



This is a repository copy of *Large-System Analysis of Correlated MIMO Multiple Access Channels with Arbitrary Signaling in the Presence of Interference*.

White Rose Research Online URL for this paper:
<http://eprints.whiterose.ac.uk/95094/>

Version: Accepted Version

Article:

Girnyk, M.A., Vehkaperä, M. orcid.org/0000-0002-3085-538X and Rasmussen, L.K. (2014) Large-System Analysis of Correlated MIMO Multiple Access Channels with Arbitrary Signaling in the Presence of Interference. *IEEE Transactions on Wireless Communications*, 13 (4). pp. 2060-2073. ISSN 1536-1276

<https://doi.org/10.1109/TWC.2014.022614.130860>

Reuse

Unless indicated otherwise, fulltext items are protected by copyright with all rights reserved. The copyright exception in section 29 of the Copyright, Designs and Patents Act 1988 allows the making of a single copy solely for the purpose of non-commercial research or private study within the limits of fair dealing. The publisher or other rights-holder may allow further reproduction and re-use of this version - refer to the White Rose Research Online record for this item. Where records identify the publisher as the copyright holder, users can verify any specific terms of use on the publisher's website.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

Large-System Analysis of Correlated MIMO Multiple Access Channels with Arbitrary Signaling in the Presence of Interference

Maksym A. Girnyk, Mikko Vehkaperä, Lars K. Rasmussen

Abstract— Presence of multiple antennas on both sides of a communication channel promises significant improvements in system throughput and power efficiency. In effect, a new class of large multiple-input multiple-output (MIMO) communication systems has recently emerged and attracted both scientific and industrial attention. To analyze these systems in realistic scenarios, one has to include such aspects as co-channel interference, multiple access and spatial correlation. In this paper, we study the properties of correlated MIMO multiple-access channels in the presence of external interference. Using the replica method from statistical physics, we derive the ergodic sum-rate of the communication for arbitrary signal constellations when the numbers of antennas at both ends of the channel grow large. Based on these asymptotic expressions, we also address the problem of sum-rate maximization using statistical channel information and linear precoding. The numerical results demonstrate that when the interfering terminals use discrete constellations, the resulting interference becomes easier to handle compared to Gaussian signals. Thus, it may be possible to accommodate more interfering transmitter-receiver pairs within the same area as compared to the case of Gaussian signals. In addition, we demonstrate numerically for the Gaussian and QPSK signaling schemes that it is possible to design precoder matrices that significantly improve the achievable rates at low-to-mid range of signal-to-noise ratios when compared to isotropic precoding.

I. INTRODUCTION

During the last decade, multi-antenna communications has received an increased interest both from academia and industry. Pioneering research by Foschini, Gans and Telatar [1], [2] on the topic suggested that the new class of multiple-input multiple-output (MIMO) systems allowed the transmission rate to be increased roughly linearly in the number of antennas available at the transmitter and receiver. Measurements both indoors [3] and outdoors [4] have also confirmed the throughput gains of the multi-antenna transmission.

The main price to pay for the benefits offered by multi-antenna transmission is the hardware and signal processing complexity at both the transmitter and receiver. It is therefore of great importance to analyze the potential performance gains of MIMO processing in realistic scenarios before employing

The research leading to these results has received funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement no 228044. The work has further been supported in parts by the ARC Grant DP0986089, and the VR grant 621-2009-4666. The material of this paper was presented in parts at the IEEE GLOBECOM Conference 2012, Anaheim, U.S.A., Dec. 2012.

The authors are with the School of Electrical Engineering and the ACCESS Linnaeus Centre at KTH Royal Institute of Technology, Stockholm, Sweden (e-mail: {mgyr, mikkov, lkra}@kth.se).

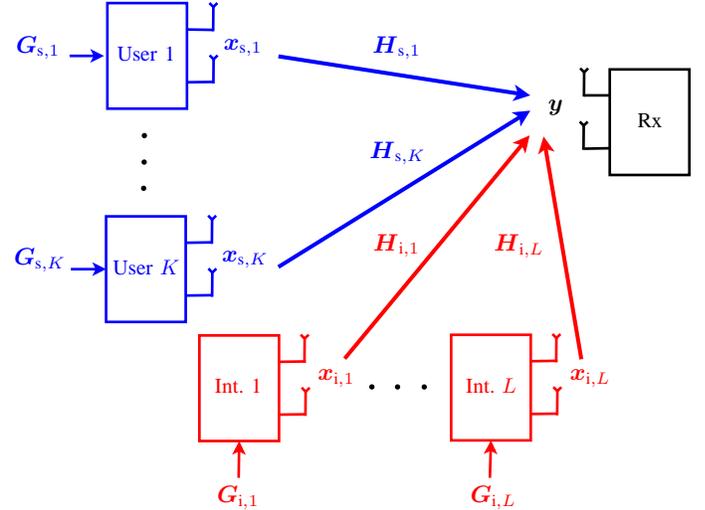


Fig. 1. MIMO multiple-access channel in the presence of interference.

the techniques in practice. For example, in the uplink of a cellular system, the effects of *co-channel interference* emerging from other cells need to be taken into account. As observed in [5], such interference has also a surprising influence on the optimal power allocation strategy at the transmitter. In addition to co-channel interference, *spatial correlation* [6] and the type of channel inputs [7] have a great impact on the achievable rate of the channel. However, analysis of realistic scenarios tend to be mathematically challenging and numerical simulations are time consuming, especially if discrete signaling is employed at the transmitter. Some simplifying assumptions are therefore needed to make the problem tractable.

Asymptotic approaches developed within the field of *random matrix theory* greatly facilitate the analysis of achievable ergodic rates (mutual information averaged over channel realizations) in MIMO systems. Such methods were used already in the early works [1], [2] to assess the capacity of multi-antenna transmission. At the same time, several approaches using random matrix theory for the analysis of the spectral efficiency of large code division multiple access (CDMA) systems [8]–[12] were reported. The multi-antenna results were later extended to the case of spatial correlation in [13] and then to *MIMO multiple-access channel* (MIMO-MAC) in [14] (*vide* Fig. 1). Some analysis of MIMO systems in the case of co-channel interference have also been carried out under the assumption of Gaussian channel inputs. The first

analytical results using random matrix theory were obtained in [15] assuming uncorrelated channels and interferers. Some later efforts [16]–[18] have extended this analysis to different assumptions about correlations and numbers of antennas present in the system.

Although the above methods assume formally that the system size grows without bound, they usually provide a good approximation also for the performance of finite-sized systems. Furthermore, the underlying large-system assumption *per se* has recently found a practical application in so-called *massive MIMO* [19]–[22]. Typically this concept entails a multiuser system where a single base station, equipped with a very large antenna array, is used to serve a smaller number of terminals simultaneously. Apart from the aforementioned throughput gains, such systems allow for a significant reduction of the transmit power and reduced-complexity signal processing [21]. Consequently, asymptotic random-matrix methods have been widely applied for the analysis of various aspects of massive MIMO [23]–[25] systems.

The aforementioned studies regarding MIMO systems with co-channel interference have all concentrated on the special case, where Gaussian signals are transmitted both by the desired user and the interfering terminals. This is in contrast to real-world systems, where discrete constellations such as QPSK and QAM are used. These realistic cases are, however, out-of-bounds for random matrix theory, except for setups where sub-optimal linear detection and per-stream decoding is considered. To investigate the performance bounds of generic systems with *non-Gaussian channel inputs*, a tool borrowed from the field of statistical physics, namely the *replica method*, has been recently used.

The replica method was invented by Kac [26], and is widely known due to its early applications to *spin glasses* [27], [28]. The replica framework provides a powerful set of mathematical tools for computing average quantities within large many-body systems and has since been applied to various problems in science and engineering. In the context of information theory, it was used to assess the spectral efficiency of large CDMA systems with antipodal signaling by Tanaka in [29], [30]. Later, Guo and Verdú generalized the approach to CDMA with arbitrary signaling [31]. Meanwhile, in [7] and [32] the method was applied to spatially correlated MIMO channels with binary inputs. These works were further generalized in [33] to the analysis of the sum-rate of a MIMO-MAC. A somewhat different approach was taken in [34], where the replica method was used to analyze the moments of mutual information of a MIMO system with co-channel interference. The results in [34] were obtained, however, under the assumption of Gaussian signaling at all terminals.

In the present paper we extend our previous work [35] and investigate the performance and sum-rate maximization of a correlated MIMO-MAC using the replica method. The analysis encompasses the presence of correlated non-Gaussian interferers and arbitrary inputs at all terminals. As in [29]–[34], the results are obtained under the technical assumption of replica symmetric ansatz. To summarize, the following contributions are reported:

- We derive an expression for the asymptotic sum-rate

of the MIMO-MAC with spatial correlation and in the presence of spatially correlated multi-antenna interferers. The analysis is valid for arbitrary channel inputs at all terminals¹ and is carried out in the *large-system limit* (LSL), where the numbers of antennas at both ends of each MIMO channel grow without bound at a constant rate. As expected, several prior results are obtained as special cases of the analysis. For instance, in the absence of interferers, our results degenerate to those reported in [14], [36] when Gaussian signals are employed and to those provided in [33] when arbitrary signal constellations are used. Finally, in the presence of interference and under assumption of Gaussian signaling, our results partly reduce to the expression of the mean mutual information derived in [34].

- We address the precoder optimization problem for both Gaussian and finite-alphabet signaling schemes under the assumptions of full channel state information (CSI) at the receiver² and *statistical CSI at the transmitter*. By using the asymptotic sum-rate as an objective function for the corresponding optimization problem, we obtain the precoding matrices for each user.

The remainder of the paper is organized as follows. In the following section, we describe the system model and formulate the main problem. In addition, we discuss the necessary details regarding the MIMO channels with perfect CSI at the transmitter. Next, in Section III, we present the main result of the paper, that is, the asymptotic sum-rate of a MIMO-MAC in the presence of interference. Section IV then addresses the precoder optimization problem, followed by Section V, where we present numerical results and discussion. Finally, in Section VI, we conclude the paper. The proofs are relegated to the appendices.

Notation: Throughout this paper we will use upper case bold-faced letters to denote matrices, e.g., \mathbf{A} , with elements denoted by $[\mathbf{A}]_{i,j}$, lower case bold-faced letters to denote column vectors, e.g., \mathbf{a} , with elements a_i , and lower case light-faced letters to denote scalar variables, e.g., a . Superscripts T and H denote transpose and Hermitian adjoint operators, respectively. Meanwhile, $\mathbf{A}^{1/2}$, $\text{tr}\{\mathbf{A}\}$ and $\det(\mathbf{A})$ denote the principal square root, the trace and the determinant of matrix \mathbf{A} . We differentiate between operators $\text{diag}(\mathbf{a})$, which denotes a diagonal matrix containing the coefficients of vector \mathbf{a} on its

¹Throughout the paper, we assume that the signaling scheme used at each interfering terminal is known to the receiver. In the cellular setting, such information can be exchanged between neighboring base stations *via* an existent backhaul link with a very small overhead. Alternatively, our results provide an upper bound to the setting where the signaling schemes are not known or the base station is misinformed about them.

²In practice, the CSI at the receiver is estimated by using known training sequences sent by the users. In the cellular setting, the same procedure can also be used to estimate the channels of the interferers if there is sufficient synchronization and the base station is informed about the set of training sequences that are in use in the neighboring cell. The latter can be achieved with a small overhead in the backhaul link. Furthermore, in order to implement soft handover between the cells, the base stations tend to establish tight synchronization with the users from other cells close to the cell border. Hence, estimating the interfering users' channels is possible also in practical systems. The mathematical model considered in this paper provides thus an upper bound for the performance of a practical systems that estimates the CSI or has only partial knowledge of the latter.

main diagonal, and $\text{Diag}(\mathbf{A})$, which denotes a column vector containing the diagonal entries of matrix \mathbf{A} . Also, \mathbf{I} , $\mathbf{0}$ and $\mathbf{1}$ denote the identity matrix, the zero matrix and the all-ones vector of appropriate sizes. Operator $\mathbb{E}\{\cdot\}$ denotes the expectation, $\mathbb{1}(\cdot)$ denotes the indicator function, \otimes represents the Kronecker product, and $\mathbf{A} \succeq \mathbf{0}$ implies that the matrix \mathbf{A} is positive semidefinite. Finally, $\text{Re}\{\cdot\}$ and $\text{Im}\{\cdot\}$ stand for the real and imaginary parts of the argument.

II. PRELIMINARIES

A. System Model

Consider the scenario where K multi-antenna terminals communicate to a single multi-antenna receiver in the presence of L multi-antenna interferers, as depicted in Fig. 1. An uplink cellular communication system in the presence of inter-cell interference can be regarded as an example of such a scenario. The numbers of antennas at transmitter k , interferer ℓ and the receiver, are denoted by $M_{s,k}$, $M_{i,\ell}$, and N , respectively. The discrete-time received vector is given by

$$\mathbf{y} = \sum_{k=1}^K \mathbf{H}_{s,k} \mathbf{x}_{s,k} + \sum_{\ell=1}^L \mathbf{H}_{i,\ell} \mathbf{x}_{i,\ell} + \mathbf{n}, \quad (1)$$

where $\mathbf{x}_{s,k} \in \mathbb{C}^{M_{s,k}}$ is the zero-mean transmitted signal vector of the k th user with covariance matrix $\mathbb{E}\{\mathbf{x}_{s,k} \mathbf{x}_{s,k}^H\} = \mathbf{P}_{s,k}$ and $\mathbf{x}_{i,\ell} \in \mathbb{C}^{M_{i,\ell}}$ is the transmitted signal vector of the ℓ th interferer having zero-mean and covariance $\mathbb{E}\{\mathbf{x}_{i,\ell} \mathbf{x}_{i,\ell}^H\} = \mathbf{P}_{i,\ell}$. To satisfy long-term power constraints at the transmitters, we require that $\text{tr}\{\mathbf{P}_{s,k}\} \leq M_{s,k}$ and $\text{tr}\{\mathbf{P}_{i,\ell}\} \leq M_{i,\ell}$. The noise vector $\mathbf{n} \in \mathbb{C}^N$ has independent circularly symmetric complex Gaussian (CSCG) entries with unit variance. Matrices $\mathbf{H}_{s,k} \in \mathbb{C}^{N \times M_{s,k}}$ and $\mathbf{H}_{i,\ell} \in \mathbb{C}^{N \times M_{i,\ell}}$ denote the MIMO channels between user k and the receiver and between interferer ℓ and the receiver, respectively. The channels are assumed to be flat-fading and are modeled *via* the Kronecker model [37], that is

$$\mathbf{H}_{s,k} = \sqrt{\frac{\rho_{s,k}}{M_{s,k}}} \mathbf{R}_{s,k}^{1/2} \mathbf{W}_{s,k} \mathbf{T}_{s,k}^{1/2}, \quad (2a)$$

$$\mathbf{H}_{i,\ell} = \sqrt{\frac{\rho_{i,\ell}}{M_{i,\ell}}} \mathbf{R}_{i,\ell}^{1/2} \mathbf{W}_{i,\ell} \mathbf{T}_{i,\ell}^{1/2}, \quad (2b)$$

where $\rho_{s,k}$ and $\rho_{i,\ell}$ represent average signal-to-noise ratios (SNRs) of the corresponding links and matrices $\mathbf{W}_{s,k}$ and $\mathbf{W}_{i,\ell}$ have i.i.d. CSCG entries of unit variance. The correlation matrices at the receive end are denoted by $\mathbf{R}_{s,k}$ and $\mathbf{R}_{i,\ell}$, while $\mathbf{T}_{s,k}$ and $\mathbf{T}_{i,\ell}$ represent the correlation matrices at the transmit end of the corresponding channels. To ensure that the correlation matrices do not influence the average path gains, they are normalized as

$$\text{tr}\{\mathbf{R}_{s,k}\} = N, \quad \text{tr}\{\mathbf{T}_{s,k}\} = M_{s,k}, \quad (3a)$$

$$\text{tr}\{\mathbf{R}_{i,\ell}\} = N, \quad \text{tr}\{\mathbf{T}_{i,\ell}\} = M_{i,\ell}. \quad (3b)$$

For later convenience, we write the input covariance matrices in terms of two *precoder matrices*, that is, $\mathbf{P}_{s,k} = \mathbf{G}_{s,k} \mathbf{G}_{s,k}^H$ and $\mathbf{P}_{i,\ell} = \mathbf{G}_{i,\ell} \mathbf{G}_{i,\ell}^H$, and let them depend on the statistical CSI, *i.e.*, the knowledge of $\{\rho_{s,k}, \mathbf{T}_{s,k}, \mathbf{R}_{s,k}\}$ and $\{\rho_{i,\ell}, \mathbf{T}_{i,\ell}, \mathbf{R}_{i,\ell}\}$, respectively. Thus, denoting $\mathbf{s}_{s,k}$ and $\mathbf{s}_{i,\ell}$

for independent vectors with i.i.d. zero-mean unit variance entries, we may write $\mathbf{x}_{s,k} = \mathbf{G}_{s,k} \mathbf{s}_{s,k}$ and $\mathbf{x}_{i,\ell} = \mathbf{G}_{i,\ell} \mathbf{s}_{i,\ell}$ without loss of generality. This formulation is especially useful when we consider the optimization of the input covariance for discrete signals. For notational simplicity, we also denote $M_s \triangleq \sum_{k=1}^K M_{s,k}$ and $M_i \triangleq \sum_{\ell=1}^L M_{i,\ell}$, and rewrite the input-output relation of the resulting MIMO channel as

$$\mathbf{y} = \mathbf{H}_s \mathbf{x}_s + \mathbf{H}_i \mathbf{x}_i + \mathbf{n}, \quad (4)$$

where $\mathbf{H}_s \triangleq [\mathbf{H}_{s,1}, \dots, \mathbf{H}_{s,K}] \in \mathbb{C}^{N \times M_s}$, $\mathbf{H}_i \triangleq [\mathbf{H}_{i,1}, \dots, \mathbf{H}_{i,L}] \in \mathbb{C}^{N \times M_i}$, $\mathbf{x}_s \triangleq [\mathbf{x}_{s,1}^T, \dots, \mathbf{x}_{s,K}^T]^T \in \mathbb{C}^{M_s}$, $\mathbf{x}_i \triangleq [\mathbf{x}_{i,1}^T, \dots, \mathbf{x}_{i,L}^T]^T \in \mathbb{C}^{M_i}$.

B. Problem Statement

Define the instantaneous CSI at the receiver as $\mathcal{H} \triangleq \{\mathbf{H}_s, \mathbf{H}_i\}$. Given that all channels are ergodic and the receiver knows \mathcal{H} , the distribution of \mathbf{x}_s , as well as the distribution of \mathbf{x}_i , we can write down the average mutual information

$$I(\mathbf{y}; \mathbf{x}_s) = h(\mathbf{y}|\mathcal{H}) - h(\mathbf{y}|\mathbf{x}_s, \mathcal{H}), \quad (5)$$

where the differential entropy terms are given by

$$h(\mathbf{y}|\mathcal{H}) = -\mathbb{E}_{\mathbf{y}, \mathcal{H}} \ln \mathbb{E}_{\mathbf{x}_s, \mathbf{x}_i} p(\mathbf{y}|\mathbf{x}_s, \mathbf{x}_i, \mathcal{H}), \quad (6a)$$

$$h(\mathbf{y}|\mathbf{x}_s, \mathcal{H}) = -\mathbb{E}_{\mathbf{y}, \mathbf{x}_s, \mathcal{H}} \ln \mathbb{E}_{\mathbf{x}_i} p(\mathbf{y}|\mathbf{x}_s, \mathbf{x}_i, \mathcal{H}), \quad (6b)$$

and the conditional distribution of the channel (4) reads

$$p(\mathbf{y}|\mathbf{x}_s, \mathbf{x}_i, \mathcal{H}) = \frac{1}{\pi^N} e^{-\|\mathbf{y} - \mathbf{H}_s \mathbf{x}_s - \mathbf{H}_i \mathbf{x}_i\|^2}. \quad (7)$$

The mutual information in (5) represents an achievable sum-rate of the MIMO-MAC (1) when the receiver does not decode the interference signal \mathbf{x}_i . Given statistical knowledge of the channels, the sum-rate could then, in principle, be maximized by designing the precoder matrices $\mathbf{G}_{s,k}$, $\forall k$. Unfortunately, the explicit expression for (5) is not known in general. Moreover, its numerical evaluation is computationally expensive due to averaging of (6a) and (6b) over the channel realizations. Even more serious difficulty arises when the data symbols are non-Gaussian. In this case, one needs to compute two sums over an exponential number (w.r.t. the numbers of transmit antennas and bits in the constellation) of terms for every realization of the channel.

The aim of the present paper is to find a computationally feasible expression for the ergodic mutual information (5) given arbitrary channel inputs. To make the analysis tractable, we consider the asymptotic regime where the system size grows large and use the replica method to compute the individual entropy terms. These expressions (*vide* Section III) are then used to optimize the covariance matrices so that the mutual information is maximized (*vide* Section IV). Finally, the large system result is used to give an approximation for the original quantity (5) when the system size is finite.

C. Mutual Information and the MMSE of a Fixed MIMO Channel

To finish this section, we discuss the problem of finding the mutual information and the minimum mean squared error

(MMSE) of a fixed MIMO channel. These results are used later in the paper to evaluate the asymptotic mutual information obtained *via* the replica method.

1) *General Case*: Consider the following multi-antenna communication channel

$$\mathbf{z} = \mathbf{A}\mathbf{x} + \mathbf{w}, \quad (8)$$

where \mathbf{A} is a fixed $N \times M$ channel matrix and $\mathbf{w} \in \mathbb{C}^N$ has i.i.d. CSCG elements of unit variance. The channel inputs $\mathbf{x} = \mathbf{G}\mathbf{s} \in \mathbb{C}^M$ are a combination of a precoder matrix \mathbf{G} and vector \mathbf{s} that has i.i.d. zero-mean unit variance entries, with constraint $\text{tr}\{\mathbf{P}\} \leq M$ where $\mathbf{E}\{\mathbf{x}\mathbf{x}^H\} = \mathbf{P}$. The conditional distribution of this fixed *single-user* (su) channel³ is given by

$$p_{\text{su}}(\mathbf{z}|\mathbf{x}, \mathbf{A}) = \frac{1}{\pi^M} e^{-\|\mathbf{z} - \mathbf{A}\mathbf{x}\|^2}. \quad (9)$$

The posterior mean estimate of \mathbf{x} is denoted $\langle \mathbf{x} \rangle \triangleq \mathbf{E}\{\mathbf{x}|\mathbf{z}, \mathbf{A}\}$, where the expectation is taken over the posterior density $p_{\text{su}}(\mathbf{x}|\mathbf{z}, \mathbf{A})$ obtained from the prior distribution $p(\mathbf{x})$ and (9) *via* Bayes' theorem. For future convenience, the conditional *MMSE matrix* is defined here through parametrization

$$\text{mmse}_{\text{su}}(\mathbf{x}, \mathbf{A}) \triangleq \mathbf{E}_{\mathbf{z}, \mathbf{x}} \left\{ (\mathbf{x} - \langle \mathbf{x} \rangle)(\mathbf{x} - \langle \mathbf{x} \rangle)^H \right\} \in \mathbb{C}^{M \times M}, \quad (10)$$

where the expectation is w.r.t. the joint distribution $p_{\text{su}}(\mathbf{z}, \mathbf{x}|\mathbf{A})$. Similarly, the mutual information reads

$$I_{\text{su}}(\mathbf{z}; \mathbf{x}|\mathbf{A}) = \mathbf{E}_{\mathbf{z}, \mathbf{x}} \ln p_{\text{su}}(\mathbf{z}|\mathbf{x}, \mathbf{A}) - \mathbf{E}_{\mathbf{z}} \ln \mathbf{E}_{\mathbf{x}} p_{\text{su}}(\mathbf{z}|\mathbf{x}, \mathbf{A}), \quad (11)$$

where the expectations are again w.r.t. $p_{\text{su}}(\mathbf{z}, \mathbf{x}|\mathbf{A})$.

Below we present two important special cases and provide the corresponding expressions for the mutual information and the MMSE.

Example 1 (Gaussian inputs). The MMSE detector becomes linear if the channel input \mathbf{x} is a CSCG vector given \mathbf{G} . In this case, the output of the MMSE detector reads

$$\langle \mathbf{x} \rangle = \left(\mathbf{P}^{-1} + \mathbf{A}^H \mathbf{A} \right)^{-1} \mathbf{A}^H \mathbf{z}, \quad (12)$$

and the MMSE matrix is given by

$$\text{mmse}_{\text{su}}(\mathbf{x}, \mathbf{A}) = \left(\mathbf{P}^{-1} + \mathbf{A}^H \mathbf{A} \right)^{-1}. \quad (13)$$

The mutual information reduces also to the well-known formula

$$I_{\text{su}}(\mathbf{z}; \mathbf{x}|\mathbf{A}) = \ln \det \left(\mathbf{I}_N + \mathbf{A}\mathbf{P}\mathbf{A}^H \right). \quad (14)$$

Example 2 (Discrete channel inputs). Let the entries of \mathbf{s} be independently drawn from a discrete constellation (*e.g.*, QPSK, QAM) of cardinality C , so that \mathbf{s} is uniformly distributed over the set $\{\mathbf{s}_1, \dots, \mathbf{s}_{C^M}\}$. For fixed \mathbf{G} we may then treat also \mathbf{x} as being uniformly drawn from a set $\{\mathbf{x}_1, \dots, \mathbf{x}_{C^M}\}$ of

possible input vectors. Denoting

$$p_{\text{su}}(\mathbf{z}|\mathbf{A}) = \frac{1}{C^M} \sum_{i=1}^{C^M} \frac{1}{\pi^M} e^{-\|\mathbf{z} - \mathbf{A}\mathbf{x}_i\|^2}, \quad (15)$$

the MMSE estimate of \mathbf{x} is by definition given as

$$\langle \mathbf{x} \rangle = \frac{1}{C^M} \sum_{i=1}^{C^M} \frac{\mathbf{x}_i p_{\text{su}}(\mathbf{z}|\mathbf{x}_i, \mathbf{A})}{p_{\text{su}}(\mathbf{z}|\mathbf{A})}. \quad (16)$$

The MMSE matrix is thus obtained from

$$\text{mmse}_{\text{su}}(\mathbf{x}, \mathbf{A}) = \mathbf{P} - \int \langle \mathbf{x} \rangle \langle \mathbf{x} \rangle^H p_{\text{su}}(\mathbf{z}|\mathbf{A}) d\mathbf{z}, \quad (17)$$

while the mutual information between \mathbf{z} and \mathbf{x} reads

$$I_{\text{su}}(\mathbf{z}; \mathbf{x}|\mathbf{A}, p_{\text{su}}) = N + M \ln C - \frac{1}{C^M} \sum_{i=1}^{C^M} \mathbf{E}_{\mathbf{w}} \left\{ \ln \sum_{j=1}^{C^M} e^{-\|\mathbf{A}(\mathbf{x}_i - \mathbf{x}_j) + \mathbf{w}\|^2} \right\}. \quad (18)$$

2) *Parallel Gaussian Channels*: Assume that the channel matrix \mathbf{A} is diagonal with real-valued entries a_1, \dots, a_M . Let \mathbf{G} also be a real diagonal matrix formed by g_1, \dots, g_M so that the MIMO channel (8) decouples into a bank of parallel channels

$$z_m = a_m x_m + w_m. \quad (19)$$

The MMSE estimate of x_m for the m th channel is then $\langle x_m \rangle = \mathbf{E}\{x_m|z_m, a_m\}$, while the MMSE matrix is diagonal with $\mathbf{E}_{z_m, x_m}\{|x_m - \langle x_m \rangle|^2|a_m\}$ as its (m, m) th element. The mutual information (11) reduces in this case to a form

$$I_{\text{su}}(\mathbf{z}; \mathbf{x}|\mathbf{A}) = \sum_{m=1}^M I(z_m; x_m|a_m, p_{\text{su}}). \quad (20)$$

In the following examples we discuss three analytically tractable special cases, which will be useful later in Section V.

Example 3 (Gaussian inputs). In this scenario, we have $x_m \sim \mathcal{CN}(0, g_m^2)$ so that the MMSE estimate of x_m becomes

$$\langle x_m \rangle = \frac{g_m^2 a_m z_m}{1 + g_m^2 a_m^2}, \quad (21)$$

leading to

$$\mathbf{E}_{z_m, x_m} \{|x_m - \langle x_m \rangle|^2|a_m\} = \frac{g_m^2}{1 + g_m^2 a_m^2}. \quad (22)$$

The mutual information between the input and output of (19), on the other hand, is quantified as

$$I_{\text{su}}(z_m; x_m|a_m) = \ln(1 + g_m^2 a_m^2), \quad (23)$$

so that using (20) we obtain the total achievable sum-rate (11) of a single-user MIMO system with fixed diagonal channel.

Example 4 (QPSK inputs). When the prior distribution of the elements of \mathbf{s} is given by $p(\mathbf{s}) = 1/4 \delta(\mathbf{s} \pm 1/\sqrt{2} \pm j/\sqrt{2})$, we have

$$\langle x_m \rangle = \frac{g_m}{\sqrt{2}} \tanh\left(\sqrt{2} g_m a_m \text{Re}\{z_m\}\right) + j \frac{g_m}{\sqrt{2}} \tanh\left(\sqrt{2} g_m a_m \text{Im}\{z_m\}\right). \quad (24)$$

³In this and the following sections, the probabilities related to the channel (8) are denoted p_{su} . This is to make a clear separation to the probabilities related to the original channel (4).

$$\mathbf{E}_{z_m, x_m} \{|x_m - \langle x_m \rangle|^2 | a_m\} = g_m^2 - \frac{g_m^2}{10\sqrt{\pi}} \int_{\mathbb{R}} \frac{\left(3e^{-\frac{4g_m^2 a_m^2}{5}} \sinh\left(\frac{6g_m a_m}{\sqrt{10}} s\right) + \sinh\left(\frac{2g_m a_m}{\sqrt{10}} s\right) \right)^2}{e^{-\frac{4g_m a_m}{5}} \cosh\left(\frac{6g_m a_m}{\sqrt{10}} s\right) + \cosh\left(\frac{2g_m a_m}{\sqrt{10}} s\right)} e^{-\frac{10s^2 + g_m^2 a_m^2}{10}} ds, \quad (27)$$

Furthermore, the (m, m) th element of the MMSE matrix is given by

$$\begin{aligned} \mathbf{E}_{z_m, x_m} \{|x_m - \langle x_m \rangle|^2 | a_m\} \\ = g_m^2 - \frac{g_m^2}{\sqrt{2\pi}} \int_{\mathbb{R}} \tanh(g_m^2 a_m^2 - g_m a_m s) e^{-\frac{s^2}{2}} ds, \end{aligned} \quad (25)$$

and the per-stream mutual information is evaluated as

$$\begin{aligned} I_{\text{su}}(z_m; x_m | a_m) = 2g_m^2 a_m^2 \\ - \sqrt{\frac{2}{\pi}} \int_{\mathbb{R}} \ln \cosh(g_m^2 a_m^2 - g_m a_m s) e^{-\frac{s^2}{2}} ds. \end{aligned} \quad (26)$$

Example 5 (16-QAM inputs). When the elements of \mathbf{s} , are uniformly drawn from the standard 16-QAM constellation, the diagonal terms of the MMSE matrix are evaluated as in (27) on the top of the page (the minor typo in [38, (27) – (28)] is corrected there). The mutual information is then obtained numerically through the I-MMSE relation [39].

III. ASYMPTOTIC ACHIEVABLE SUM-RATE

In this section, we present the main findings of the paper, namely, the asymptotic sum-rate of reliable communication over a multi-access MIMO channel in the presence of interferers. The expression is derived in the large-system limit (LSL), where the numbers of antennas at each terminal grow without bounds at constant ratios, *i.e.*, $\beta_{s,k} M_{s,k} = N \rightarrow \infty$, $\forall k \in \{1, \dots, K\}$ and $\beta_{i,\ell} M_{i,\ell} = N \rightarrow \infty$, $\forall \ell \in \{1, \dots, L\}$, where $\beta_{s,k}$ and $\beta_{i,\ell}$ are finite positive constants.

In the remainder of the section, we first present the asymptotic result for a general (correlated) MIMO-MAC, and then specialize to the uncorrelated case where the expression are much simpler.

A. General Case

The main result of the paper is given in the following proposition⁴.

Proposition 1. *Let the input distributions $p(\mathbf{x}_{s,k})$ and $p(\mathbf{x}_{i,\ell})$, as well as the spatial correlation matrices $\mathbf{T}_{s,k}$, $\mathbf{R}_{s,k}$, $\mathbf{T}_{i,\ell}$ and $\mathbf{R}_{i,\ell}$, be given. Then, the ergodic mutual information (5) normalized by the number of antennas at the receiver reads in the LSL as*

$$\frac{1}{N} I(\mathbf{y}; \mathbf{x}_s) = h_s - h_i + \mathcal{O}\left(\frac{1}{N}\right), \quad (28)$$

⁴Even though some of the steps in the replica method are still lacking rigorous proof, the key results of the present paper are presented as propositions, being a convention in the replica calculus literature.

where

$$\begin{aligned} h_s = \frac{1}{N} \sum_{k=1}^K I_{\text{su}}(\mathbf{z}_{s,k}; \mathbf{x}_{s,k} | \mathbf{A}_{s,k}) - \sum_{k=1}^K \frac{1}{\beta_{s,k}} \xi_{s,k} \varepsilon_{s,k} \\ + \frac{1}{N} \sum_{\ell=1}^L I_{\text{su}}(\mathbf{z}_{i,\ell}; \mathbf{x}_{i,\ell} | \mathbf{A}_{i,\ell}) - \sum_{\ell=1}^L \frac{1}{\beta_{i,\ell}} \xi_{i,\ell} \varepsilon_{i,\ell} \\ + \frac{1}{N} \ln \det \left(\mathbf{I}_N + \sum_{k=1}^K \varepsilon_{s,k} \mathbf{R}_{s,k} + \sum_{\ell=1}^L \varepsilon_{i,\ell} \mathbf{R}_{i,\ell} \right) \\ + 1 + \ln \pi, \end{aligned} \quad (29)$$

$$\begin{aligned} h_i = \frac{1}{N} \sum_{\ell=1}^L I_{\text{su}}(\bar{\mathbf{z}}_{i,\ell}; \mathbf{x}_{i,\ell} | \bar{\mathbf{A}}_{i,\ell}) - \sum_{\ell=1}^L \frac{1}{\beta_{i,\ell}} \bar{\xi}_{i,\ell} \bar{\varepsilon}_{i,\ell} \\ + \frac{1}{N} \ln \det \left(\mathbf{I}_N + \sum_{\ell=1}^L \bar{\varepsilon}_{i,\ell} \mathbf{R}_{i,\ell} \right) + 1 + \ln \pi. \end{aligned} \quad (30)$$

The parameters $\varepsilon_{s,k}$, $\xi_{s,k}$, $\varepsilon_{i,\ell}$, $\xi_{i,\ell}$, $\bar{\varepsilon}_{i,\ell}$ and $\bar{\xi}_{i,\ell}$ satisfy the following set of fixed-point equations⁵

$$\xi_{s,k} = \frac{1}{M_{s,k}} \text{tr} \left\{ \mathbf{R}_{s,k} \left[\mathbf{I}_N + \sum_{k=1}^K \varepsilon_{s,k} \mathbf{R}_{s,k} + \sum_{\ell=1}^L \varepsilon_{i,\ell} \mathbf{R}_{i,\ell} \right]^{-1} \right\}, \quad (31a)$$

$$\xi_{i,\ell} = \frac{1}{M_{i,\ell}} \text{tr} \left\{ \mathbf{R}_{i,\ell} \left[\mathbf{I}_N + \sum_{k=1}^K \varepsilon_{s,k} \mathbf{R}_{s,k} + \sum_{\ell=1}^L \varepsilon_{i,\ell} \mathbf{R}_{i,\ell} \right]^{-1} \right\}, \quad (31b)$$

$$\bar{\xi}_{i,\ell} = \frac{1}{M_{i,\ell}} \text{tr} \left\{ \mathbf{R}_{i,\ell} \left[\mathbf{I}_N + \sum_{\ell=1}^L \bar{\varepsilon}_{i,\ell} \mathbf{R}_{i,\ell} \right]^{-1} \right\}, \quad (31c)$$

$$\varepsilon_{s,k} = \frac{\rho_{s,k}}{M_{s,k}} \text{tr} \{ \mathbf{mmse}_{\text{su}}(\mathbf{x}_{s,k}, \mathbf{A}_{s,k}) \mathbf{T}_{s,k} \}, \quad (31d)$$

$$\varepsilon_{i,\ell} = \frac{\rho_{i,\ell}}{M_{i,\ell}} \text{tr} \{ \mathbf{mmse}_{\text{su}}(\mathbf{x}_{i,\ell}, \mathbf{A}_{i,\ell}) \mathbf{T}_{i,\ell} \}, \quad (31e)$$

$$\bar{\varepsilon}_{i,\ell} = \frac{\rho_{i,\ell}}{M_{i,\ell}} \text{tr} \{ \mathbf{mmse}_{\text{su}}(\mathbf{x}_{i,\ell}, \bar{\mathbf{A}}_{i,\ell}) \mathbf{T}_{i,\ell} \}, \quad (31f)$$

where $\mathbf{A}_{s,k} = \sqrt{\rho_{s,k} \xi_{s,k}} \mathbf{T}_{s,k}^{1/2}$, $\mathbf{A}_{i,\ell} = \sqrt{\rho_{i,\ell} \xi_{i,\ell}} \mathbf{T}_{i,\ell}^{1/2}$, $\bar{\mathbf{A}}_{i,\ell} = \sqrt{\rho_{i,\ell} \bar{\xi}_{i,\ell}} \mathbf{T}_{i,\ell}^{1/2}$ and the MMSE matrices are obtained via (10).

Proof: In the mutual information expression of (5), term (6a) represents the sum-rate of an uplink system with all interferers considered as being desired users and Gaussian noise being the only source of disturbance. The asymptotic

⁵In general, the fixed-point equations may have more than one set of solutions. Among those, the one minimizing both (29) and (30), corresponding to entropy terms $h(\mathbf{y}|\mathcal{H})$ and $h(\mathbf{y}|\mathbf{x}_s, \mathcal{H})$, respectively. In physics, this phenomenon is referred to as *phase coexistence* [40]. Note, however, that according to [41], one should expect the number of coexisting solutions to be finite since the phase space of a related problem in CDMA is known to be simple.

$$\begin{aligned} \frac{1}{N}I(\mathbf{y}; \mathbf{x}_s) &= \sum_{k=1}^K \frac{1}{\beta_{s,k}} I_{\text{su}} \left(z_{s,k}; x_{s,k} \mid \sqrt{\rho_{s,k} \xi_{s,k}} \right) + \sum_{\ell=1}^L \frac{1}{\beta_{i,\ell}} I_{\text{su}} \left(z_{i,\ell}; x_{i,\ell} \mid \sqrt{\rho_{i,\ell} \xi_{i,\ell}} \right) + \ln \left(1 + \sum_{k=1}^K \varepsilon_{s,k} + \sum_{\ell=1}^L \varepsilon_{i,\ell} \right) \\ &- \sum_{k=1}^K \frac{1}{\beta_{s,k}} \xi_{s,k} \varepsilon_{s,k} - \sum_{\ell=1}^L \frac{1}{\beta_{i,\ell}} \xi_{i,\ell} \varepsilon_{i,\ell} - \sum_{\ell=1}^L \frac{1}{\beta_{i,\ell}} I_{\text{su}} \left(\bar{z}_{i,\ell}; x_{i,\ell} \mid \sqrt{\rho_{i,\ell} \bar{\xi}_{i,\ell}} \right) - \ln \left(1 + \sum_{\ell=1}^L \bar{\varepsilon}_{i,\ell} \right) + \sum_{\ell=1}^L \frac{1}{\beta_{i,\ell}} \bar{\xi}_{i,\ell} \bar{\varepsilon}_{i,\ell} + \mathcal{O} \left(\frac{1}{N} \right), \end{aligned} \quad (32)$$

expression (29), corresponding to this term can therefore be derived following the lines of [33], where the sum-rate of a MIMO-MAC without interferers is considered. Hence, we omit the part of the proof related to (29) and present only the part related to (30) in the Appendix. \blacksquare

Here the mutual information terms in (29) and (30), as well as terms (31d), (31e) and (31e), are associated with two different fixed channels given by (8) with corresponding channel matrices $\mathbf{A}_{s,k}$, $\mathbf{A}_{i,\ell}$ and $\bar{\mathbf{A}}_{i,\ell}$. Hence, the terms $\varepsilon_{s,k}$, $\varepsilon_{i,\ell}$ and $\bar{\varepsilon}_{i,\ell}$ include the MMSE matrix of a fixed single-user channel discussed in Subsection II-C1, and transmit correlation of the original multiuser-MIMO channel (2). Despite looking a bit cumbersome, the two asymptotic expressions above have a simple interpretation. For instance, (29) represents the contributions of both users and interferers to the sum-rate of the MIMO-MAC presented in (1). Meanwhile, (30) describes the amount of information discarded at the receiver due to noise and interference removal.

Here we emphasize the difference from the MAC system studied in [33], where interferers were absent and white noise was the only source of disturbance. In contrast, our result, given in Proposition 1, describes the sum-rate of the MIMO-MAC in the presence of interference. In the case of a single user and a single interferer, both using Gaussian signals, the above result immediately reduces to the mean mutual information reported in [34], where it was obtained in a slightly different way. Note, though, that in contrast to [34], our result is not restricted to Gaussian channel inputs, while the authors there computed also higher moments of mutual information.

B. Uncorrelated Channels

The next result provides the sum-rate for the special case where $\mathbf{T}_{s,k}$, $\mathbf{T}_{i,\ell}$, $\mathbf{R}_{s,k}$ and $\mathbf{R}_{i,\ell}$ are all identities.

Corollary 1. *In the LSL, the asymptotic average sum-rate of an uncorrelated MIMO-MAC given in (4) in the presence of interference is given by (32) on the top of the page, where parameters $\varepsilon_{s,k}$, $\xi_{s,k}$, $\varepsilon_{i,\ell}$, $\xi_{i,\ell}$, $\bar{\varepsilon}_{i,\ell}$ and $\bar{\xi}_{i,\ell}$ satisfy the following set of fixed-point equations*

$$\xi_{s,k} = \beta_{s,k} \left(1 + \sum_{k=1}^K \varepsilon_{s,k} + \sum_{\ell=1}^L \varepsilon_{i,\ell} \right)^{-1}, \quad (33a)$$

$$\xi_{i,\ell} = \beta_{i,\ell} \left(1 + \sum_{k=1}^K \varepsilon_{s,k} + \sum_{\ell=1}^L \varepsilon_{i,\ell} \right)^{-1}, \quad (33b)$$

$$\varepsilon_{s,k} = \frac{\rho_{s,k}}{M_{s,k}} \text{tr} \left\{ \mathbf{mmse}_{\text{su}} \left(\mathbf{x}_{s,k}, \sqrt{\rho_{s,k} \xi_{s,k}} \mathbf{I}_{M_{s,k}} \right) \right\}, \quad (33c)$$

$$\varepsilon_{i,\ell} = \frac{\rho_{i,\ell}}{M_{i,\ell}} \text{tr} \left\{ \mathbf{mmse}_{\text{su}} \left(\mathbf{x}_{i,\ell}, \sqrt{\rho_{i,\ell} \xi_{i,\ell}} \mathbf{I}_{M_{i,\ell}} \right) \right\}, \quad (33d)$$

$$\bar{\xi}_{i,\ell} = \beta_{i,\ell} \left(1 + \sum_{\ell=1}^L \bar{\varepsilon}_{i,\ell} \right)^{-1}, \quad (33e)$$

$$\bar{\varepsilon}_{i,\ell} = \frac{\rho_{i,\ell}}{M_{i,\ell}} \text{tr} \left\{ \mathbf{mmse}_{\text{su}} \left(\mathbf{x}_{i,\ell}, \sqrt{\rho_{i,\ell} \bar{\xi}_{i,\ell}} \mathbf{I}_{M_{i,\ell}} \right) \right\}, \quad (33f)$$

and the mutual information terms are obtained using (20).

Proof: The proof follows directly from Proposition 1. The result was also reported in our previous work [35]. \blacksquare

IV. PRECODER OPTIMIZATION

As each transmitter has statistical CSI, by carefully choosing the precoder matrix, the transmitters could, in principle, maximize the sum mutual information between the inputs and outputs of their channels. The corresponding optimization problem is described as

$$\begin{aligned} \max_{\mathcal{G}_s} \quad & I(\mathbf{y}; \mathbf{x}_s) \\ \text{s.t.} \quad & \text{tr} \{ \mathbf{G}_{s,k} \mathbf{G}_{s,k}^H \} \leq M_{s,k}, \quad k \in \mathcal{K} \\ & \mathbf{G}_{s,k} \mathbf{G}_{s,k}^H \succeq \mathbf{0}_{M_{s,k}}, \quad k \in \mathcal{K}, \end{aligned} \quad (34)$$

where $\mathcal{K} \triangleq \{1, \dots, K\}$, $\mathcal{G}_s \triangleq \{ \mathbf{G}_{s,k}, \forall k \in \mathcal{K} \}$ and the objective function is given in (5). However, as mentioned in previous sections, working directly with the ergodic mutual information (5) is difficult. Thus, we next use the asymptotic results obtained in the previous section to simplify the optimization problem.

When examining Proposition 1, we see that the random parts of the channels, $\mathbf{W}_{s,k}$ and $\mathbf{W}_{i,\ell}$, play no role in the mutual information when the system is sufficiently large. Therefore, instead of the objective function (5), we maximize its asymptotic counterpart. Problem (34) then decouples into a set of individual per-transmitter optimization problems

$$\begin{aligned} \max_{\mathbf{G}_{s,k}} \quad & I_{\text{su}} \left(z_{s,k}; \mathbf{x}_{s,k} \mid \mathbf{A}_{s,k} \right) \\ \text{s.t.} \quad & \text{tr} \{ \mathbf{G}_{s,k} \mathbf{G}_{s,k}^H \} \leq M_{s,k}, \\ & \mathbf{G}_{s,k} \mathbf{G}_{s,k}^H \succeq \mathbf{0}_{M_{s,k}}, \end{aligned} \quad (35)$$

where $\mathbf{A}_{s,k} = \sqrt{\rho_{s,k} \xi_{s,k}} \mathbf{T}_{s,k}^{1/2}$. Namely, each transmitter k adjusts its own precoder matrix $\mathbf{G}_{s,k}$ according to its own transmit correlation matrix $\mathbf{T}_{s,k}$, which is available by the statistical-CSI assumption.

Note that here we need to maximize only the term h_s in (29) since h_i does not depend on the precoding matrices $\mathbf{G}_{s,k}$, $\forall k \in \mathcal{K}$. On the other hand, the parameters $\varepsilon_{s,k}$, $\xi_{s,k}$, $\forall k$, and $\varepsilon_{i,\ell}$, $\xi_{i,\ell}$, $\forall \ell$ do depend on the precoders in \mathcal{G}_s .

To obtain a feasible point satisfying the KKT conditions [42], we have to set to zero the derivatives of the objective w.r.t. the precoder matrices $\mathbf{G}_{s,k}$, given by

$$\begin{aligned} \frac{1}{N} \nabla_{\mathbf{G}_{s,k}} I(\mathbf{y}; \mathbf{x}_s) &= \nabla_{\mathbf{G}_{s,k}} h_s(\mathbf{G}_{s,k}) \\ &+ \sum_{j=1}^K \frac{\partial h_s}{\partial \varepsilon_{s,j}} \nabla_{\mathbf{G}_{s,k}} \varepsilon_{s,j}(\mathbf{G}_{s,k}) + \sum_{j=1}^K \frac{\partial h_s}{\partial \xi_{s,j}} \nabla_{\mathbf{G}_{s,k}} \xi_{s,j}(\mathbf{G}_{s,k}) \\ &+ \sum_{\ell=1}^L \frac{\partial h_s}{\partial \varepsilon_{i,\ell}} \nabla_{\mathbf{G}_{s,k}} \varepsilon_{i,\ell}(\mathbf{G}_{s,k}) + \sum_{\ell=1}^L \frac{\partial h_s}{\partial \xi_{i,\ell}} \nabla_{\mathbf{G}_{s,k}} \xi_{i,\ell}(\mathbf{G}_{s,k}). \end{aligned} \quad (36)$$

However, since h_s represents the free energy given by an equation similar to (66) in the Appendix, parameters $\varepsilon_{s,k}$, $\xi_{s,k}$, $\varepsilon_{i,\ell}$ and $\xi_{i,\ell}$ constitute its *saddle-point*. Therefore, the corresponding derivatives reduce to (see also [43] for further discussion)

$$\frac{\partial h_s}{\partial \varepsilon_{s,k}} = \frac{\partial h_s}{\partial \xi_{s,k}} = \frac{\partial h_s}{\partial \varepsilon_{i,\ell}} = \frac{\partial h_s}{\partial \xi_{i,\ell}} = 0, \quad \forall k, \ell \quad (37)$$

and it follows that

$$\nabla_{\mathbf{G}_{s,k}} I(\mathbf{y}; \mathbf{x}_s) = \nabla_{\mathbf{G}_{s,k}} I_{\text{su}}(\mathbf{z}_{s,k}; \mathbf{x}_{s,k} | \mathbf{A}_{s,k}), \quad \forall k. \quad (38)$$

Thus, for the sum-rate maximization we may consider $\varepsilon_{s,k}$, $\xi_{s,k}$, $\varepsilon_{i,\ell}$ and $\xi_{i,\ell}$ as being independent of \mathbf{G}_s .

Albeit optimization problem (35) is, in general, non-convex, it can still be efficiently solved for the following (most practically relevant) special cases.

A. Gaussian Inputs

In the case of Gaussian channel inputs, it is convenient to work with covariance matrices instead of precoders since the objective function of the optimization problem (35) reduces to

$$I_{\text{su}}(\mathbf{z}_{s,k}; \mathbf{x}_{s,k} | \mathbf{A}_{s,k}) = \ln \det \left(\mathbf{I}_M + \mathbf{A}_{s,k} \mathbf{P}_{s,k} \mathbf{A}_{s,k}^H \right). \quad (39)$$

Let the *singular-value decomposition* (SVD) of the effective fixed channel be given by $\mathbf{A}_{s,k} = \mathbf{U}_{\mathbf{A}_{s,k}} \boldsymbol{\Sigma}_{\mathbf{A}_{s,k}} \mathbf{V}_{\mathbf{A}_{s,k}}^H$, where $\mathbf{U}_{\mathbf{A}_{s,k}}$ and $\mathbf{V}_{\mathbf{A}_{s,k}}$ are orthonormal matrices and $\boldsymbol{\Sigma}_{\mathbf{A}_{s,k}} = \text{diag}([\sigma_1(\mathbf{A}_{s,k}), \dots, \sigma_{M_{s,k}}(\mathbf{A}_{s,k})]^T)$ is the matrix with the singular values on the diagonal. Given the solution to the fixed-point equations ($\xi_{s,k}$ and $\varepsilon_{s,k}$), the optimal input covariance matrix is then given by the *water-filling* solution [44]

$$\mathbf{P}_{s,k}^* = \mathbf{V}_{\mathbf{A}_{s,k}} \boldsymbol{\Sigma}_{\mathbf{P}_{s,k}} \mathbf{V}_{\mathbf{A}_{s,k}}^H, \quad (40)$$

where $\boldsymbol{\Sigma}_{\mathbf{P}_{s,k}}$ is a diagonal matrix whose non-zero entries are

$$[\boldsymbol{\Sigma}_{\mathbf{P}_{s,k}}]_{m,m} = \left[\frac{1}{\nu} - \frac{1}{\sigma_m(\mathbf{A}_{s,k})} \right]^+, \quad (41)$$

where ν is chosen so that the power constraint $\text{tr}\{\mathbf{P}_{s,k}\} = M_{s,k}$ is satisfied.

We remark here that, as pointed out in [14], to obtain the optimal transmit covariance matrix one has to iterate the solution to the fixed-point equations with the above statistical water-filling until the stopping criterion is reached.

B. Discrete Inputs

Unlike the previous case, finding the optimal precoder for discrete constellations is a difficult task. For such cases (35) is no longer a convex optimization problem. It has been shown in [45], [46] that the mutual information is a concave function in the quadratic form $\mathbf{F}_{s,k} \triangleq \mathbf{A}_{s,k} \mathbf{G}_{s,k} \mathbf{G}_{s,k}^H \mathbf{A}_{s,k}^H$; yet, due to the power constraint, $\text{tr}\{\mathbf{G}_{s,k} \mathbf{G}_{s,k}^H\} = M_{s,k}$, one cannot directly apply convex optimization tools for solving the problem. For instance, when using the *gradient ascent* method for solving (35), one updates $\mathbf{F}_{s,k}$ iteratively as

$$\mathbf{F}_{s,k}^{(l+1)} = \mathbf{F}_{s,k}^{(l)} + \mu_F \Delta \mathbf{F}_{s,k}, \quad (42)$$

with μ_F being the step size and $\Delta \mathbf{F}_{s,k}$ being the gradient of $I_{\text{su}}(\mathbf{z}_{s,k}; \mathbf{x}_{s,k} | \mathbf{A}_{s,k})$ w.r.t. $\mathbf{F}_{s,k}$. It is shown in [47] that the gradient of the mutual information in the single-user setup is

$$\nabla_{\mathbf{F}_{s,k}} I_{\text{su}}(\mathbf{z}_{s,k}; \mathbf{x}_{s,k} | \mathbf{A}_{s,k}) = \mathbf{E}_{s,k}, \quad (43)$$

where we denoted for notational convenience $\mathbf{E}_{s,k} \triangleq \mathbf{mmse}_{\text{su}}(\mathbf{x}_{s,k}, \mathbf{A}_{s,k})$ for the MMSE matrix defined in (10). Nonetheless, in practice, precoder matrix $\mathbf{G}_{s,k}$ is updated subject to a power constraint, which limits its feasible region and complicates the problem.

Here we apply an algorithm similar to that proposed in [45], [46], based on the alternating optimization between the following two subproblems.

1) *Per-Eigenmode Power Allocation*: Let the SVD of the precoder matrix be given by $\mathbf{G}_{s,k} = \mathbf{U}_{\mathbf{G}_{s,k}} \boldsymbol{\Sigma}_{\mathbf{G}_{s,k}} \mathbf{V}_{\mathbf{G}_{s,k}}^H$. For fixed $\mathbf{V}_{\mathbf{G}_{s,k}}$ the first subproblem is

$$\begin{aligned} \max_{\boldsymbol{\Sigma}_{\mathbf{G}_{s,k}}^2} \quad & I(\mathbf{z}_{s,k}; \mathbf{x}_{s,k} | \mathbf{A}_{s,k}, p_{\text{su}}) \\ \text{s.t.} \quad & \text{tr}\{\boldsymbol{\Sigma}_{\mathbf{G}_{s,k}}^2\} \leq M_{s,k}, \\ & \boldsymbol{\Sigma}_{\mathbf{G}_{s,k}}^2 \succeq \mathbf{0}_{M_{s,k}}. \end{aligned} \quad (44)$$

According to [48], this problem is convex, provided that the precoder has an optimal structure.

Since the matrix of interest, $\boldsymbol{\Sigma}_{\mathbf{G}_{s,k}}^2$, is a diagonal matrix, we introduce for notational convenience a vector \mathbf{g}_k , such that $\boldsymbol{\Sigma}_{\mathbf{G}_{s,k}}^2 = \text{diag}(\mathbf{g}_k)$. We then choose an initial value for \mathbf{g}_k , e.g., $\mathbf{g}_k^{(1)} = 1/M_{s,k} \mathbf{1}_{M_{s,k}}$, and perform the gradient update⁶

$$\mathbf{g}_k^{(l+1)} = \mathbf{g}_k^{(l)} + \mu_g \left(\text{Diag}(\boldsymbol{\Sigma}_{\mathbf{G}_{s,k}}^2 \mathbf{V}_{\mathbf{G}_{s,k}}^H \mathbf{E}_{s,k} \mathbf{V}_{\mathbf{G}_{s,k}}) - \gamma \mathbf{1}_{M_{s,k}} \right), \quad (45)$$

with $\gamma = 1/M_{s,k} \mathbf{1}_{M_{s,k}}^T \text{Diag}(\boldsymbol{\Sigma}_{\mathbf{G}_{s,k}}^2 \mathbf{V}_{\mathbf{G}_{s,k}}^H \mathbf{E}_{s,k} \mathbf{V}_{\mathbf{G}_{s,k}})$ and μ_g being an appropriately chosen step size, e.g., obtained by the *backtracking line search* algorithm [42]. If $\mathbf{g}_k^{(l+1)}$ has negative entries, one sets those to zero and renormalizes $\mathbf{g}_k^{(l+1)}$ so that the power constraint is satisfied and then sets $\boldsymbol{\Sigma}_{\mathbf{G}_{s,k}}^2 = \text{diag}(\mathbf{g}_k^{(l+1)})$.

2) *Optimization of the Eigenvectors of $\mathbf{F}_{s,k}$* : In this subproblem, for fixed $\boldsymbol{\Sigma}_{\mathbf{G}_{s,k}}^2$ we optimize the eigenvectors of the quadratic form $\mathbf{F}_{s,k} = \mathbf{A}_{s,k} \mathbf{G}_{s,k} \mathbf{G}_{s,k}^H \mathbf{A}_{s,k}^H$. Let $\boldsymbol{\Lambda}_{\mathbf{F},k}$ be the diagonal matrix, whose entries are eigenvalues of $\mathbf{F}_{s,k}$. The

⁶We remind the reader that $\text{Diag}(\mathbf{A})$ denotes a column vector containing the diagonal entries of matrix \mathbf{A} , whereas $\text{diag}(\mathbf{a})$ denotes a diagonal matrix containing the entries of vector \mathbf{a} .

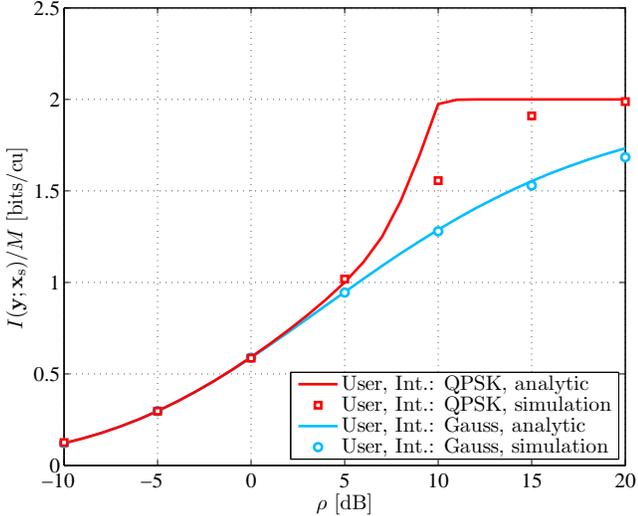


Fig. 2. Average mutual information per dimension vs. SNR for the single-user single-interferer scenario. Both, the user and interferer, have the same type of signaling. The terminals are equipped with $M = N = 4$ antennas. Solid curves denote analytic results, markers denote the results of Monte-Carlo simulation.

second subproblem is then formulated as

$$\begin{aligned} \max_{\mathbf{F}_{s,k}} \quad & I(\mathbf{z}_{s,k}; \mathbf{x}_{s,k} | \mathbf{A}_{s,k}, p_{\text{su}}) \\ \text{s.t.} \quad & \mathbf{\Lambda}_{\mathbf{F}_{s,k}} = \mathbf{\Sigma}_{\mathbf{G}_{s,k}}^2 \mathbf{\Sigma}_{\mathbf{A}_{s,k}}^2. \end{aligned} \quad (46)$$

The gradient of the mutual information is given by (43), and hence the gradient update for $\mathbf{F}_{s,k}$ becomes

$$\mathbf{F}_{s,k}^{(l+1)} = \mathbf{F}_{s,k}^{(l)} + \mu_{\mathbf{F}} \mathbf{E}_{s,k}. \quad (47)$$

The obtained update has to be further projected into a matrix with the prescribed eigenvalues, which is as close to $\mathbf{F}_{s,k}^{(l+1)}$ as possible [45], [46].

V. NUMERICAL RESULTS

In this section, we provide numerical results alongside with some discussion. For the simulations, the spatial correlation at the transmitter side is assumed to be generated by a uniform linear antenna array with *Gaussian power azimuth spectrum* [43]. Hence, correlation matrices ($\mathbf{T}_{s,k}$ and $\mathbf{T}_{i,\ell}$) consist of entries given by

$$[\mathbf{T}]_{a,b} = \frac{1}{2\pi\delta^2} \int_{-\pi}^{\pi} e^{2\pi j d_{\lambda}(a-b) \sin(\varphi) - \frac{(\varphi-\theta)^2}{2\delta^2}} d\varphi, \quad (48)$$

where d_{λ} is the nearest neighbor antenna spacing (in wavelengths λ), θ is the mean angle and δ^2 is the mean-square angle spread. For the sake of simplicity, we assume that there is no correlation at the receiver side, that is, $\mathbf{R}_{s,k} = \mathbf{I}_{M_{s,k}}$, $\forall k$ and $\mathbf{R}_{i,\ell} = \mathbf{I}_{M_{i,\ell}}$, $\forall \ell$.

A. Uncorrelated Channels

To begin with, we complement the obtained expression (32) for the uncorrelated case with Monte-Carlo simulations [49]. We consider the setup, where a single user transmits its signal towards the receiver in the presence of a single interferer.

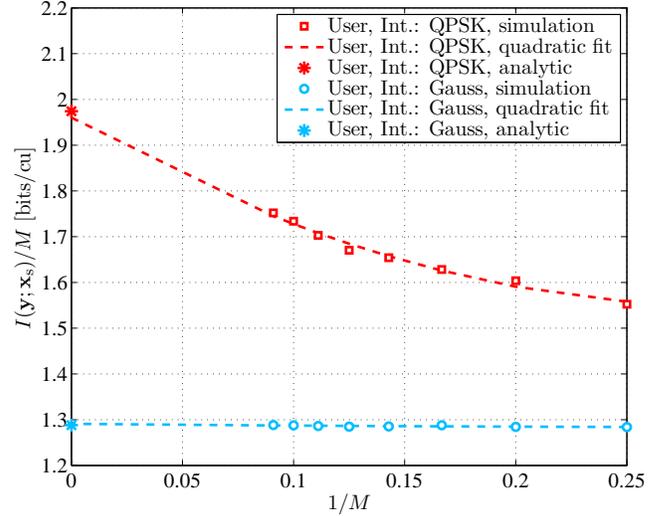


Fig. 3. Average mutual information per dimension vs. the inverse of the number of antennas $M = N \in \{4, \dots, 11\}$ at terminals for both Gaussian and QPSK signaling schemes at SNR $\rho = 10$ dB. The asterisk markers at $1/M \rightarrow 0$ denote the predictions obtained by the replica analysis in the LSL.

All terminals have equal numbers of antennas, that is, $N = M_s = M_i = M$. Both the user and interferer utilize the same type of signals (either Gaussian or QPSK), and have the same total transmit power, that is, $\rho_s = \rho_i = \rho$. In Fig. 2, we plot the average mutual information per transmit antenna in bits per channel use (cu) as a function of SNR. Both the asymptotic results obtained *via* the replica method and Monte Carlo simulations for $M = 4$ antennas are shown. For QPSK, the simulations and asymptotic results are the farthest apart at SNRs around $\rho = 10$ dB due to the *phase transition* phenomenon. Namely, in this region the system instantly switches from one state to another, mimicking the “water-ice” transition in physics [7]. For Gaussian inputs, the plotted curve does not experience a phase transition and the asymptotic results are accurate already for small numbers of antennas.

To illustrate how the small scale simulations converge to the asymptotic result obtained using the replica method, Fig. 3 plots the simulated values of the mutual information (5) vs. $1/M$ for $M \in \{4, \dots, 11\}$ at $\rho = 10$ dB. The markers at $1/M = 0$ represent the results obtained using Corollary 1 and quadratic curves are fitted to the simulated data using non-linear least-squares regression. From the extrapolation we observe that the simulated per-antenna mutual information approaches close to the value predicted by the replica analysis also in the region nearby the phase transition.

Next, we consider the effect of signal constellations on the achievable sum-rate. Fig. 4 depicts the mutual information of the desired user when the Gaussian and/or QPSK signaling are used by the terminals. The interference-free case is also drawn for comparison. We directly see that for the desired user it is always best to employ Gaussian signaling. On the other hand, Gaussian signaling, when used by the interferer, creates more disturbance. Hence, in a cellular system where inter-cell interference is present, the network might be able

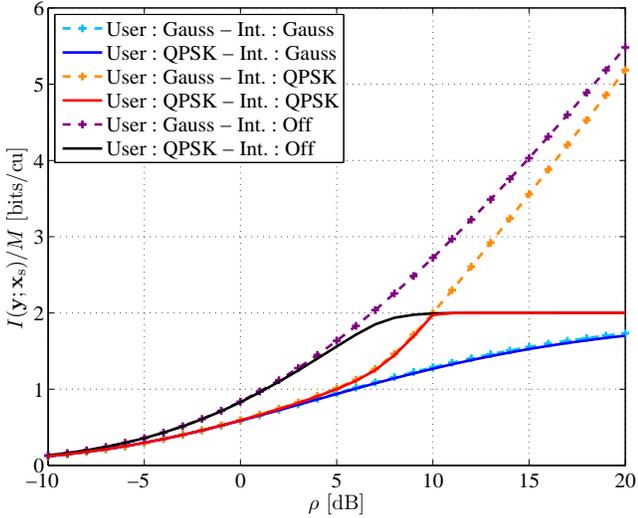


Fig. 4. Average mutual information per dimension vs. SNR for different combinations of user's and interferer's signaling schemes. The terminals are equipped with $M = N = 4$ antennas.

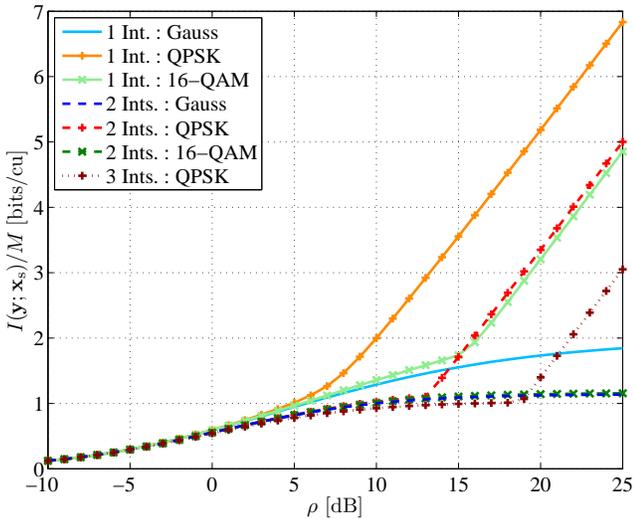


Fig. 5. Average mutual information per dimension vs. SNR for a single user with Gaussian signaling and $L \in \{1, 2, 3\}$ interferers using different signaling schemes. The terminals have $M = N = 4$ antennas.

to handle more users if some of them are assigned discrete constellations. This is due to the fact that the most severe (unoptimized) interference is in fact Gaussian⁷ [51].

Fig. 5 presents the average mutual information per transmit antenna of a single user in the presence of $L \in \{1, 2, 3\}$ interferers using different signaling schemes (Gaussian, QPSK and 16-QAM). We see that $L = 2$ interferers using QPSK constellations create roughly the same performance degradation as a single interferer with 16-QAM signaling at high SNR. On the other hand, at SNR higher than 20 dB, $L = 3$

⁷Note that here we do not consider the optimization of the interferer's signal constellation aiming to jam the user. In the latter case, Gaussian signaling would not cause the worst-case interference, whereas an optimized discrete signaling would degrade the user's performance the most [50].

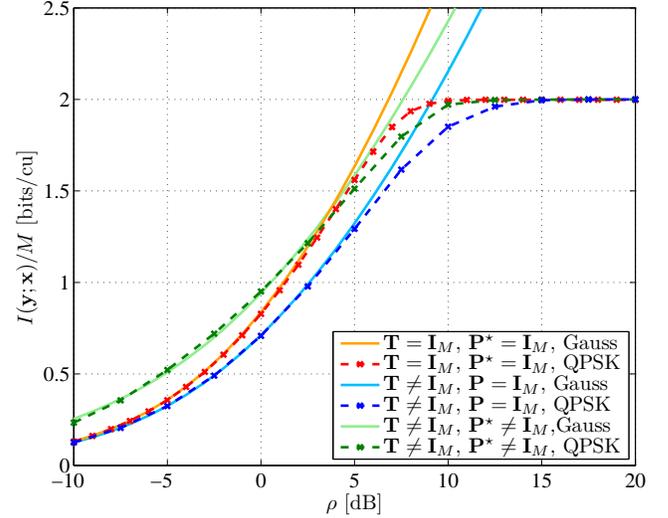


Fig. 6. Average mutual information per dimension vs. SNR for a single-user system without interference. Both correlated ($T \neq I_M$) and i.i.d. ($T = I_M$) MIMO channels with ($P \neq I_M$) and without ($P = I_M$) precoding are considered. Gaussian or QPSK signaling is employed by the terminals that each are equipped with $M = N = 3$ antennas.

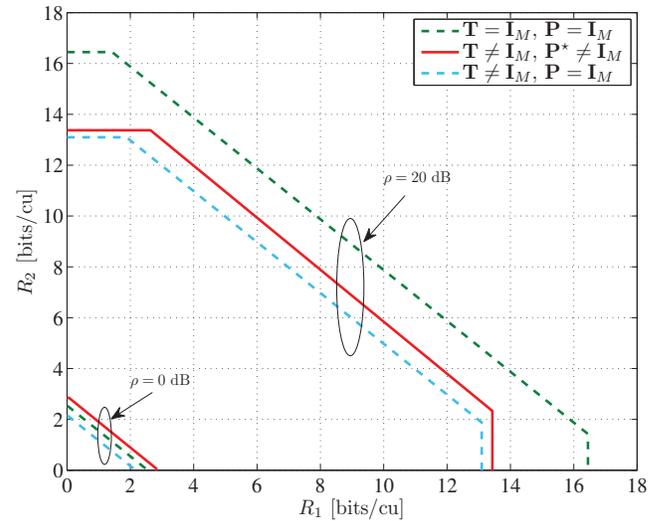


Fig. 7. Achievable rate region for the 2-user MIMO-MAC under a power constraint of $\rho \in \{0, 20\}$ dB. Both correlated ($T \neq I_M$) and i.i.d. ($T = I_M$) channels with ($P \neq I_M$) and without ($P = I_M$) precoding are considered. The terminals have $M = N = 3$ antennas.

interferers with QPSK create smaller performance degradation than $L = 2$ interferers with 16-QAM. Again, we see that Gaussian signaling causes the worst-case degradation in the desired user's performance.

B. Correlated Channels

In this section, we study the behavior of the system under spatial correlation and quantify the gains of precoding. Fig. 6 depicts the normalized ergodic mutual information given Gaussian and QPSK inputs as a function of SNR of a single-user (no interference) MIMO channel with $N = M_s = M = 3$ antennas under various conditions. Namely, we consider the cases of correlated and uncorrelated channels with and without

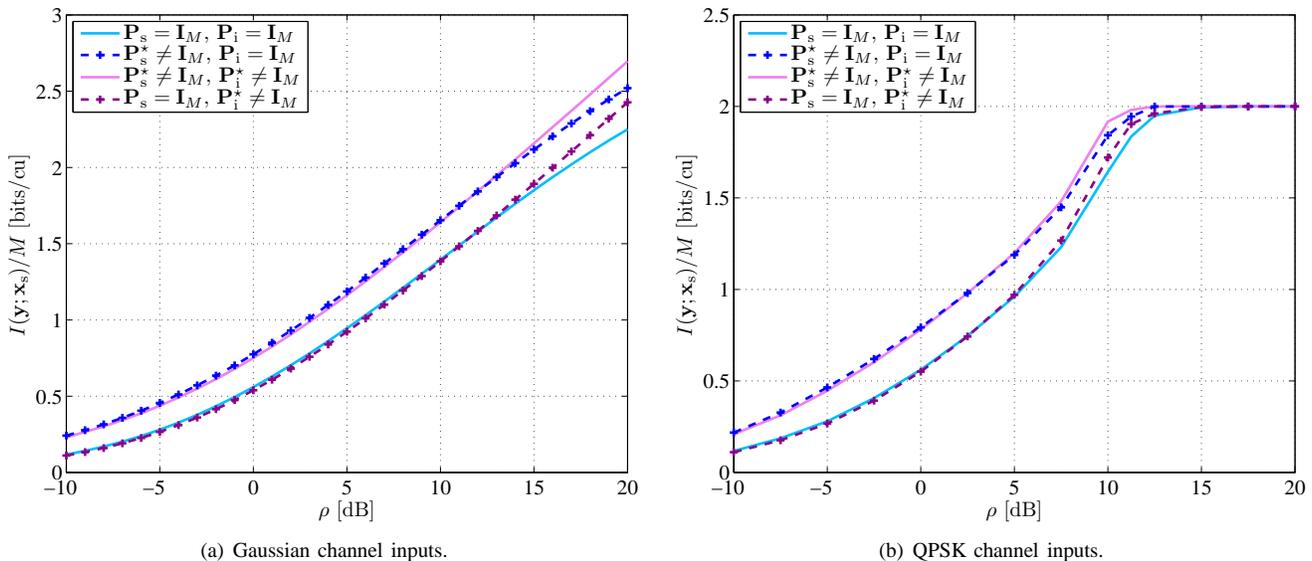


Fig. 8. Average mutual information per dimension vs. SNR for a single-user correlated MIMO channel with ($\mathbf{P}_s \neq \mathbf{I}_M$) and without ($\mathbf{P}_s = \mathbf{I}_M$) precoding. A single interferer using precoded ($\mathbf{P}_i \neq \mathbf{I}_M$) or isotropic ($\mathbf{P}_i = \mathbf{I}_M$) channel inputs is present. The terminals have $M = N = 3$ antennas.

precoding at the transmitter. The transmit side correlation parameters are set to the terminals as follows. The antenna spacing is set to $d_\lambda = 1$, the mean angle is set to $\theta = 0^\circ$ and the root-mean-square angle spread is chosen to be $\delta = 5^\circ$. The receive side correlation matrix is set to identity, *i.e.*, $\mathbf{R} = \mathbf{I}_N$. As expected, at low SNR the curves representing the two constellations coincide. Moreover, for the case of Gaussian inputs, transmit correlation decreases the achievable rate at high SNR regardless of precoding. Quite remarkably though, at low SNR transmit correlation together with precoding based on the statistical water-filling (40) is beneficial in terms of the mutual information. Somewhat similar behavior is observed for the case of QPSK signals. At low SNR a precoder in combination with transmitter-side correlation allows to improve the system performance as compared to the case of an uncorrelated channel. However, since in this case the per-stream mutual information saturates at 2 bits/cu at high SNR, transmit correlation does not affect the rates too much in this region. To optimize the precoder matrix for the case of QPSK signals, we have used the algorithm described in Subsection IV-B.

Next, we investigate the performance of a correlated MIMO-MAC. The rate region of a generic K -user MIMO-MAC using Gaussian signaling is given by [52]

$$\mathcal{C}_{\text{MAC}} = \bigcup_{\substack{\text{tr}\{\mathbf{P}_{s,k}\} \leq M_{s,k} \\ \mathbf{P}_{s,k} \succeq \mathbf{0}_{M_{s,k}} \\ \forall k \in \mathcal{K}}} \left\{ \{R_k\}, \forall k \in \mathcal{K} : \right. \\ \left. \sum_{i \in \mathcal{S}} R_i \leq \ln \det \left(\mathbf{I}_N + \sum_{i \in \mathcal{S}} \mathbf{H}_{s,i} \mathbf{P}_{s,i} \mathbf{H}_{s,i}^H \right), \forall \mathcal{S} \subset \mathcal{K} \right\}. \quad (49)$$

Note that the corresponding large-system ergodic mutual information terms can be directly obtained from Proposition 1. To illustrate this region, we consider a symmetric setup with

two users who both have $M = M_{s,1} = M_{s,2} = 3$ antennas. We further fix the available transmit powers $\rho = \rho_{s,1} = \rho_{s,2}$ to $\rho \in \{0, 20\}$ dB and evaluate the achievable rate regions for the given 2-user MAC. The result is depicted in Fig. 7 where both uncorrelated and correlated channels with and without precoding are present. It is clear that using precoding at both terminals is beneficial when transmit correlation is present. As expected though, at high SNR the rate region is largest for the uncorrelated MAC. On the contrary, at low SNR the largest rate region is achieved in the presence of correlation and optimal precoding.

To finish this section, we return to the case of one desired user and add an interferer with $M_i = 3$ antennas, having the same transmit power, $\rho_i = \rho$, and same correlation parameters. Fig. 8 depicts the average mutual information as a function of SNR for this scenario under Gaussian and QPSK signaling schemes. Both the user and interferer either do or do not realize precoding. Note that the scenario is symmetric and hence the precoders used by the terminals are the same. Moreover, the terminals adapt to their own correlation matrices aiming to increase their own rates. From Fig. 8(a) we see that, quite expectedly, for the case of Gaussian signals, precoding at the user increases its own ergodic rate. At the same time, we also see that utilizing the optimal precoder at the interferer results in higher rate at the user's terminal at high SNR, degrading the performance of the latter in the low-SNR region only slightly. This observation falls along the lines of [53], where spatially colored noise was shown to be less harmful than white Gaussian noise. Interestingly, similar behavior is observed for the case of QPSK inputs (cf. Fig. 8(b)), apart from expected saturation at 2 bits/cu at high SNR.

VI. CONCLUSIONS

In this paper, we derived an explicit expression for the asymptotic achievable sum-rate of the MIMO multiple-access channel in the presence of interference. The result accounts

for the spatial correlation at the terminals and, in contrast to the previous results, is not restricted to Gaussian signals. Although derived in the large system limit, it approximates relatively well the achievable sum-rate of small systems. We have also studied the impact of the number of interferers, their signaling scheme and spatial correlation structure on the system performance. For instance, Gaussian signaling is seen to create the worst-case (unoptimized) interference. Thus, the system may handle more interferers if they use discrete signal constellation, as compared to the case of Gaussian interferers. The obtained large-system approximation has been further used to find precoder matrices for maximizing the sum-rate for both Gaussian and finite-alphabet signaling schemes. It has been demonstrated that properly optimized precoder significantly increases the achievable rates. More interestingly, in the low-SNR region the presence of spatial correlation, in combination with an optimal precoder, is beneficial and can in fact improve the system performance as compared to the case of uncorrelated channels. The proposed approach is general and degenerates to many well-known results as special cases.

APPENDIX

In general, direct computation of (6b) is very difficult if we allow arbitrary channel input distributions. To overcome this obstacle we use the replica method to compute the entropy in the LSL. To stay coherent with the existing work, we partially keep the statistical physics terminology, avoiding unnecessary jargon whenever possible.

Let us define the *partition function* related to (6b) as

$$Z(\mathbf{y}, \mathbf{x}_s, \mathcal{H}) \triangleq \mathbf{E}_{\mathbf{x}_i} \left\{ \frac{1}{\pi^N} e^{-\|\mathbf{y} - \mathbf{H}_s \mathbf{x}_s - \mathbf{H}_i \mathbf{x}_i\|^2} \right\}. \quad (50)$$

In statistical physics, virtually all interesting macroscopic quantities can be derived from the partition function of the system. Often, however, it is more convenient to work with the logarithm of the partition function, or *free energy*, instead. If we further average the (normalized) free energy w.r.t. the remaining randomness in the MIMO setup (50), we get

$$\mathcal{F} \triangleq -\frac{1}{N} \mathbf{E}_{\mathbf{y}, \mathbf{x}_s, \mathcal{H}} \ln Z(\mathbf{y}, \mathbf{x}_s, \mathcal{H}), \quad (51)$$

that is, just the normalized equivocation term, $\frac{1}{N} h(\mathbf{y} | \mathbf{x}_s, \mathcal{H})$. Then, we take the first step towards making the evaluation of \mathcal{F} solvable by writing

$$\mathcal{F} = -\frac{1}{N} \lim_{u \rightarrow 0^+} \frac{\partial}{\partial u} \ln \mathbf{E}_{\mathbf{y}, \mathbf{x}_s, \mathcal{H}} \{ Z^u(\mathbf{y}, \mathbf{x}_s, \mathcal{H}) \}, \quad (52)$$

and implicitly assuming that the system size also grows large as discussed in Section III. This identity is exact when u is a real number, but on its own it does not solve the problem. Thus, we invoke the *replica trick* and write the under-log term as

$$\begin{aligned} & \mathbf{E}_{\mathbf{y}, \mathbf{x}_s, \mathcal{H}} \{ Z^u(\mathbf{y}, \mathbf{x}_s, \mathcal{H}) \} \\ &= \mathbf{E}_{\mathbf{x}_s, \mathbf{X}_i, \mathcal{H}} \left\{ \int \frac{1}{\pi^N} \prod_{a=0}^u e^{-\|\mathbf{y} - \mathbf{H}_s \mathbf{x}_s - \mathbf{H}_i \mathbf{x}_i^{(a)}\|^2} d\mathbf{y} \right\}, \quad (53) \end{aligned}$$

where $\mathbf{x}_i^{(a)}$ is the a th replica of the signal vector transmitted

by the interferers. Its distribution, $p(\mathbf{x}_i^{(a)})$, is identical to $p(\mathbf{x}_i)$ and conditionally independent for $a \in \{0, 1, \dots, u\}$ given \mathbf{y} and \mathcal{H} . For ease of exposition, we have also defined the vector $\mathbf{X}_i \triangleq [\mathbf{x}_i^{(0)\top}, \dots, \mathbf{x}_i^{(u)\top}]^\top \in \mathbb{C}^{LM_i(u+1)}$ that contains the replicated interferers' signals. Note that since the partition function (50) has an expectation over \mathbf{x}_i and not \mathbf{x}_s , we introduce only the replicas related to the former.

After applying the replica trick, the problem of finding the free energy reduces to evaluating $\mathbf{E}_{\mathbf{y}, \mathbf{x}_s, \mathcal{H}} \{ Z^u(\mathbf{y}, \mathbf{x}_s, \mathcal{H}) \}$ for integer-valued u using techniques from large deviations theory and then assuming that the result generalizes to real positive values, at least in the vicinity of zero⁸.

Let us now define the following set of random vectors

$$\mathbf{v}_{s,k} \triangleq \sqrt{\frac{\rho_{s,k}}{M_{s,k}}} \mathbf{H}_{s,k} \mathbf{x}_{s,k} \in \mathbb{C}^N, \quad (54a)$$

$$\mathbf{v}_{i,\ell}^{(a)} \triangleq \sqrt{\frac{\rho_{i,\ell}}{M_{i,\ell}}} \mathbf{H}_{i,\ell} \mathbf{x}_{i,\ell}^{(a)} \in \mathbb{C}^N. \quad (54b)$$

Denote also $\mathbf{v}_s \triangleq \sum_{k=1}^K \mathbf{v}_{s,k}$ and $\mathbf{v}_i^{(a)} \triangleq \sum_{\ell=1}^L \mathbf{v}_{i,\ell}^{(a)}$, and group them into a concatenated vector

$$\mathbf{V} \triangleq [\mathbf{v}_s^\top + \mathbf{v}_i^{(0)\top}, \dots, \mathbf{v}_s^\top + \mathbf{v}_i^{(u)\top}]^\top \in \mathbb{C}^{N(u+1)}. \quad (55)$$

Conditioned on the interferers' signals \mathbf{X}_i , we know by the central limit theorem that as the dimensions of the channel matrices $\mathbf{H}_{s,k}$ and $\mathbf{H}_{i,\ell}$ grow large, \mathbf{V} converges to a zero-mean Gaussian random vector with conditional covariance

$$\mathbf{Q} = \sum_{k=1}^K (\mathbf{Q}_{s,k} \otimes \mathbf{R}_{s,k}) + \sum_{\ell=1}^L (\mathbf{Q}_{i,\ell} \otimes \mathbf{R}_{i,\ell}). \quad (56)$$

The auxiliary matrix $\mathbf{Q}_{i,\ell}$ has entries

$$[\mathbf{Q}_{i,\ell}]_{a,b} = \frac{\rho_{i,\ell}}{M_{i,\ell}} \mathbf{x}_{i,k}^{(b)\top} \mathbf{T}_{i,\ell} \mathbf{x}_{i,\ell}^{(a)}, \quad (57)$$

for $a, b \in \{0, 1, \dots, u\}$, while $\mathbf{Q}_{s,k} = q_{s,k} \mathbf{1}_{u+1} \mathbf{1}_{u+1}^\top$ with

$$q_{s,k} = \frac{\rho_{s,k}}{M_s} \mathbf{x}_{s,k}^\top \mathbf{T}_{s,k} \mathbf{x}_{s,k}. \quad (58)$$

Note that here we used (2) and the assumption that $\mathbf{W}_{s,k}$ and $\mathbf{W}_{i,\ell}$ have i.i.d. CSCG entries of unit variance to derive the result.

Let us define $\mathcal{Q} \triangleq \{(\mathbf{Q}_{s,k}, \mathbf{Q}_{i,\ell}), \forall k, \ell\}$, so that the expectation over replicated vectors \mathbf{X}_i may be rewritten as an integral over a probability measure of \mathcal{Q} . Treating \mathbf{V} as a Gaussian random vector, it can be shown *via* the Edgeworth expansion that in the LSL [30]

$$\mathbf{E}_{\mathbf{y}, \mathbf{x}_s, \mathcal{H}} \{ Z^u(\mathbf{y}, \mathbf{x}_s, \mathcal{H}) \} = \int e^{G^{(u)}(\mathcal{Q})} d\mu^{(u)}(\mathcal{Q}), \quad (59)$$

where we have omitted constant terms and $\mu^{(u)}(\mathcal{Q})$ reads

$$\mu^{(u)}(\mathcal{Q}) = \mathbf{E} \left\{ \prod_{k=1}^K \mathbb{1} \left(\rho_{s,k} \mathbf{x}_{s,k}^\top \mathbf{T}_{s,k} \mathbf{x}_{s,k} - M_s q_{s,k} \right) \right\}$$

⁸Note that mathematical rigor of this step is still an open problem. However, some results obtained by the replica method are confirmed to match the ones derived *via* systematic approaches (e.g., [54], [40]). Moreover, the results can be further verified *via* Monte-Carlo simulations, as we saw in Section V. Therefore, we regard the replica analysis as a valid mathematical tool.

$$T_1^{(u)}(\mathcal{Q}, \tilde{\mathcal{Q}}) = u \ln \det \left(\mathbf{I}_{N(u+1)} + \sum_{\ell=1}^L (p_{i,\ell} - q_{i,\ell}) \mathbf{R}_{i,\ell} \right), \quad (67a)$$

$$T_2^{(u)}(\mathcal{Q}, \tilde{\mathcal{Q}}) = \sum_{k=1}^K M_{s,k} (u+1) (\tilde{p}_{s,k} + u \tilde{q}_{s,k}) q_{s,k} + \sum_{\ell=1}^L M_{i,\ell} (u+1) (\tilde{p}_{i,\ell} p_{i,\ell} + u \tilde{q}_{i,\ell} q_{i,\ell}), \quad (67b)$$

$$T_3^{(u)}(\mathcal{Q}, \tilde{\mathcal{Q}}) = \sum_{k=1}^K M_{s,k} \ln \left(1 - u(u+1) \bar{\mathbf{A}}_{s,k}^2 \right) - \sum_{\ell=1}^L \ln \mathbf{E}_{\mathbf{X}_i} \left\{ \exp \left(\left\| \sum_{a=0}^u \bar{\mathbf{A}}_{i,\ell} \mathbf{x}_{i,\ell}^{(a)} \right\|^2 - \sum_{a=0}^u \mathbf{x}_{i,\ell}^{(a)\text{H}} (\bar{\mathbf{A}}_{i,\ell}^2 - \bar{\mathbf{B}}_{i,\ell}) \mathbf{x}_{i,\ell}^{(a)} \right) \right\} \quad (67c)$$

$$\times \prod_{a,b=0}^u \prod_{\ell=1}^L \mathbb{1} \left(\rho_{i,\ell} \mathbf{x}_{i,k}^{(b)\text{H}} \mathbf{T}_{i,\ell} \mathbf{x}_{i,\ell}^{(a)} - M_i[\mathcal{Q}_{i,\ell}]_{a,b} \right) \Big\}. \quad (60)$$

with the above expectation being w.r.t. $\{\mathbf{X}_i, \mathcal{H}\}$. If we plug \mathbf{V} into (53) and assess the expectations w.r.t. \mathbf{V} and \mathbf{y} using Gaussian integration, we get

$$G^{(u)}(\mathcal{Q}) = -Nu \ln \pi - N \ln(u+1) - \ln \det \left[\mathbf{I}_{N(u+1)} + \sum_{k=1}^K \mathbf{Q}_{s,k} \boldsymbol{\Sigma} \otimes \mathbf{R}_{s,k} + \sum_{\ell=1}^L \mathbf{Q}_{i,\ell} \boldsymbol{\Sigma} \otimes \mathbf{R}_{i,\ell} \right], \quad (61)$$

where $\boldsymbol{\Sigma} \triangleq \mathbf{I}_{u+1} - \frac{1}{u+1} \mathbf{1}_{u+1} \mathbf{1}_{u+1}^\text{T} \in \mathbb{R}^{(u+1) \times (u+1)}$.

To compute the integral in (59), we note that since both $\mathbf{Q}_{s,k}$ and $\mathbf{Q}_{i,\ell}$ are formed by summing independent random variables, measure (60) satisfies the large deviations property and by Varadhan's theorem [55]

$$\frac{1}{N} \ln \mathbf{E}_{\mathbf{y}, \mathbf{x}_s, \mathcal{H}} \{ Z^u(\mathbf{y}, \mathbf{x}_s, \mathcal{H}) \} - \frac{1}{N} \max_{\mathcal{Q}} \left(G^{(u)}(\mathcal{Q}) - I^{(u)}(\mathcal{Q}) \right) \rightarrow 0, \quad (62)$$

in the LSL. The second term inside the maximization is referred to as the rate function and can be obtained via Cramér's theorem [55]

$$I^{(u)}(\mathcal{Q}) = \max_{\tilde{\mathcal{Q}}} \left\{ \sum_{k=1}^K M_{s,k} \text{tr} \{ \tilde{\mathbf{Q}}_{s,k} \mathbf{Q}_{s,k} \} + \sum_{\ell=1}^L M_{i,\ell} \text{tr} \{ \tilde{\mathbf{Q}}_{i,\ell} \mathbf{Q}_{i,\ell} \} - \ln M^{(u)}(\tilde{\mathcal{Q}}) \right\}, \quad (63)$$

where the moment-generating function of $\mu^{(u)}(\mathcal{Q}_s, \mathcal{Q}_i)$ reads

$$M^{(u)}(\tilde{\mathcal{Q}}) = \mathbf{E}_{\mathbf{X}_i} \left\{ \prod_{k=1}^K e^{\rho_{s,k} \mathbf{X}_{s,k}^\text{H} (\tilde{\mathbf{Q}}_{s,k} \otimes \mathbf{T}_{s,k}) \mathbf{X}_{s,k}} \times \prod_{\ell=1}^L e^{\rho_{i,\ell} \mathbf{X}_{i,\ell}^\text{H} (\tilde{\mathbf{Q}}_{i,\ell} \otimes \mathbf{T}_{i,\ell}) \mathbf{X}_{i,\ell}} \right\}, \quad (64)$$

and we denoted $\mathbf{X}_{s,k} \triangleq [\mathbf{x}_{s,k}^\text{T}, \dots, \mathbf{x}_{s,k}^\text{T}]^\text{T} \in \mathbb{C}^{M_{s,k}(u+1)}$, $\mathbf{X}_{i,\ell} \triangleq [\mathbf{x}_{i,\ell}^{(0)\text{T}}, \dots, \mathbf{x}_{i,\ell}^{(u)\text{T}}]^\text{T} \in \mathbb{C}^{M_{i,\ell}(u+1)}$. As before, we group the auxiliary "Q-matrices" as $\tilde{\mathcal{Q}} \triangleq \{(\tilde{\mathbf{Q}}_{s,k}, \tilde{\mathbf{Q}}_{i,\ell}), \forall k, \ell\}$.

To make the optimization problems in (62) and (63) tractable, we next assume that the saddle-point solutions are

invariant under the permutation of the replica indices. This is known as the *replica symmetric* (RS) ansatz⁹ and here it implies that we can write the members of \mathcal{Q} and $\tilde{\mathcal{Q}}$ as

$$\tilde{\mathbf{Q}}_{s,k} = \tilde{q}_{s,k} \mathbf{1}_{u+1} \mathbf{1}_{u+1}^\text{T} + (\tilde{p}_{s,k} - \tilde{q}_{s,k}) \mathbf{I}_{u+1}, \quad (65a)$$

$$\mathbf{Q}_{i,\ell} = q_{i,\ell} \mathbf{1}_{u+1} \mathbf{1}_{u+1}^\text{T} + (p_{i,\ell} - q_{i,\ell}) \mathbf{I}_{u+1}, \quad (65b)$$

$$\tilde{\mathbf{Q}}_{i,\ell} = \tilde{q}_{i,\ell} \mathbf{1}_{u+1} \mathbf{1}_{u+1}^\text{T} + (\tilde{p}_{i,\ell} - \tilde{q}_{i,\ell}) \mathbf{I}_{u+1}. \quad (65c)$$

Under the RS assumption, the free energy in the LSL becomes

$$\mathcal{F} = 1 + \ln \pi + \frac{1}{N} \lim_{u \rightarrow 0^+} \frac{\partial}{\partial u} \min_{\mathcal{Q}} \max_{\tilde{\mathcal{Q}}} \sum_{i=1}^3 T_i^{(u)}(\mathcal{Q}, \tilde{\mathcal{Q}}), \quad (66)$$

where the terms of $T_i^{(u)}(\mathcal{Q}, \tilde{\mathcal{Q}})$ are given in (67), on the top of the page, where we denoted $\bar{\mathbf{A}}_{s,k} \triangleq \sqrt{\rho_{s,k} \tilde{q}_{s,k}} \mathbf{T}_{s,k}^{1/2}$, $\bar{\mathbf{A}}_{i,\ell} \triangleq \sqrt{\rho_{i,\ell} \tilde{q}_{i,\ell}} \mathbf{T}_{i,\ell}^{1/2}$ and $\bar{\mathbf{B}}_{i,\ell} \triangleq \rho_{i,\ell} \tilde{p}_{i,\ell} \mathbf{T}_{i,\ell}$. We also assumed there that all terminals have independent channel inputs. Performing the Hubbard-Stratonovich transform [58], [59] on (67c), we decouple the quadratic terms

$$\begin{aligned} T_3^{(u)}(\mathcal{Q}, \tilde{\mathcal{Q}}) &= \sum_{k=1}^K M_{s,k} \ln \det \left(1 - u(u+1) \bar{\mathbf{A}}_{s,k}^2 \right) \\ &\quad - \sum_{\ell=1}^L \ln \frac{1}{\pi M_{i,\ell}} \int \mathbf{E}_{\mathbf{x}_{i,\ell}} \left\{ e^{-\|\bar{\mathbf{z}}_{i,\ell} - \bar{\mathbf{A}}_{i,\ell} \mathbf{x}_{i,\ell}\|^2} e^{\mathbf{x}_{i,\ell}^\text{H} \bar{\mathbf{B}}_{i,\ell} \mathbf{x}_{i,\ell}} \right\} \\ &\quad \times \left[\mathbf{E}_{\mathbf{x}_{i,\ell}} \left\{ e^{2\text{Re}\{\bar{\mathbf{z}}_{i,\ell}^\text{H} \bar{\mathbf{A}}_{i,\ell} \mathbf{x}_{i,\ell}\} - \mathbf{x}_{i,\ell}^\text{H} (\bar{\mathbf{A}}_{i,\ell}^2 - \bar{\mathbf{B}}_{i,\ell}) \mathbf{x}_{i,\ell}} \right\} \right]^u d\mathbf{z}_{i,\ell}, \end{aligned} \quad (68)$$

where $\bar{\mathbf{z}}_{i,\ell} \in \mathbb{C}^{M_{i,\ell}}$ is an auxiliary variable.

To solve (66), we find the conditions under which the derivatives of the argument w.r.t. all the RS parameters vanish. After taking $u \rightarrow 0$, we get

$$\tilde{p}_{s,k} = \tilde{q}_{s,k} = \tilde{p}_{i,\ell} = 0, \quad \forall k, \ell, \quad (69a)$$

$$\tilde{q}_{i,\ell} = \frac{\rho_{i,\ell}}{M_{i,\ell}} \text{tr} \{ \mathbf{S}^{-1} \mathbf{R}_{i,\ell} \}, \quad (69b)$$

$$p_{i,\ell} - q_{i,\ell} = \frac{1}{M_{i,\ell}} \text{tr} \{ \text{mmse}_{\text{su}}(\mathbf{x}_{i,\ell}, \bar{\mathbf{A}}_{i,\ell}) \mathbf{T}_{i,\ell} \}, \quad (69c)$$

⁹This assumption has been widely accepted in the field of statistical physics [40] and information theory [7], [31], [33]. However, the cases of *replica-symmetry breaking* are known in the literature [56], [57].

where we wrote for notational convenience

$$\mathbf{S} \triangleq \mathbf{I}_N + \sum_{\ell=1}^L (p_{i,\ell} - q_{i,\ell}) \mathbf{R}_{i,\ell}, \quad (70)$$

and the MMSE matrix is defined in (10). Finally, taking the derivative w.r.t. u and letting $u \rightarrow 0$ we get

$$\begin{aligned} \mathcal{F} = & \ln \det \mathbf{S} + \sum_{\ell=1}^L I_{\text{su}}(\bar{\mathbf{z}}_{i,\ell}; \mathbf{x}_{i,\ell} | \bar{\mathbf{A}}_{i,\ell}) \\ & + \sum_{\ell=1}^L M_{i,\ell} \tilde{q}_{i,\ell} (q_{i,\ell} - p_{i,\ell}) + 1 + \ln \pi, \end{aligned} \quad (71)$$

where we used the fact that $p_{i,\ell} = \rho_{i,\ell}$. Denoting $\bar{\xi}_{i,\ell} \triangleq \tilde{q}_{i,\ell}$ and $\bar{\varepsilon}_{i,\ell} \triangleq p_{i,\ell} - q_{i,\ell}$, we obtain h_i from (30), as well as a system of fixed-point equations given by (31f) and (31c).

ACKNOWLEDGEMENT

The authors thank the anonymous reviewers for their suggestions that have greatly improved the quality of the manuscript. In addition, the authors are grateful to Peter Larsson for the discussions on practical considerations regarding the scenario investigated here.

REFERENCES

- [1] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Commun.*, vol. 6, no. 3, pp. 311–335, Mar. 1998.
- [2] E. Telatar, "Capacity of multi-antenna Gaussian channels," *Europ. Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, Nov. 1999.
- [3] K. Werner, H. Asplund, D. V. Figueiredo, N. Jaldén, and B. Halvarsson, "LTE-advanced 8 × 8 MIMO measurements in an indoor scenario," in *Proc. Int. Symp. Antennas and Propagation (ISAP)*, Nanjing, China, May 2012, pp. 750–753.
- [4] C. Oikonomopoulos-Zachos, T. Ould, and M. Arnold, "Outdoor channel characterization of MIMO-LTE antenna configurations through measurements," in *Proc. IEEE Vehic. Tech. Conf. (VTC Spring)*, Yokohama, Japan, May 2012, pp. 1–4.
- [5] R. S. Blum, "MIMO capacity with interference," *IEEE J. Select. Areas Commun.*, vol. 21, no. 5, pp. 793–801, June 2003.
- [6] A. M. Tulino, A. Lozano, and S. Verdú, "Impact of antenna correlation on the capacity of multiantenna channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2491–2509, 2005.
- [7] R. R. Müller, "Channel capacity and minimum probability of error in large dual antenna array systems with binary modulation," *IEEE Trans. Signal Process.*, vol. 51, no. 11, pp. 2821–2828, Nov. 2003.
- [8] A. J. Grant and P. D. Alexander, "Random sequence multisets for synchronous code-division multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 44, no. 7, pp. 2832–2836, Mar. 1998.
- [9] S. Verdú and S. Shamai, "Spectral efficiency of CDMA with random spreading," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 622–640, Dec. 1999.
- [10] D. N. C. Tse and S. V. Hanly, "Linear multiuser receivers: Effective interference, effective bandwidth and user capacity," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 641–657, Mar. 1999.
- [11] P. B. Rapajic and D. Popescu, "Information capacity of a random signature multiple-input multiple-output channel," *IEEE Trans. Commun.*, vol. 48, no. 8, pp. 1245–1248, Aug. 2000.
- [12] J. Evans and D. N. C. Tse, "Large system performance of linear multiuser receivers in multipath fading channels," *IEEE Trans. Inf. Theory*, vol. 46, no. 6, pp. 2059–2078, Sep. 2000.
- [13] 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.
- [14] R. Couillet, M. Debbah, and J. W. Silverstein, "A deterministic equivalent for the analysis of correlated MIMO multiple access channels," *IEEE Trans. Inf. Theory*, vol. 57, no. 6, pp. 3493–3514, Jun. 2011.
- [15] A. Lozano and A. M. Tulino, "Capacity of multiple-transmit multiple-receive antenna architectures," *IEEE Trans. Inf. Theory*, vol. 48, no. 12, pp. 3117–3128, Dec. 2002.
- [16] M. Kang, L. Yang, and M.-S. Alouini, "Capacity of MIMO channels in the presence of co-channel interference," *Wirel. Commun. Mob. Comput.*, vol. 7, no. 1, pp. 113–125, 2007.
- [17] Y. Wang and D.-W. Yue, "Capacity of MIMO Rayleigh fading channels in the presence of interference and receive correlation," *IEEE Trans. Veh. Technol.*, vol. 58, no. 8, pp. 4398–4405, Oct. 2009.
- [18] M. Chiani, M. Z. Win, and H. Shin, "MIMO networks: The effects of interference," *IEEE Trans. Inf. Theory*, vol. 56, no. 1, pp. 336–349, Jan. 2010.
- [19] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.
- [20] H. Huh, G. Caire, H. Papadopoulos, and S. Ramprasad, "Achieving "massive MIMO" spectral efficiency with a not-so-large number of antennas," *IEEE Trans. Wireless Commun.*, vol. 11, no. 9, pp. 3226–3239, Sep. 2012.
- [21] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.
- [22] E. G. Larsson, F. Tufvesson, O. Edfors, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *arXiv preprint arXiv:1304.6690*, 2013.
- [23] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?" *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 160–171, Feb. 2013.
- [24] R. R. Müller, M. Vehkaperä, and L. Cottatellucci, "Blind pilot decontamination," in *Proc. ITG Workshop on Smart Antennas (WSA)*, Stuttgart, Germany, Mar. 2013.
- [25] J. Zhang, C. K. Wen, S. Jin, X. Gao, and K. K. Wong, "On capacity of large-scale mimo multiple access channels with distributed sets of correlated antennas," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 133–148, Feb. 2013.
- [26] M. Kac, "On certain Toeplitz-like matrices and their relation to the problem of lattice vibrations," in *Ark. Fys. Seminar i Trondheim*, Trondheim, Norway, Nov. 1968, pp. 1–21.
- [27] S. F. Edwards and P. W. Anderson, "Theory of spin glasses," *J. Phys. S. F. Metal Phys.*, vol. 5, pp. 965–974, May 1975.
- [28] D. Sherrington and S. Kirkpatrick, "Solvable model of a spin-glass," *Phys. Rev. Lett.*, vol. 35, pp. 1792–1796, Dec. 1975.
- [29] T. Tanaka, "Statistical mechanics of CDMA multiuser demodulation," *Europhys. Lett.*, vol. 54, no. 4, pp. 540–546, May 2001.
- [30] —, "A statistical-mechanics approach to large-system analysis of CDMA multiuser detectors," *IEEE Trans. Inf. Theory*, vol. 48, no. 11, pp. 2888–2910, 2002.
- [31] D. Guo and S. Verdú, "Randomly spread CDMA: Asymptotics via statistical physics," *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 1983–2010, June 2005.
- [32] K. Takeda, A. Hatabu, and Y. Kabashima, "Statistical mechanical analysis of the linear vector channel in digital communication," *J. Phys. A*, vol. 40, no. 47, p. 14085, 2007.
- [33] C.-K. Wen and K.-K. Wong, "Asymptotic analysis of spatially correlated MIMO multiple-access channels with arbitrary signaling inputs for joint and separate decoding," *IEEE Trans. Inf. Theory*, vol. 53, no. 1, pp. 252–268, Jan. 2007.
- [34] A. L. Moustakas, S. H. Simon, and A. M. Sengupta, "MIMO capacity through correlated channels in the presence of correlated interferers and noise: A (not so) large N analysis," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2545–2561, Oct. 2003.
- [35] M. A. Girnyk, M. Vehkaperä, and L. K. Rasmussen, "On the asymptotic sum-rate of uplink MIMO cellular systems in the presence of non-Gaussian inter-cell interference," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Anaheim, U.S.A., Dec. 2012.
- [36] D. Aktas, M. N. Bacha, J. S. Evans, and S. V. Hanly, "Scaling results on the sum capacity of cellular networks with MIMO links," *IEEE Trans. Inf. Theory*, vol. 52, no. 7, pp. 3264–3274, Jul. 2006.
- [37] D. Chizhik, F. Rashid-Farrokhi, J. Ling, and A. Lozano, "Effect of antenna separation on the capacity of BLAST in correlated channels," *IEEE Commun. Lett.*, vol. 4, no. 11, pp. 337–339, Nov. 2000.
- [38] A. Lozano, A. M. Tulino, and S. Verdú, "Optimum power allocation for parallel Gaussian channels with arbitrary input distributions," *IEEE Trans. Inf. Theory*, vol. 52, no. 7, pp. 3033–3051, Jul. 2006.

- [39] D. Guo, S. Shamai, and S. Verdú, "Mutual information and minimum mean-square error in Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1261–1282, Apr. 2005.
- [40] H. Nishimori, *Statistical physics of spin glasses and information processing: An introduction*. Clarendon Press, 2001, vol. 111.
- [41] —, "Comment on 'Statistical mechanics of CDMA multiuser demodulation'," *Europhys. Lett.*, vol. 57, no. 2, pp. 302–303, 2002.
- [42] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge University Press, 2004.
- [43] C.-K. Wen, K.-K. Wong, and C. T. Ng, "On the asymptotic properties of amplify-and-forward MIMO relay channels," *IEEE Trans. Commun.*, vol. 59, no. 2, pp. 590–602, 2011.
- [44] T. M. Cover and J. A. Thomas, *Elements of information theory*, 2nd, Ed. Wiley-Interscience, 2006.
- [45] M. Lamarca, "Linear precoding for mutual information maximization in MIMO systems," in *Proc. IEEE Int. Symp. Wireless Commun. Systems (ISWCS)*, Siena, Italy, Sep. 2009, pp. 26–30.
- [46] C. Xiao, Y. R. Zheng, and Z. Ding, "Globally optimal linear precoders for finite alphabet signals over complex vector Gaussian channels," *IEEE Trans. Signal Process.*, vol. 59, no. 7, pp. 3301–3314, Jul. 2011.
- [47] D. P. Palomar and S. Verdú, "Gradient of mutual information in linear vector Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 1, pp. 141–154, Jan. 2006.
- [48] M. Payaró and D. P. Palomar, "Hessian and concavity of mutual information, differential entropy, and entropy power in linear vector Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 8, pp. 3613–3628, Aug. 2009.
- [49] W. Tranter, K. Shanmugan, T. Rappaport, and K. Kosbar, *Principles of communication systems simulation with wireless applications*. Prentice Hall Press, 2003.
- [50] S. Shamai and S. Verdú, "Worst-case power-constrained noise for binary-input channels," *IEEE Trans. Inf. Theory*, vol. 38, no. 5, pp. 1494–1511, Sep. 1992.
- [51] C. Cahn, "Worst interference for coherent binary channel," *IEEE Trans. Inf. Theory*, vol. 17, no. 2, pp. 209–210, Mar. 1971.
- [52] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 5, pp. 684–702, Jun. 2003.
- [53] A. Lozano, A. M. Tulino, and S. Verdú, "Multiple-antenna capacity in the low-power regime," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2527–2544, 2003.
- [54] W. Hachem, O. Khorunzhiy, P. Loubaton, J. Najim, and L. Pastur, "A new approach for mutual information analysis of large dimensional multi-antenna channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 9, pp. 3987–4004, Sep. 2008.
- [55] R. Ellis, *Entropy, large deviations, and statistical mechanics*. Springer, 2005, vol. 1431, no. 821.
- [56] B. M. Zaidel, R. R. Müller, A. L. Moustakas, and R. de Miguel, "Vector precoding for Gaussian MIMO broadcast channels: Impact of replica symmetry breaking," *IEEE Trans. Inf. Theory*, vol. 58, no. 3, pp. 1413–1440, Mar. 2012.
- [57] R. R. Müller, D. Guo, and A. L. Moustakas, "Vector precoding for wireless MIMO systems and its replica analysis," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 3, pp. 530–540, Apr. 2008.
- [58] R. L. Stratonovich, "On a method of calculating quantum distribution functions," in *Dokl. Sov. Akad. Nauk.*, vol. 115, no. 6, 1957, pp. 1097–1100.
- [59] J. Hubbard, "Calculation of partition functions," *Phys. Rev. Lett.*, vol. 3, no. 2, p. 77, 1959.