,j}}{1+\bar{b}_k\zeta_{kj}(\sigma^2)} + \sigma^2\mathbf{I}_N\right)^{-1}$$
$$\mathbf{T}^{(\alpha)}(\sigma^2) = \left(\frac{1}{N}\sum_{k=1}^{K}\sum_{j=1}^{N_k}\frac{\bar{b}_k^{(\alpha)}\mathbf{R}_{k,j}}{1+\bar{b}_k^{(\alpha)}\zeta_{kj}^{(\alpha)}(\sigma^2)} + \sigma^2\mathbf{I}_N\right)^{-1}.$$

By Lemma 6, it is now sufficient to show that $\left(\mathbf{T}^{(\alpha)}(z)\right)^{-1} \succ (\alpha\mathbf{T}(z))^{-1}$. Write therefore

$$\begin{aligned}&\left(\mathbf{T}^{(\alpha)}(\sigma^2)\right)^{-1} - \left(\alpha\mathbf{T}(\sigma^2)\right)^{-1}\\ =\ &\sigma^2\left(1-\frac{1}{\alpha}\right)\mathbf{I}_N + \frac{1}{N}\sum_{k=1}^{K}\sum_{j=1}^{N_k}\frac{\left[\alpha\bar{b}_k^{(\alpha)}-\bar{b}_k\right]+\bar{b}_k^{(\alpha)}\bar{b}_k\left[\alpha\zeta_{kj}(\sigma^2)-\zeta_{kj}^{(\alpha)}(\sigma^2)\right]}{\alpha\left[1+\bar{b}_k\zeta_{kj}(\sigma^2)\right]\left[1+\bar{b}_k^{(\alpha)}\zeta_{kj}^{(\alpha)}(\sigma^2)\right]}\mathbf{R}_{kj}.\end{aligned}$$

The first summand is positive definite since $\sigma^2 > 0$ and $\alpha > 1$. All other terms are also positive definite since $\alpha\bar{b}_k^{(\alpha)} - \bar{b}_k > 0$ from (49) and $\alpha\bar{b}_k^{(\alpha)}\bar{b}_k\zeta_{kj}(\sigma^2) > \bar{b}_k\bar{b}_k^{(\alpha)}\zeta_{kj}^{(\alpha)}(\sigma^2)$, since $\alpha\bar{b}_k^{(\alpha)} > \bar{b}_k$ and $\bar{b}_k\zeta_{kj}(\sigma^2) > \bar{b}_k^{(\alpha)}\zeta_{kj}^{(\alpha)}(\sigma^2)$ by (48)-(ii) and (47)-(i). Since the sum of positive definite matrices is also positive definite, we have $\alpha\zeta_{kj}(\sigma^2) - \zeta_{kj}^{(\alpha)}(\sigma^2) > 0$. This terminates the proof of scalability.

Thus, we have shown $\mathbf{h} : (x_1,\ldots,x_K) \mapsto (h_1,\ldots,h_K)$ to be a standard function. Moreover, from the series convergence in Theorem 1 and Theorem 9, we have the following algorithm to compute $\bar{b}_k$ and $\zeta_{kj}(\sigma^2)$:

$$\bar{b}_k = \lim_{t\to\infty}\bar{b}_k^{(t)}, \qquad \zeta_{kj}(\sigma^2) = \lim_{t\to\infty}\zeta_{kj}^{(t)}(\sigma^2)$$

where

$$\bar{b}_k^{(t)} = \frac{1}{N}\mathrm{tr}\mathbf{P}_k\Big(x_k\mathbf{P}_k + \left[\bar{c}_k - x_k\bar{b}_k^{(t-1)}\right]\mathbf{I}_{n_k}\Big)^{-1}$$
$$\zeta_{kj}^{(t)}(\sigma^2) = \frac{1}{N}\mathrm{tr}\mathbf{R}_{kj}\left(\frac{1}{N}\sum_{k=1}^{K}\sum_{j=1}^{N_k}\frac{\bar{b}_k\mathbf{R}_{k,j}}{1+\bar{b}_k\zeta_{kj}^{(t-1)}(\sigma^2)} + \sigma^2\mathbf{I}_N\right)^{-1}$$

and $\bar{b}_k^{(0)}$ can take any value in $[0, c_k\bar{c}_k/x_k)$ and $\zeta_{kj}^{(0)}(\sigma^2) = 1/\sigma^2$ for all k, j.

Appendix E
Proof of Theorem 4

We begin by proving the following result:

$$\max_k |\bar{a}_k(\sigma^2) - \bar{b}_k(\sigma^2)| \xrightarrow{\text{a.s.}} 0 \tag{50}$$
$$\max_k |a_k(\sigma^2) - b_k(\sigma^2)| \xrightarrow{\text{a.s.}} 0 \tag{51}$$

where $\bar{a}_k(\sigma^2)$, $a_k(\sigma^2)$ are defined in Theorem 1 and $\bar{b}_k(\sigma^2)$, $b_k(\sigma^2)$ are defined in Theorem 2, assuming that the matrices $\mathbf{H}_k$ are random and modeled as described in (2). For notational simplicity, we will drop from now on the dependence on σ^2. From standard lemmas of matrix analysis, we have

$$\begin{aligned}
a_k &= \frac{1}{N}\mathrm{tr}\mathbf{H}_k\mathbf{H}_k^{\mathsf{H}}\left(\sum_{i=1}^{K}\bar{a}_i\mathbf{H}_i\mathbf{H}_i^{\mathsf{H}}+\sigma^2\mathbf{I}_N\right)^{-1} \\
&= \frac{1}{N}\sum_{j=1}^{N_k}\mathbf{h}_{kj}^{\mathsf{H}}\left(\sum_{i=1}^{K}\bar{a}_i\mathbf{H}_i\mathbf{H}_i^{\mathsf{H}}+\sigma^2\mathbf{I}_N\right)^{-1}\mathbf{h}_{kj} \\
&= \frac{1}{N}\sum_{j=1}^{N_k}\frac{\mathbf{h}_{kj}^{\mathsf{H}}\left(\sum_{i=1}^{K}\bar{a}_i\mathbf{H}_i\mathbf{H}_i^{\mathsf{H}}-\bar{a}_k\mathbf{h}_{kj}\mathbf{h}_{kj}^{\mathsf{H}}+\sigma^2\mathbf{I}_N\right)^{-1}\mathbf{h}_{kj}}{1+\bar{a}_k\mathbf{h}_{kj}^{\mathsf{H}}\left(\sum_{i=1}^{K}\bar{a}_i\mathbf{H}_i\mathbf{H}_i^{\mathsf{H}}-\bar{a}_k\mathbf{h}_{kj}\mathbf{h}_{kj}^{\mathsf{H}}+\sigma^2\mathbf{I}_N\right)^{-1}\mathbf{h}_{kj}}
\end{aligned}$$

where the last step follows from Lemma 3. If $\bar{a}_i$ were not dependent on $\mathbf{h}_{kj}$, we could now simply proceed by applying Lemma 4 to the individual quadratic forms, i.e.:

$$\mathbf{h}_{kj}^{\mathsf{H}}\left(\sum_{i=1}^{K}\bar{a}_i\mathbf{H}_i\mathbf{H}_i^{\mathsf{H}}-\bar{a}_k\mathbf{h}_{kj}\mathbf{h}_{kj}^{\mathsf{H}}+\sigma^2\mathbf{I}_N\right)^{-1}\mathbf{h}_{kj}\asymp\frac{1}{N}\mathrm{tr}\mathbf{R}_{kj}\left(\sum_{i=1}^{K}\bar{a}_i\mathbf{H}_i\mathbf{H}_i^{\mathsf{H}}-\bar{a}_k\mathbf{h}_{kj}\mathbf{h}_{kj}^{\mathsf{H}}+\sigma^2\mathbf{I}_N\right)^{-1}$$

where, in the following, for $\{a_N\}$ and $\{b_N\}$ 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 $N \to \infty$.

However, in order to show that this step is correct, in a similar manner as in the proof of Theorem 7, we need the following intermediate arguments. Define $\bar{a}_{i,kj}$ and $a_{i,kj}$ as the unique solutions to the following fixed-point equations:

$$\begin{aligned}
a_{i,kj} &= \frac{1}{N}\mathrm{tr}\mathbf{H}_{i,kj}\mathbf{H}_{i,kj}^{\mathsf{H}}\left(\sum_{l=1}^{K}\bar{a}_{l,kj}\mathbf{H}_{l,kj}\mathbf{H}_{l,kj}^{\mathsf{H}}+\sigma^2\mathbf{I}_N\right)^{-1} \\
\bar{a}_{i,kj} &= \frac{1}{N}\mathrm{tr}\mathbf{P}_i\left(a_{i,kj}\mathbf{P}_i+[\bar{c}_k-a_{i,kj}\bar{a}_{i,kj}\mathbf{I}_{n_i}]\right)^{-1}
\end{aligned}$$

for $i \in \{1,\ldots,K\}$, where

$$\mathbf{H}_{i,kj} = \begin{cases}\mathbf{H}_i, & k \neq i \\ \left(\mathbf{h}_{k1}\cdots\mathbf{h}_{kj-1}\mathbf{h}_{kj+1}\cdots\mathbf{h}_{kN_i}\right), & k = i\end{cases}.$$

Thus, $\bar{a}_{i,kj}$ and $a_{i,kj}$ are independent of $\mathbf{h}_{kj}$. Following similar steps as in the proof of Theorem 7 (Step 3), one can show that for $i \in \{1,\ldots,K\}$ and all k,j,

$$a_{i,kj} - a_i \xrightarrow{\text{a.s.}} 0, \quad \bar{a}_{i,kj} - \bar{a}_i \xrightarrow{\text{a.s.}} 0. \tag{52}$$

Thus, we have

$$
\begin{aligned}
&\frac{1}{N}\sum_{j=1}^{N_k}\frac{\mathbf{h}_{kj}^{\mathsf{H}}\left(\sum_{i=1}^{K}\bar{a}_i\mathbf{H}_i\mathbf{H}_i^{\mathsf{H}}-\bar{a}_k\mathbf{h}_{kj}\mathbf{h}_{kj}^{\mathsf{H}}+\sigma^2\mathbf{I}_N\right)^{-1}\mathbf{h}_{kj}}{1+\bar{a}_k\mathbf{h}_{kj}^{\mathsf{H}}\left(\sum_{i=1}^{K}\bar{a}_i\mathbf{H}_i\mathbf{H}_i^{\mathsf{H}}-\bar{a}_k\mathbf{h}_{kj}\mathbf{h}_{kj}^{\mathsf{H}}+\sigma^2\mathbf{I}_N\right)^{-1}\mathbf{h}_{kj}}\\
\overset{(a)}{\asymp}\;&\frac{1}{N}\sum_{j=1}^{N_k}\frac{\mathbf{h}_{kj}^{\mathsf{H}}\left(\sum_{i=1}^{K}\bar{a}_{i,kj}\mathbf{H}_i\mathbf{H}_i^{\mathsf{H}}-\bar{a}_{k,kj}\mathbf{h}_{kj}\mathbf{h}_{kj}^{\mathsf{H}}+\sigma^2\mathbf{I}_N\right)^{-1}\mathbf{h}_{kj}}{1+\bar{a}_k\mathbf{h}_{kj}^{\mathsf{H}}\left(\sum_{i=1}^{K}\bar{a}_{i,kj}\mathbf{H}_i\mathbf{H}_i^{\mathsf{H}}-\bar{a}_{k,kj}\mathbf{h}_{kj}\mathbf{h}_{kj}^{\mathsf{H}}+\sigma^2\mathbf{I}_N\right)^{-1}\mathbf{h}_{kj}}\\
\overset{(b)}{\asymp}\;&\frac{1}{N}\sum_{j=1}^{N_k}\frac{\frac{1}{N}\mathrm{tr}\mathbf{R}_{kj}\left(\sum_{i=1}^{K}\bar{a}_{i,kj}\mathbf{H}_i\mathbf{H}_i^{\mathsf{H}}-\bar{a}_{k,kj}\mathbf{h}_{kj}\mathbf{h}_{kj}^{\mathsf{H}}+\sigma^2\mathbf{I}_N\right)^{-1}}{1+\bar{a}_k\frac{1}{N}\mathrm{tr}\mathbf{R}_{kj}\left(\sum_{i=1}^{K}\bar{a}_{i,kj}\mathbf{H}_i\mathbf{H}_i^{\mathsf{H}}-\bar{a}_{k,kj}\mathbf{h}_{kj}\mathbf{h}_{kj}^{\mathsf{H}}+\sigma^2\mathbf{I}_N\right)^{-1}}\\
\overset{(c)}{\asymp}\;&\frac{1}{N}\sum_{j=1}^{N_k}\frac{\frac{1}{N}\mathrm{tr}\mathbf{R}_{kj}\left(\sum_{i=1}^{K}\bar{a}_i\mathbf{H}_i\mathbf{H}_i^{\mathsf{H}}+\sigma^2\mathbf{I}_N\right)^{-1}}{1+\bar{a}_k\frac{1}{N}\mathrm{tr}\mathbf{R}_{kj}\left(\sum_{i=1}^{K}\bar{a}_i\mathbf{H}_i\mathbf{H}_i^{\mathsf{H}}+\sigma^2\mathbf{I}_N\right)^{-1}}\\
\overset{(d)}{\asymp}\;&\frac{1}{N}\sum_{j=1}^{N_k}\frac{\frac{1}{N}\mathrm{tr}\mathbf{R}_{kj}\bar{\mathbf{T}}}{1+\bar{a}_k\frac{1}{N}\mathrm{tr}\mathbf{R}_{kj}\bar{\mathbf{T}}}
\end{aligned}
\tag{53}
$$

where (a) follows from (52), (b) follows from Lemma 4 and Lemma 8, (c) is again due to (52) and Lemma 7, and (d) follows from an application of Theorem 9, where we have defined

$$
\bar{\mathbf{T}}=\left(\frac{1}{N}\sum_{k=1}^{K}\sum_{j=1}^{N_k}\frac{\bar{a}_k\mathbf{R}_{kj}}{1+\bar{a}_k\frac{1}{N}\mathrm{tr}\mathbf{R}_{kj}\bar{\mathbf{T}}}+\sigma^2\mathbf{I}_N\right)^{-1}.
$$

Note again that Theorem 9 cannot be directly applied here since the quantities $\bar{a}_i$ depend on the matrices $\mathbf{H}_i$. However, it is immediate to show that the result extends in this case, by replacing $\bar{a}_i$ by $\bar{a}_{i,kj}$ at each necessary step of the proof.

Hence, we can write

$$
a_k=\frac{1}{N}\mathrm{tr}\mathbf{H}_k\mathbf{H}_k^{\mathsf{H}}\left(\sum_{i=1}^{K}\bar{a}_i\mathbf{H}_i\mathbf{H}_i^{\mathsf{H}}+\sigma^2\mathbf{I}_N\right)^{-1}=\frac{1}{N}\sum_{j=1}^{N_k}\frac{\frac{1}{N}\mathrm{tr}\mathbf{R}_{kj}\bar{\mathbf{T}}}{1+\bar{a}_k\frac{1}{N}\mathrm{tr}\mathbf{R}_{kj}\bar{\mathbf{T}}}+\epsilon_{N,k}
$$

for some sequences of reals $\epsilon_{N,k}$, satisfying $\epsilon_{N,k}\to 0$.

Recall now the following definitions for $k=1,\ldots,K$:

$$
\begin{aligned}
a_k&=\frac{1}{N}\sum_{j=1}^{N_k}\frac{\frac{1}{N}\mathrm{tr}\mathbf{R}_{kj}\bar{\mathbf{T}}}{1+\bar{a}_k\frac{1}{N}\mathrm{tr}\mathbf{R}_{kj}\bar{\mathbf{T}}}+\epsilon_{N,k}\\
b_k&=\frac{1}{N}\sum_{j=1}^{N_k}\frac{\frac{1}{N}\mathrm{tr}\mathbf{R}_{kj}\mathbf{T}}{1+\bar{b}_k\frac{1}{N}\mathrm{tr}\mathbf{R}_{kj}\mathbf{T}}\\
\bar{a}_k&=\frac{1}{N}\sum_{j=1}^{n_k}\frac{p_{kj}}{\bar{c}_k-a_k\bar{a}_k+a_kp_{kj}},\qquad 0\le\bar{a}_k<c_k\bar{c}_k/a_k\\
\bar{b}_k&=\frac{1}{N}\sum_{j=1}^{n_k}\frac{p_{kj}}{\bar{c}_k-b_k\bar{b}_k+b_kp_{kj}},\qquad 0\le\bar{b}_k<c_k\bar{c}_k/b_k
\end{aligned}
$$

where

$$\bar{\mathbf{T}} = \left(\frac{1}{N} \sum_{k=1}^{K} \sum_{j=1}^{N_k} \frac{\bar{a}_k \mathbf{R}_{kj}}{1 + \bar{f}_{N,k} \frac{1}{N} \mathrm{tr} \mathbf{R}_{kj} \bar{\mathbf{T}}} + \sigma^2 \mathbf{I}_N \right)^{-1}$$

$$\mathbf{T} = \left(\frac{1}{N} \sum_{k=1}^{K} \sum_{j=1}^{N_k} \frac{\bar{b}_k \mathbf{R}_{kj}}{1 + \bar{b}_k \frac{1}{N} \mathrm{tr} \mathbf{R}_{kj} \mathbf{T}} + \sigma^2 \mathbf{I}_N \right)^{-1}.$$

A. Case: $\limsup c_k < 1$

We will first assume that $\limsup c_k < 1$ for all k. The case $\limsup c_k = 1$ will be treated separately in the subsequent section. Denote $P = \max_k \{\limsup \|\mathbf{P}_k\|\}$, $R = \max_m \{\limsup \|\tilde{\mathbf{R}}_m\|\}$, $c_+ = \max_k \{\limsup c_k\}$ and $\bar{c}_- = \min_k \{\liminf \bar{c}_k\}$, $\bar{c}_+ = \max_k \{\limsup \bar{c}_k\}$. Since we are interested in the asymptotic limit $N \to \infty$, we assume from the beginning that N is sufficiently large, so that the following inequalities hold for all k:

$$c_k \le c_+, \quad \bar{c}_- \le \bar{c}_k \le \bar{c}_+, \quad \|\mathbf{P}_k\| \le P, \quad \|\mathbf{R}_{kj}\| \le R.$$

We then have the following properties:

$$\bar{a}_k \le \frac{P}{(1 - c_+)\bar{c}_-}, \quad \bar{b}_k \le \frac{P}{(1 - c_+)\bar{c}_-}, \quad b_k \bar{b}_k < c_+ \bar{c}_+, \quad a_k \bar{a}_k < c_+ \bar{c}_+. \tag{54}$$

For notational simplicity, we define the following quantities:

$$\xi = \max_k |a_k - b_k|, \qquad \bar{\xi} = \max_k |\bar{a}_k - \bar{b}_k|.$$

We will show in the sequel that $\xi \xrightarrow{\text{a.s.}} 0$ and $\bar{\xi} \xrightarrow{\text{a.s.}} 0$ as $N \to \infty$.

Consider first the following difference:

$$\begin{aligned}
\sup_{k,j} \left| \frac{1}{N} \mathrm{tr} \mathbf{R}_{kj} \left(\mathbf{T} - \bar{\mathbf{T}} \right) \right| &= \sup_{k,j} \left| \frac{1}{N} \mathrm{tr} \mathbf{R}_{kj} \mathbf{T} \left(\frac{1}{N} \sum_{l=1}^{K} \sum_{m=1}^{N_l} \frac{\bar{a}_l \mathbf{R}_{lm}}{1 + \bar{a}_l \frac{1}{N} \mathrm{tr} \mathbf{R}_{lm} \bar{\mathbf{T}}} - \frac{\bar{b}_l \mathbf{R}_{lm}}{1 + \bar{b}_l \frac{1}{N} \mathrm{tr} \mathbf{R}_{lm} \bar{\mathbf{T}}} \right) \bar{\mathbf{T}} \right| \\
&= \sup_{k,j} \left| \frac{1}{N} \sum_{l=1}^{K} \sum_{m=1}^{N_l} \frac{\bar{a}_l - \bar{b}_l + \bar{a}_l \bar{b}_l \left(\frac{1}{N} \mathrm{tr} \mathbf{R}_{lm} \mathbf{T} - \frac{1}{N} \mathrm{tr} \mathbf{R}_{lm} \bar{\mathbf{T}} \right)}{\left(1 + \bar{a}_l \frac{1}{N} \mathrm{tr} \mathbf{R}_{lm} \bar{\mathbf{T}} \right) \left(1 + \bar{b}_l \frac{1}{N} \mathrm{tr} \mathbf{R}_{lm} \bar{\mathbf{T}} \right)} \frac{1}{N} \mathrm{tr} \mathbf{R}_{kj} \bar{\mathbf{T}} \mathbf{R}_{lm} \mathbf{T} \right| \\
&\le \frac{R^2}{\sigma^4} K \max_k \bar{c}_k \left[\max_k |\bar{a}_k - \bar{b}_k| + \max_k |\bar{a}_k \bar{b}_k| \sup_{k,j} \left| \frac{1}{N} \mathrm{tr} \mathbf{R}_{kj} \left(\mathbf{T} - \bar{\mathbf{T}} \right) \right| \right] \\
&\le \frac{R^2}{\sigma^4} K \bar{c}_+ \left[\bar{\xi} + \frac{P^2}{(1 - c_+)^2 \bar{c}_-^2} \sup_{k,j} \left| \frac{1}{N} \mathrm{tr} \mathbf{R}_{kj} \left(\mathbf{T} - \bar{\mathbf{T}} \right) \right| \right]
\end{aligned}$$

where the first equality follows from Lemma 2. Rearranging the terms yields:

$$\sup_{k,j} \left| \frac{1}{N} \mathrm{tr} \mathbf{R}_{kj} \left(\mathbf{T} - \bar{\mathbf{T}} \right) \right| \le \frac{P^2 K \bar{c}_+}{\sigma^4 - \frac{R^2 P^2}{(1-c_+)^2 \bar{c}_-^2}} \ \bar{\xi} \tag{55}$$

for $\sigma^2 > \frac{RP}{(1-c_+)\bar{c}_-}$.

Consider now the term $\xi = \max_k |a_k - b_k|$:

$$\begin{aligned}
\xi &= \max_k \left| \frac{1}{N} \sum_{j=1}^{N_k} \frac{\frac{1}{N}\mathrm{tr}\mathbf{R}_{kj}\left(\bar{\mathbf{T}} - \mathbf{T}\right) + (\bar{b}_k - \bar{a}_k)\frac{1}{N}\mathrm{tr}\mathbf{R}_{kj}\frac{1}{N}\mathrm{tr}\mathbf{R}_{kj}\bar{\mathbf{T}}}{\left(1 + \bar{a}_k \frac{1}{N}\mathrm{tr}\mathbf{R}_{kj}\bar{\mathbf{T}}\right)\left(1 + \bar{b}_k \frac{1}{N}\mathrm{tr}\mathbf{R}_{kj}\mathbf{T}\right)} + \epsilon_{N,k} \right| \\
&\leq \bar{c}_+ \sup_{kj} \left| \frac{1}{N}\mathrm{tr}\mathbf{R}_{kj}\left(\mathbf{T} - \bar{\mathbf{T}}\right) \right| + \bar{c}_+ \frac{R^2}{\sigma^4} \max_k |\bar{a}_k - \bar{b}_k| + \max_k |\epsilon_{N,k}| \\
&\leq \frac{P^2 K \bar{c}_+^2}{\sigma^4 - \frac{R^2P^2}{(1-c_+)^2\bar{c}_-^2}} \xi + \frac{\bar{c}_+ R^2}{\sigma^4}\bar{\xi} + \max_k |\epsilon_{N,k}| \\
&= \left[\frac{P^2 K \bar{c}_+^2}{\sigma^4 - \frac{R^2P^2}{(1-c_+)^2\bar{c}_-^2}} + \frac{\bar{c}_+ R^2}{\sigma^4} \right] \bar{\xi} + \max_k |\epsilon_{N,k}| \qquad (56)
\end{aligned}$$

where the last inequality follows from (55). Similarly, we have for $\bar{\xi} = \max_k |\bar{a}_k - \bar{b}_k|$:

$$\begin{aligned}
\bar{\xi} &= \max_k \left| \frac{1}{N} \sum_{j=1}^{n_k} p_{kj} \frac{a_k \bar{a}_k - b_k \bar{b}_k + p_{kj}(b_k - a_k)}{(\bar{c}_k - a_k\bar{a}_k + a_k p_{kj})(\bar{c}_k - b_k\bar{b}_k + b_k p_{kj})} \right| \\
&\leq \frac{1}{N} \sum_{j=1}^{n_k} \frac{p_{kj}^2 \max_k |a_k - b_k|}{(1-c_+)^2\bar{c}_-^2} + p_{kj} \frac{\max_k \left[\bar{a}_k |a_k - b_k]\,| + \max_k \left[b_k|\bar{a}_k - \bar{b}_k|\right]}{(1-c_+)^2\bar{c}_-^2} \\
&\leq \frac{P^2}{(1-c_+)^2\bar{c}_-^2} \left(1 + \frac{1}{(1-c_+)\bar{c}_-}\right) \xi + \frac{PR\bar{c}_+}{\sigma^2(1-c_+)^2\bar{c}_-^2} \bar{\xi}.
\end{aligned}$$

Thus, for $\sigma^2 \geq \max\left\{ \frac{2PR\bar{c}_+}{(1-c_+)^2\bar{c}_-^2}, \frac{RP}{(1-c_+)\bar{c}_-} \right\}$, we have

$$\bar{\xi} \leq \frac{2P^2}{(1-c_+)^2\bar{c}_-^2} \left(1 + \frac{1}{(1-c_+)\bar{c}_-}\right) \xi. \qquad (57)$$

Replacing (57) in (56) leads to

$$\xi \leq \left[\frac{P^2 K \bar{c}_+^2}{\sigma^4 - \frac{R^2P^2}{(1-c_+)^2\bar{c}_-^2}} + \frac{\bar{c}_+ R^2}{\sigma^4} \right] \frac{2P^2}{(1-c_+)^2\bar{c}_-^2} \left(1 + \frac{1}{(1-c_+)\bar{c}_-}\right) \xi + \max_k |\epsilon_{N,k}|.$$

For σ^2 sufficiently large, we therefore have

$$0 \leq \xi \leq C\epsilon_{N,k} \to 0$$

for some $C > 0$. This implies that $\xi \xrightarrow{\text{a.s.}} 0$ and, by (57), that $\bar{\xi} \xrightarrow{\text{a.s.}} 0$. Since $b_k, \bar{a}_k, \bar{b}_k$ are uniformly bounded for σ^2 spanning any closed subset of $\mathbb{R}_+$ and a_k is almost surely uniformly bounded for σ^2 spanning any closed subset of $\mathbb{R}_+$, we have from Vitali's convergence theorem [37] that the almost sure convergence holds true for all $\sigma^2 \in \mathbb{R}_+$. This terminates the proof for $c_k < 1$.

B. *Case:* $\limsup c_k = 1$

It was shown in Appendix A (reminding that $\bar{a}_k(\sigma^2) = \bar{e}_k(-\sigma^2)$) that the following refined inequalities hold for $c_k = 1$:

$$\bar{a}_k \leq P, \quad \bar{b}_k \leq P.$$

Using these inequalities instead of (54) in the proof for the case $c_k < 1$, one can show that $\xi \xrightarrow{\text{a.s.}} 0$ and $\bar{\xi} \xrightarrow{\text{a.s.}} 0$ as $N \to \infty$.

C. Convergence of the mutual information

Consider now the first term of $V_N(\sigma^2)$ in Theorem 10. Due to the convergence of $\bar{a}_k - \bar{b}_k \xrightarrow{\text{a.s.}} 0$ and the almost sure boundedness of the $\mathbf{H}_k\mathbf{H}_k^{\mathsf{H}}$ matrices, $\|\sum_{k=1}^K (\bar{a}_k - \bar{b}_k)\mathbf{H}_k\mathbf{H}_k^{\mathsf{H}}\| \xrightarrow{\text{a.s.}} 0$, and we have immediately that

$$\frac{1}{N}\log\det\left(\mathbf{I}_N + \frac{1}{\sigma^2}\sum_{k=1}^K \bar{a}_k\mathbf{H}_k\mathbf{H}_k^{\mathsf{H}}\right) - \frac{1}{N}\log\det\left(\mathbf{I}_N + \frac{1}{\sigma^2}\sum_{k=1}^K \bar{b}_k\mathbf{H}_k\mathbf{H}_k^{\mathsf{H}}\right) \xrightarrow{\text{a.s.}} 0.$$

Applying Corollary 3 to the second term yields

$$\frac{1}{N}\log\det\left(\mathbf{I}_N + \frac{1}{\sigma^2}\sum_{k=1}^K \bar{b}_k\mathbf{H}_k\mathbf{H}_k^{\mathsf{H}}\right) - \bar{V}_N(\sigma^2) \xrightarrow{\text{a.s.}} 0. \tag{58}$$

Consider now $\bar{I}_N^{(a)}(\sigma^2)$ and $\bar{I}_N^{(b)}(\sigma^2)$ as defined in Theorems 3 and 4. It follows from (50), (51) and (58), that

$$\bar{I}_N^{(a)}(\sigma^2) - \bar{I}_N^{(b)}(\sigma^2) \xrightarrow{\text{a.s.}} 0.$$

This implies also that

$$I_N^{(b)}(\sigma^2) - \bar{I}_N^{(b)}(\sigma^2) \xrightarrow{\text{a.s.}} 0. \tag{59}$$

To prove convergence in the mean, we can no longer use the fact that $I_N^{(b)}(\sigma^2)$ is bounded for all N as in Appendix B, which is now untrue. Instead, we will use the same arguments as in [5]. Denote

$$m_N^{(b)}(z) = \frac{1}{N}\operatorname{tr}(\mathbf{B}_N - z\mathbf{I}_N)^{-1}, \quad \bar{m}_N^{(b)}(z) = \frac{1}{N}\operatorname{tr}\left(\frac{1}{N}\sum_{k=1}^K\sum_{j=1}^{N_k}\frac{\bar{b}_k(-z)\mathbf{R}_{k,j}}{1+b_k(-z)\zeta_{kj}(-z)} - z\mathbf{I}_N\right)^{-1}$$

where $m_N^{(b)}(z)$ is the Stieltjes transform of $\mathbf{B}_N$. It is easy to see that

$$\mathbb{E}I_N^{(b)}(\sigma^2) - \bar{I}_N^{(b)}(\sigma^2) = \int_{\sigma^2}^{\infty}\left(\left[\frac{1}{\omega} - \mathbb{E}m_N^{(b)}(-\omega)\right] - \left[\frac{1}{\omega} - \bar{m}_N^{(b)}(-\omega)\right]\right)d\omega.$$

We now apply the argument from [5, pp. 923] which shows that

$$\begin{aligned}&\left|\int_{\sigma^2}^{\infty}\left(\left[\frac{1}{\omega} - \mathbb{E}m_N^{(b)}(-\omega)\right] - \left[\frac{1}{\omega} - \bar{m}_N^{(b)}(-\omega)\right]\right)d\omega\right| \\ \leq\ &\int_{\sigma^2}^{\infty}\frac{1}{\omega^2}\left(\left|\mathbb{E}\int_0^{\infty} t dF_N^{(b)}(t)\right| + \left|\frac{1}{N}\operatorname{tr}\left(\frac{1}{N}\sum_{k=1}^K\sum_{j=1}^{N_k}\frac{\bar{b}_k(\omega)\mathbf{R}_{k,j}}{1+b_k(\omega)\zeta_{kj}(\omega)}\right)\right|\right)d\omega\end{aligned}$$

the right-hand side of which exists for all N and is uniformly bounded by $\frac{2}{\sigma^2}(KPR)$. Since $m_N^{(b)}(-\omega) - \bar{m}_N^{(b)}(-\omega) \xrightarrow{\text{a.s.}} 0$ (as a consequence of the convergence $\bar{a}_k - \bar{b}_k \xrightarrow{\text{a.s.}} 0$), the boundedness of $m_N^{(b)}(-\omega)$ then ensures (by dominated convergence) that $\mathbb{E}m_N^{(b)}(-\omega) - \bar{m}_N^{(b)}(-\omega) \to 0$. Since the integrand tends to zero and is summable independently of N, the dominated convergence theorem now ensures that

$$\mathbb{E}I_N^{(b)}(\sigma^2) - \bar{I}_N^{(b)}(\sigma^2) \to 0.$$

We now turn to the proof of Proposition 1.

Proof of Proposition 1: By the chain rule of differentiation, we first have

$$\frac{d\bar{I}_N^{(b)}(\sigma^2)}{dp_{kj}} = \frac{\partial\bar{I}_N^{(b)}(\sigma^2)}{\partial p_{kj}} + \sum_{i=1}^K\left[\frac{\partial\bar{I}_N^{(b)}(\sigma^2)}{\partial b_i}\frac{\partial b_i}{\partial p_{kj}} + \frac{\partial\bar{I}_N^{(b)}(\sigma^2)}{\partial\bar{b}_i}\frac{\bar{b}_i}{\partial p_{kj}}\right].$$

Consider now the partial derivative:

$$\begin{aligned}
\frac{\partial \bar{I}_N^{(b)}(\sigma^2)}{\partial b_i} &= \frac{\partial \bar{V}_N(\sigma^2)}{\partial b_i} + \frac{1}{N}\mathrm{tr}\mathbf{P}_i\left(b_i\mathbf{P}_i + \left[\bar{c}_i - b_i\bar{b}_i\right]\mathbf{I}_{n_i}\right)^{-1} - \bar{b}_i\frac{1}{N}\sum_{j=1}^{n_i}\frac{1}{\bar{c}_i - b_i\bar{b}_i + b_ip_{ij}} - \bar{b}_i\frac{(1-c_i)\bar{c}_i}{\bar{c}_i - b_i\bar{b}_i}\\
&= \frac{\partial \bar{V}_N(\sigma^2)}{\partial b_i} + \bar{b}_i\left(1 - \frac{(1-c_i)\bar{c}_i}{\bar{c}_i - b_i\bar{b}_i} - \frac{1}{N}\sum_{j=1}^{n_i}\frac{1}{\bar{c}_i - b_i\bar{b}_i + b_ip_{ij}}\right)\\
&= \frac{\partial \bar{V}_N(\sigma^2)}{\partial b_i}
\end{aligned}$$

where the last equality follows from

$$\begin{aligned}
0 &= \bar{c}_i - \frac{1}{N}\sum_{j=1}^{n_i}\frac{\bar{c}_i - b_i\bar{b}_i + b_ip_{ij}}{\bar{c}_i - b_i\bar{b}_i + b_ip_{ij}} - \frac{1}{N}\sum_{j=1}^{N_i-n_i}\frac{\bar{c}_i - b_i\bar{b}_i}{\bar{c}_i - b_i\bar{b}_i}\\
&= \bar{c}_i - (\bar{c}_i - b_i\bar{b}_i)\frac{1}{N}\sum_{j=1}^{n_i}\frac{1}{\bar{c}_i - b_i\bar{b}_i + b_ip_{ij}} + b_i\frac{1}{N}\sum_{j=1}^{n_i}\frac{p_{ij}}{\bar{c}_i - b_i\bar{b}_i + b_ip_{ij}} - (\bar{c}_i - b_i\bar{b}_i)\frac{\frac{N_i}{N} - \frac{n_i}{N}}{\bar{c}_i - b_i\bar{b}_i}\\
&= (\bar{c}_i - b_i\bar{b}_i)\left(1 - \frac{(1-c_i)\bar{c}_i}{\bar{c}_i - b_i\bar{b}_i} - \frac{1}{N}\sum_{j=1}^{n_i}\frac{1}{\bar{c}_i - b_i\bar{b}_i + b_ip_{ij}}\right)
\end{aligned}$$

and $\bar{c}_i \geq \bar{c}_ic_i > b_i\bar{b}_i$ by definition.

Similarly, we have

$$\begin{aligned}
\frac{\partial \bar{I}_N^{(b)}(\sigma^2)}{\partial \bar{b}_i} &= \frac{\partial \bar{V}_N(\sigma^2)}{\partial \bar{b}_i} - b_i\left(\frac{(1-c_i)\bar{c}_i}{\bar{c}_i - b_i\bar{b}_i} + \frac{1}{N}\sum_{j=1}^{n_i}\frac{1}{\bar{c}_i - b_i\bar{b}_i + b_ip_{ij}}\right)\\
&= \frac{\partial \bar{V}_N(\sigma^2)}{\partial \bar{b}_i} - b_i.
\end{aligned}$$

It remains now to calculate the partial derivatives $\frac{\partial \bar{V}_N(\sigma^2)}{\partial b_i}$ and $\frac{\partial \bar{V}_N(\sigma^2)}{\partial \bar{b}_i}$. To this end, note that

$$\begin{aligned}
1 = \frac{1}{N}\mathrm{tr}\mathbf{T}\mathbf{T}^{-1} &= \sigma^2\frac{1}{N}\mathrm{tr}\mathbf{T} + \sum_{k=1}^{K}\bar{b}_k\frac{1}{N}\sum_{j=1}^{N_k}\frac{\frac{1}{N}\mathrm{tr}\mathbf{R}_{k,j}\mathbf{T}}{1 + \bar{b}_k\frac{1}{N}\mathrm{tr}\mathbf{R}_{k,j}\mathbf{T}}\\
&= \sigma^2\frac{1}{N}\mathrm{tr}\mathbf{T} + \sum_{k=1}^{K}\bar{b}_kb_k.
\end{aligned}$$

Replacing $\sum_{k=1}^{K}\bar{b}_kb_k$ in (8) by $\frac{1}{N}\mathrm{tr}\mathbf{T}\mathbf{T}^{-1} - \sigma^2\frac{1}{N}\mathrm{tr}\mathbf{T}$ yields

$$\bar{V}_N(\sigma^2) = -\frac{1}{N}\log\det\left(\sigma^2\mathbf{T}\right) - \frac{1}{N}\mathrm{tr}\mathbf{T}\mathbf{T}^{-1} + \sigma^2\frac{1}{N}\mathrm{tr}\mathbf{T} + \frac{1}{N}\sum_{k=1}^{K}\sum_{j=1}^{N_k}\log\left(1 + \bar{b}_k\frac{1}{N}\mathrm{tr}\mathbf{R}_{k,j}\mathbf{T}\right).$$

Taking the derivative with respect to $\bar{b}_i$ and denoting $\mathbf{T}' = \frac{\partial \mathbf{T}}{\partial \bar{b}_i}$ leads to

$$\begin{aligned}
\frac{\partial \bar{V}_N(\sigma^2)}{\partial \bar{b}_i} &= -\frac{1}{N}\mathrm{tr}\mathbf{T}^{-1}\mathbf{T}' + \sigma^2\frac{1}{N}\mathrm{tr}\mathbf{T}' + \frac{1}{N}\sum_{k=1}^{K}\sum_{j=1}^{N_k}\frac{\bar{b}_k\frac{1}{N}\mathrm{tr}\mathbf{R}_{k,j}\mathbf{T}'}{1 + \bar{b}_k\frac{1}{N}\mathrm{tr}\mathbf{R}_{k,j}\mathbf{T}} + \frac{1}{N}\sum_{j=1}^{N_i}\frac{\frac{1}{N}\mathrm{tr}\mathbf{R}_{i,j}\mathbf{T}}{1 + \bar{b}_i\frac{1}{N}\mathrm{tr}\mathbf{R}_{i,j}\mathbf{T}}\\
&= -\sigma^2\frac{1}{N}\mathrm{tr}\mathbf{T}' - \frac{1}{N}\sum_{k=1}^{K}\sum_{j=1}^{N_k}\frac{\bar{b}_k\frac{1}{N}\mathrm{tr}\mathbf{R}_{k,j}\mathbf{T}'}{1 + \bar{b}_k\frac{1}{N}\mathrm{tr}\mathbf{R}_{k,j}\mathbf{T}} + \sigma^2\frac{1}{N}\mathrm{tr}\mathbf{T}' + \frac{1}{N}\sum_{k=1}^{K}\sum_{j=1}^{N_k}\frac{\bar{b}_k\frac{1}{N}\mathrm{tr}\mathbf{R}_{k,j}\mathbf{T}'}{1 + \bar{b}_k\frac{1}{N}\mathrm{tr}\mathbf{R}_{k,j}\mathbf{T}} + b_i\\
&= b_i.
\end{aligned}$$

This implies that $\frac{\partial \bar{I}_N^{(b)}(\sigma^2)}{\partial \bar{b}_i} = 0$. We similarly have

$$\frac{\partial \bar{V}_N(\sigma^2)}{\partial b_i} = 0$$

and hence $\frac{\partial \bar{I}_N^{(b)}(\sigma^2)}{\partial b_i} = 0$. Putting the last results together yields

$$\frac{d\bar{I}_N^{(b)}(\sigma^2)}{dp_{kj}} = \frac{\partial \bar{I}_N^{(b)}(\sigma^2)}{\partial p_{kj}} = \frac{b_k}{N\left(\bar{c}_k - b_k\bar{b}_k + b_k p_{kj}\right)}. \tag{60}$$

We can calculate the second derivative in a similar manner:

$$\frac{d^2\bar{I}_N^{(b)}(\sigma^2)}{dp_{kj}^2} = \frac{\partial^2 \bar{I}_N^{(b)}(\sigma^2)}{\partial p_{kj}^2} = -\frac{b_k^2}{N\left(\bar{c}_k - b_k\bar{b}_k + b_k p_{kj}\right)^2} \leq 0$$

since $b_k \geq 0$. Thus, $\bar{I}_N^{(b)}(\sigma^2)$ is a concave function in p_{kj} for all k, j. It is straightforward to verify that also $I_N^{(b)}(\sigma^2)$ is concave in all p_{kj}.

Consider now the Lagrangian functions related to the power constraints (I) and (II):

$$\mathcal{L}(\lambda, \lambda_1, \ldots, \lambda_K, p_{11}, \ldots, p_{KN_K}) = \begin{cases} \bar{I}_N^{(b)}(\sigma^2) - \sum_{k=1}^K \lambda_k \left(\frac{1}{n_k}\sum_{j=1}^{n_k} p_{kj} - P_k\right) & (I) \\ \bar{I}_N^{(b)}(\sigma^2) - \lambda\left(\sum_{k=1}^K \frac{1}{n_k}\sum_{j=1}^{n_k} p_{kj} - P\right) & (II) \end{cases} \tag{61}$$

We have from (60)

$$\frac{\partial \mathcal{L}}{\partial p_{kj}} = \begin{cases} \frac{b_k}{N\left(\bar{c}_k - b_k\bar{b}_k + b_k p_{kj}\right)} - \frac{\lambda_k}{n_k} & (I) \\ \frac{b_k}{N\left(\bar{c}_k - b_k\bar{b}_k + b_k p_{kj}\right)} - \frac{\lambda}{n_k} & (II). \end{cases} \tag{62}$$

Solving for the Karush-Kuhn-Tucker conditions [40] for both cases yields the desired result.

Take now the optimal solutions $\bar{\mathbf{P}}^\star \triangleq (\bar{\mathbf{P}}_1^\star, \ldots, \bar{\mathbf{P}}_K^\star)$ and $\mathbf{P}^\star \triangleq (\mathbf{P}_1^\star, \ldots, \mathbf{P}_K^\star)$ and consider the following difference:

$$I_N^{(b)}(\mathbf{P}^\star) - I_N^{(b)}(\bar{\mathbf{P}}^\star) = \left[I_N^{(b)}(\mathbf{P}^\star) - \bar{I}_N^{(b)}(\mathbf{P}^\star)\right] + \left[\bar{I}_N^{(b)}(\mathbf{P}^\star) - \bar{I}_N^{(b)}(\bar{\mathbf{P}}^\star)\right] + \left[\bar{I}_N^{(b)}(\bar{\mathbf{P}}^\star) - I_N^{(b)}(\bar{\mathbf{P}}^\star)\right]$$

where we used $I_N^{(b)}(\mathbf{P}^\star)$ and $\bar{I}_N^{(b)}(\bar{\mathbf{P}}^\star)$ to denote $I_N^{(b)}(\sigma^2)$ and $\bar{I}_N^{(b)}(\sigma^2)$ evaluated for the matrices $(\bar{\mathbf{P}}_1^\star, \ldots, \bar{\mathbf{P}}_K^\star)$ and $(\mathbf{P}_1^\star, \ldots, \mathbf{P}_K^\star)$, respectively. Assuming that $\max_K \limsup_N \|\mathbf{P}_k^\star\| \leq \infty$, we have from Theorem 4

$$I_N^{(b)}(\mathbf{P}^\star) - \bar{I}_N^{(b)}(\mathbf{P}^\star) \xrightarrow{\text{a.s.}} 0$$
$$\bar{I}_N^{(b)}(\bar{\mathbf{P}}^\star) - I_N^{(b)}(\bar{\mathbf{P}}^\star) \xrightarrow{\text{a.s.}} 0.$$

Since $I_N^{(b)}(\mathbf{P}^\star) - I_N^{(b)}(\bar{\mathbf{P}}^\star) \geq 0$ and $\bar{I}_N^{(b)}(\mathbf{P}^\star) - \bar{I}_N^{(b)}(\bar{\mathbf{P}}^\star) \leq 0$, we can conclude that

$$I_N^{(b)}(\mathbf{P}^\star) - I_N^{(b)}(\bar{\mathbf{P}}^\star) \xrightarrow{\text{a.s.}} 0.$$

It remains now to show that the matrices $\mathbf{P}_k^\star$ satisfy indeed $\max_K \limsup_N \|\mathbf{P}_k^\star\| \leq \infty$. Consider therefore the following expression:

$$\mathbb{E}I_N^{(b)} = \mathbb{E}\frac{1}{N}\log\det\left(\mathbf{I}_N + \frac{1}{\sigma^2}\mathbf{B}_{N(k,j)} + \frac{p_{kj}}{\sigma^2}\mathbf{H}_k\mathbf{w}_{kj}\mathbf{w}_{kj}^{\mathsf{H}}\mathbf{H}_k^{\mathsf{H}}\right)$$

which is clearly strictly concave in p_{kj} for all k, j. The corresponding derivative with respect to p_{kj} reads

$$\begin{aligned}\frac{\partial \mathbb{E} I_N^{(b)}}{\partial p_{kj}} &= \mathbb{E}\frac{1}{N}\mathrm{tr}\left(\mathbf{I}_N + \frac{1}{\sigma^2}\mathbf{B}_N\right)^{-1}\frac{1}{\sigma^2}\mathbf{H}_k\mathbf{w}_{kj}\mathbf{w}_{kj}^{\mathsf{H}}\mathbf{H}_k^{\mathsf{H}} \\ &= \mathbb{E}\frac{1}{\sigma^2 N}\mathbf{w}_{kj}^{\mathsf{H}}\mathbf{H}_k^{\mathsf{H}}\left(\mathbf{I}_N + \frac{1}{\sigma^2}\mathbf{B}_N\right)^{-1}\mathbf{H}_k\mathbf{w}_{kj}.\end{aligned}$$

Similar to (62), the derivative of the Lagrangian to the optimization problem (9) is given as

$$\frac{\partial \mathcal{L}}{\partial p_{kj}} = \begin{cases}\mathbb{E}\frac{1}{\sigma^2 N}\mathbf{w}_{kj}^{\mathsf{H}}\mathbf{H}_k^{\mathsf{H}}\left(\mathbf{I}_N + \frac{1}{\sigma^2}\mathbf{B}_N\right)^{-1}\mathbf{H}_k\mathbf{w}_{kj} - \frac{\lambda_k}{n_k} & (I) \\ \mathbb{E}\frac{1}{\sigma^2 N}\mathbf{w}_{kj}^{\mathsf{H}}\mathbf{H}_k^{\mathsf{H}}\left(\mathbf{I}_N + \frac{1}{\sigma^2}\mathbf{B}_N\right)^{-1}\mathbf{H}_k\mathbf{w}_{kj} - \frac{\lambda}{n_k} & (II)\end{cases}.$$

Consider now constraint (I). At the optimal point, we need to have $\frac{\partial \mathcal{L}}{\partial p_{kj}} = 0$, and therefore

$$\mathbb{E}\frac{1}{\sigma^2 N}\mathbf{w}_{kj}^{\mathsf{H}}\mathbf{H}_k^{\mathsf{H}}\left(\mathbf{I}_N + \frac{1}{\sigma^2}\mathbf{B}_N\right)^{-1}\mathbf{H}_k\mathbf{w}_{kj} = \frac{\lambda_k}{n_k}.$$

Since the right-hand side is independent of j, it follows that $\mathbf{P}_k^\star = p_k\mathbf{I}_{n_k}$ where p_k is a parameter to be optimized. Since $\frac{1}{n_k}\mathrm{tr}\mathbf{P}_k^\star = p_k \leq P_k$, we have $\max_K \limsup_N \|\mathbf{P}_k^\star\| = p_k \leq P_k < \infty$. The same arguments hold for the sum-power constraint (II). ■

Proof of Theorem 6: The proof follows directly from (50), (51), and Theorem 5. ■

Proof of Corollary 2: The almost sure convergence follows directly from Theorem 6 and the continuous mapping theorem [39, Theorem 2.3]. For the convergence in mean, note first that, $R_N^{(b)}(\sigma^2) \leq I_N^{(b)}(\sigma^2)$ and also $\bar{R}_N^{(b)}(\sigma^2) \leq \bar{I}_N^{(b)}(\sigma^2)$. Thus, for N sufficiently large, we have

$$\left|R_N^{(b)}(\sigma^2) - \bar{R}_N^{(b)}(\sigma^2)\right| \leq \limsup_N I_N^{(b)}(\sigma^2) + \bar{I}_N^{(b)}(\sigma^2) \triangleq \phi(\sigma^2).$$

Since $\mathbb{E}\phi(\sigma^2) < \infty$, it follows from the dominated convergence theorem that

$$\mathbb{E}R_N^{(b)}(\sigma^2) - \bar{R}_N^{(b)}(\sigma^2) \xrightarrow[N\to\infty]{} 0.$$

■

APPENDIX F
FUNDAMENTAL LEMMAS

Lemma 1 (Defining properties of Stieltjes transforms, Theorem 3.2 in [11]): If m is a function analytic on $\mathbb{C}^+$ such that $m(z) \in \mathbb{C}^+$ if $z \in \mathbb{C}^+$ and

$$\lim_{y\to\infty} -iy\ m(iy) = 1 \tag{63}$$

then m is the Stieltjes transform of a distribution function F given by

$$F(b) - F(a) = \lim_{y\to 0}\frac{1}{\pi}\int_a^b \mathrm{Im}[m(x+iy)]dx.$$

If, moreover, $zm(z) \in \mathbb{C}^+$ for $z \in \mathbb{C}^+$, then $F(0^-) = 0$, in which case m has an analytic continuation on $\mathbb{C} \setminus \mathbb{R}^+$.

Lemma 2 (Resolvent identity): For invertible matrices $\mathbf{A}$ and $\mathbf{B}$, we have the following identity:

$$\mathbf{A}^{-1} - \mathbf{B}^{-1} \;=\; \mathbf{A}^{-1}(\mathbf{B}-\mathbf{A})\mathbf{B}^{-1} \;.$$

Lemma 3 (A matrix inversion lemma, Equation (2.2) in [41]): 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

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

Lemma 4 (Trace lemma [30, Lemma 2.7]): Let $\mathbf{A}_1, \mathbf{A}_2, \ldots$, with $\mathbf{A}_N \in \mathbb{C}^{N\times N}$, be a sequence of matrices with uniformly bounded spectral norm and let $\mathbf{x}_N = \in \mathbb{C}^N$ be random vectors of i.i.d. entries with zero mean, variance $1/N$ and eighth order moment of order $\mathcal{O}(1/N^4)$, independent of $\mathbf{A}_N$. Then, as $N \to \infty$,

$$\mathbf{x}_N^{\mathsf{H}}\mathbf{A}_N\mathbf{x}_N - \frac{1}{N}\mathrm{tr}\mathbf{A}_N \xrightarrow{\text{a.s.}} 0. \tag{64}$$

Lemma 5 (Trace lemma for isometric matrices, [8]): Let $\mathbf{W}$ be $n < N$ columns of an $N \times N$ Haar matrix and suppose $\mathbf{w}$ is a column of $\mathbf{W}$. Let $\mathbf{B}_N$ be an $N \times N$ random matrix, which is a function of all columns of $\mathbf{W}$ except $\mathbf{w}$ and $B = \sup_N \|\mathbf{B}_N\| < \infty$, then

$$\mathbb{E}\left[\left|\mathbf{w}^{\mathsf{H}}\mathbf{B}_N\mathbf{w} - \frac{1}{N-n}\,\mathrm{tr}(\mathbf{\Pi}\mathbf{B}_N)\right|^4\right] \le \frac{C}{N^2},$$

where $\mathbf{\Pi} = \mathbf{I}_N - \mathbf{W}\mathbf{W}^{\mathsf{H}} + \mathbf{w}\mathbf{w}^{\mathsf{H}}$ and C is a constant which depends only on B and $\frac{n}{N}$.

Lemma 6 (Trace inequality): Let $\mathbf{A}, \mathbf{B}, \mathbf{R} \in \mathbb{C}^{N\times N}$, where $\mathbf{A}$ and $\mathbf{B}$ are nonnegative-definite, satisfying $\mathbf{B} \succ \mathbf{A}$, and $\mathbf{R}$ is nonnegative-definite. Then

$$\mathrm{tr}\mathbf{R}\left(\mathbf{A}^{-1} - \mathbf{B}^{-1}\right) > 0. \tag{65}$$

Proof: Note that $\mathbf{B} \succ \mathbf{A}$ implies by [42, Corollary 7.7.4] $\mathbf{B}^{-1} \prec \mathbf{A}^{-1}$. Thus, for any vector $\mathbf{x} \in \mathbb{C}^N$,

$$\mathbf{x}^{\mathsf{H}}\left(\mathbf{A}^{-1} - \mathbf{B}^{-1}\right)\mathbf{x} > 0. \tag{66}$$

Consider now the eigenvalue decomposition of the matrix $\mathbf{R} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^{\mathsf{H}}$, where $\mathbf{U} = [\mathbf{u}_1, \ldots, \mathbf{u}_N]$ and $\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_N)$. Since $\lambda_i \ge 0\ \forall i$, we have

$$\mathrm{tr}\mathbf{R}\left(\mathbf{A}^{-1} - \mathbf{B}^{-1}\right) = \sum_{i=1}^{N}\lambda_i\mathbf{u}_i^{\mathsf{H}}\left(\mathbf{A}^{-1} - \mathbf{B}^{-1}\right)\mathbf{u}_i > 0. \tag{67}$$

■

Lemma 7 (Rank-1 perturbation lemma [41]): 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,

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

Lemma 8: [15, Lemma 1] Denote a_N, $\overline{a}_N$, b_N and $\overline{b}_N$ four infinite sequences of complex random variables indexed by N and assume $a_N \asymp \overline{a}_N$ and $b_N \asymp \overline{b}_N$. If $|a_N|$, $|\overline{b}_N|$ and/or $|\overline{a}_N|$,$|b_N|$ are uniformly bounded above over N (almost surely), then $a_N b_N \asymp \overline{a}_N\overline{b}_N$. Similarly, if $|a_N|$, $|\overline{b}_N|^{-1}$ and/or $|\overline{a}_N|$,$|b_N|^{-1}$ are uniformly bounded above over N (almost surely), then $a_N/b_N \asymp \overline{a}_N/\overline{b}_N$.

Lemma 9 (Tonelli theorem [36, Theorem 18.3]): If $(\Omega, \mathcal{F}, P)$ and $(\Omega', \mathcal{F}', P')$ are two probability spaces, then for f an integrable function with respect to the product measure Q on $\mathcal{F} \times \mathcal{F}'$,

$$\int_{\Omega\times\Omega'} f(x,y)Q(d(x,y)) = \int_{\Omega}\left[\int_{\Omega'} f(x,y)P'(dy)\right]P(dx)$$

and

$$\int_{\Omega\times\Omega'} f(x,y)Q(d(x,y)) = \int_{\Omega'}\left[\int_{\Omega} f(x,y)P(dy)\right]P'(dx).$$

APPENDIX G

RELATED RESULTS

Theorem 9 ([1, Theorem 1]): Let $\mathbf{B}_N = \mathbf{X}\mathbf{X}^{\mathsf{H}}$, where $\mathbf{X} \in \mathbb{C}^{N\times n}$ is random. The jth column $\mathbf{x}_j$ of $\mathbf{X}$ is given as $\mathbf{x}_j = \mathbf{R}_j^{\frac{1}{2}}\mathbf{z}_j$, where the entries of $\mathbf{z}_j \in \mathbb{C}^N$ are i.i.d. with zero mean, variance $1/N$ and finite moment of order $4+\epsilon$, for some common $\epsilon > 0$, and $\mathbf{R}_j \in \mathbb{C}^{N\times N}$ are Hermitian nonnegative definite matrices. Let $\mathbf{D}_N \in \mathbb{C}^{N\times N}$ be a deterministic Hermitian matrix. Assume that both $\mathbf{R}_j$ and $\mathbf{D}_N$ have uniformly bounded spectral norms (with respect to N). Then, as $n, N \to \infty$ such that $0 < \liminf N/n \le \limsup N/n < \infty$, the following holds for any $z \in \mathbb{C} \setminus \mathbb{R}_+$:

$$\frac{1}{N}\mathrm{tr}\mathbf{D}_N\left(\mathbf{B}_N - z\mathbf{I}_N\right)^{-1} - \frac{1}{N}\mathrm{tr}\mathbf{D}_N\mathbf{T}_N(z) \xrightarrow{\text{a.s.}} 0$$

where $\mathbf{T}_N(z) \in \mathbb{C}^{N\times N}$ is defined as

$$\mathbf{T}_N(z) = \left(\frac{1}{N}\sum_{j=1}^{n}\frac{\mathbf{R}_j}{1+\delta_j(z)} - z\mathbf{I}_N\right)^{-1}$$

and where $\delta_1(z), \ldots, \delta_n(z)$ are given as the unique solution to the following set of implicit equations:

$$\delta_j(z) = \frac{1}{N}\mathrm{tr}\mathbf{R}_j\left(\frac{1}{N}\sum_{j=1}^{n}\frac{\mathbf{R}_j}{1+\delta_j(z)} - z\mathbf{I}_N\right)^{-1}, \qquad j = 1,\ldots,n \tag{68}$$

such that $(\delta_1(z), \ldots, \delta_n(z)) \in \mathcal{S}^n$. For $z < 0$, $\delta_1(z), \ldots, \delta_{N,n}(z)$ are the unique nonnegative solutions to (68) and can be obtained by a standard fixed-point algorithm with initial values $\delta_j^{(0)}(z) = -1/z$ for $j = 1, \ldots, n$. Moreover, let F_N be the empirical spectral distribution (e.s.d.) of $\mathbf{B}_N$ and denote by $\bar{F}_N$ the distribution function with Stieltjes transform $\frac{1}{N}\mathrm{tr}\mathbf{T}_N(z)$. Then, almost surely,

$$F_N - \bar{F}_N \Rightarrow 0.$$

Theorem 10 ([43]): Under the assumptions of Theorem 9, let $\sigma^2 > 0$ and define $V_N(\sigma^2) = \frac{1}{N}\log\det\left(\mathbf{I}_N + \frac{1}{\sigma^2}\mathbf{B}_N\right)$. Then, as $N, n \to \infty$,

$$\mathbb{E}V_N(\sigma^2) - \bar{V}_N(\sigma^2) \xrightarrow{\text{a.s.}} 0$$

where

$$\bar{V}_N(\sigma^2) = \frac{1}{N}\log\det\left(\mathbf{I}_N + \frac{1}{\sigma^2}\frac{1}{N}\sum_{j=1}^{n}\frac{\mathbf{R}_j}{1+\delta_j}\right) + \frac{1}{N}\sum_{j=1}^{n}\log\left(1+\delta_j\right) - \frac{1}{N}\sum_{j=1}^{n}\frac{\delta_j}{1+\delta_j}$$

and where $\delta_j = \delta_j(-\sigma^2)$ for $j = 1, \ldots, n$ are given by Theorem 9.

Corollary 3: Under the assumptions of Theorem 10, assume additionally that the matrices $\mathbf{R}_j$, $j = 1, \ldots, n$, are drawn from a finite set of Hermitian nonnegative-definite matrices. Then, as $N, n \to \infty$,

$$V_N(\sigma^2) - \bar{V}_N(\sigma^2) \xrightarrow{\text{a.s.}} 0 \tag{69}$$

where $V_N(\sigma^2)$ and $\bar{V}_N(\sigma^2)$ are defined as in Theorem 10.

Proof: It was shown in [44, Proof of Theorem 3] that $\mathbf{B}_N$ has almost surely uniformly bounded spectral norm as $N, n \to \infty$ if the matrices $\mathbf{R}_j$ are drawn from a finite set of matrices. Thus, F_N and $\bar{F}_N$ as defined in Theorem 9 have (almost surely) bounded support. Consider now a set $A \subset \Omega$, Ω generating the matrices $\mathbf{B}_N$, for which $\mathbf{B}_N$ has bounded spectral norm, and a set $B \subset \Omega$ for which $F_N - \bar{F}_N \Rightarrow 0$. Since $P(A) = P(B) = P(A \cap B) = 1$, it follows from [45, Theorem 25.8 (ii)], that, as $N, n \to \infty$

$$\int \log(1 + x^{-1}\lambda) dF_N(\lambda) - \int \log(1 + x^{-1}\lambda) d\overline{F}_N(\lambda) \xrightarrow{\text{a.s.}} 0 \tag{70}$$

which is equivalent to stating that $V_N(x) - \bar{V}_N(x) \xrightarrow{\text{a.s.}} 0$. ■

REFERENCES

[1] 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

[2] J. Hoydis, M. Debbah, and M. Kobayashi, "Asymptotic moments for interference mitigation in correlated fading channels," in *Proc. IEEE International Symposium on Information Theory (ISIT'11)*, Saint Petersburg, Russia, Aug. 2011. [Online]. Available: http://arxiv.org/abs/1104.4911

[3] C. N. Chuah, D. N. C. Tse, J. M. Kahn, and R. A. Valenzuela, "Capacity scaling in MIMO wireless systems under correlated fading," *IEEE Trans. Inf. Theory*, vol. 48, no. 3, pp. 637–650, Mar. 2002.

[4] 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.

[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] ——, "A CLT for information theoretic statistics of Gram random matrices with a given variance profile," *The Annals of Probability*, vol. 18, no. 6, pp. 2071–2130, Dec. 2008.

[7] 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.

[8] M. Debbah, W. Hachem, P. Loubaton, and M. de Courville, "MMSE analysis of certain large isometric random precoded systems," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1293–1311, May 2003.

[9] W. Hachem, "Simple polynomial MMSE receivers for CDMA transmissions on frequency selective channels," *IEEE Trans. Inf. Theory*, pp. 164–172, Jan. 2004.

[10] ——, "An expression for $\int \log(t/\sigma^2 + 1)\mu \boxplus \tilde{\mu}(dt)$," 2008, unpublished.

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

[12] M. J. M. Peacock, I. B. Collings, and M. L. Honig, "Asymptotic analysis of LMMSE multiuser receivers for multi-signature multicarrier CDMA in Rayleigh fading," *IEEE Trans. Commun.*, vol. 52, no. 6, pp. 964–972, Jun. 2004.

[13] ——, "Asymptotic spectral efficiency of multiuser multisignature CDMA in frequency-selective channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 3, pp. 1113–1129, Mar. 2006.

[14] R. Couillet and M. Debbah, "Uplink capacity of self-organizing clustered orthogonal CDMA networks in flat fading channels," in *Proc. IEEE Information Theory Workshop (ITW Fall'09)*, Taormina, Sicily, 2009.

[15] M. J. M. Peacock, I. B. Collings, and M. L. Honig, "Eigenvalue distributions of sums and products of large random matrices via incremental matrix expansions," *IEEE Trans. Inf. Theory*, vol. 54, no. 5, pp. 2123–2138, 2008.

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

[17] J. Dumont, W. Hachem, P. Loubaton, and J. Najim, "On the asymptotic analysis of mutual information of MIMO Rician correlated channels," in *IEEE-EURASIP International Symposium on Control, Communication, and Signal Processing (ISCCSP'06)*, Marrakech, Morocco, Mar. 2006.

[18] R. S. Blum, J. H. Winters, and N. R. Sollenberger, "On the Capacity of Cellular Systems with MIMO," *IEEE Commun. Lett.*, vol. 6, pp. 242–244, Jun. 2002.

[19] R. S. Blum, "MIMO Capacity with Interference," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 5, pp. 793–801, Jun. 2003.

[20] P. Viswanath, D. N. C. Tse, and R. Laroia, "Opportunistic Beamforming Using Dumb Antennas," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1277–1294, Jun. 2002.

[21] D. J. Love and R. W. Heath, "Limited Feedback Unitary Precoding for Spatial Multiplexing Systems," *IEEE Trans. Inf. Theory*, vol. 51, no. 8, pp. 2967 –2976, Aug. 2005.

[22] J. Lee, J. K. Han, and J. Zhang, "MIMO Technologies in 3GPP LTE and LTE-Advanced," *EURASIP Journal on Wireless Communications and Networking*, vol. 2009, May 2009.

[23] K. Huang, J. G. Andrews, and R. W. Heath, "Performance of orthogonal beamforming for SDMA with limited feedback," *IEEE Trans. Veh. Technol.*, vol. 58, no. 1, pp. 152–164, Jan. 2009.

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

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

[26] D. Voiculescu, K. J. Dykema, and A. Nica, "Free random variables," *American Mathematical Society*, 1992.

[27] F. Hiai and D. Petz, *The Semicircle Law, Free Random Variables and Entropy - Mathematical Surveys and Monographs No. 77*. Providence, RI, USA: American Mathematical Society, 2006.

[28] R. Speicher, C. Vargas, and T. Mai, "Free deterministic equivalents, rectangular random matrix models, and operator-valued free probability theory," 2011. [Online]. Available: http://arxiv.org/abs/1110.1237

[29] S. Verdu and S. Shamai, "Spectral Efficiency of CDMA with Random Spreading," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 622–640, Mar. 1999.

[30] 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.

[31] D. Chizhik, F. Rashid-Farrokhi, J. Ling, and A. Lozano, "Effects of antenna separation on the capacity of BLAST in correlated channels," *IEEE Communications Letters*, vol. 4, pp. 337–339, Nov. 2000.

[32] K. I. Pedersen, P. E. Mogensen, and B. H. Fleury, "A Stochastic Model of the Temporal and Azimuthal Dispersion Seen at the Base Station in Outdoor Propagation Environments," *IEEE Trans. Inf. Theory*, vol. 49, no. 2, p. 437, Mar. 2000.

[33] A. L. Moustakas, H. U. Baranger, L. Balents, A. M. Sengupta, and S. H. Simon, "Communication through a diffusive medium: Coherence and capacity," *Science*, vol. 287, pp. 287–290, 2000.

[34] Z. D. Bai and J. W. Silverstein, *Spectral analysis of large dimensional random matrices*, 2nd ed. New York, NY, USA: Springer Series in Statistics, 2009.

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

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

[37] E. C. Titchmarsh, *The Theory of Functions*. New York, NY, USA: Oxford University Press, 1939.

[38] W. Rudin, *Real and Complex Analysis*, 3rd ed. McGraw-Hill Series in Higher Mathematics, May 1986.

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

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

[41] 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.

[42] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 1985.

[43] S. Wagner, "MU-MIMO Transmission and Reception Techniques for the Next Generation of Cellular Wireless Standards (LTE-A)," Ph.D. dissertation, EURECOM, 2229 route des crétes, BP 193 F-06560 Sophia-Antipolis cedex, 2011. [Online]. Available: http://www.eurecom.fr/people/cifre_wagner.en.htm

[44] J. Hoydis, M. Debbah, and M. Kobayashi, "Asymptotic Moments for Interference Mitigation in Correlated Fading Channels," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, Saint Petersburg, Russia, Jul. 31 - Aug. 5 2011. [Online]. Available: http://arxiv.org/abs/1104.4911

[45] P. Billingsley, *Probability and Measure*, 3rd ed. John Wiley & Sons, Inc., 1995.