On the Ergodic Capacity of Correlated Rician Fading MIMO Channels With Interference

Giorgio Taricco, *Fellow, IEEE*, and Erwin Riegler, *Member, IEEE*

***Abstract*—An asymptotic approach to derive the ergodic capacity achieving covariance matrix for a multiple-input multiple-output (MIMO) channel is presented. The method is applicable to MIMO channels affected by separately correlated Rician fading and co-channel interference. It is assumed that the number of transmit, receive and interfering antennas grows asymptotically while their ratios, as well as the SNR and the SIR, approach finite constants. Nevertheless, it is shown that the asymptotic results represent an accurate approximation in the case of a finitely many antennas and can be used to derive the ergodic channel capacity. This is accomplished by using an iterative power allocation algorithm based on a *water-filling* approach. The convergence of a similar algorithm (nicknamed *frozen water-filling*) was conjectured in a work by Dumont *et al.* Here, we show that, in the Rayleigh case, the frozen water-filling algorithm may not converge while, in those cases, our proposed algorithm converges. Finally, numerical results are included in order to assess the accuracy of the asymptotic method proposed, which is compared to equivalent results obtained via Monte-Carlo simulations.**

***Index Terms*—Co-channel interference, multiple-input multiple-output (MIMO) system capacity, random matrices, replica method, Rician fading channels.**

I. Introduction

ONE of the key motivations for the use of multiple antennas in communication systems was the promise to achieve a capacity scaling proportional to the number of antennas employed. This is the core of the early seminal works by Winters [46], Telatar [40], and Foschini and Gans [16], [17]. Specifically, the problem of capacity evaluation was addressed and solved completely for the ergodic and independent Rayleigh fading multiple-input-multiple-output (MIMO) channel.

One basic assumption in the early literature on this subject was the presence of a very large number of propagation paths driving the transmitted signals to the receive antennas and this physical condition was referred to by the name of *rich scattering*. The rich scattering assumption implied that the MIMO channel could be modeled by a set of independent and Rayleigh-distributed gains, whereby independence was referred to as lack of spatial correlation. The spatially-uncorrelated Rayleigh fading MIMO channel was then recognized to achieve, in the asymptotic SNR regime, an ergodic capacity which scaled proportionally to the minimum number of transmit and receive antennas.

Further studies showed that the rich scattering assumption was not always realistic and spatial correlation came often into play [20]. They showed that antenna (or spatial) correlation changes drastically with the scattering environment, the distance between the transmitter and the receiver, the antenna configurations and the Doppler spread [1], [36]. Chua *et al.* [9] showed that spatial correlation reduces the growth rate of channel capacity either in the presence of perfect channel state information at the receiver (CSIR) and perfect channel state information at the transmitter (CSIT), or in the presence of perfect CSIR only (with iid power allocation at the transmitter). Another physical condition affecting negatively the channel capacity is the reduction of the rank of the channel matrix. In certain conditions, channels with very low spatial correlation can exhibit a *keyhole* effect, which may reduce the channel matrix rank significantly and limit consequently the available capacity gain [8]. Nevertheless, this effect is not prevalent in most environments [19].

The derivation of MIMO channel capacity with and without correlation has been addressed extensively in the literature and an excellent account of recent results addressing the Rayleigh fading channel model (with or without correlation) is given in [19]. This paper discusses the effects of channel state information at the receiver and the transmitter (referred to as CSIR and CSIT, respectively) and of channel distribution information at the receiver and the transmitter (referred to as CDIR and CDIT, respectively). An important result in this area, relevant to the case of separately correlated Rayleigh fading channels, consists in the fact that the eigenvectors of the capacity achieving input covariance matrix coincide with those of the transmit correlation matrix [24], [44]. The eigenvalues of the capacity achieving input covariance matrix can be found by numerical optimization techniques [26].

Another effect having an important impact on the MIMO channel capacity is the presence of line-of-sight (LOS) propagation which leads to consider a spatially correlated Rician fading model. Many experimental and theoretical studies have shown that, in order to encompass all channel characteristics, correlated Rician fading models have to be considered [18], [19], [33].

Nevertheless, the Rician fading channel model received more limited attention in the literature, as far as concerns its capacity, with some notable exceptions, which are reported in the following.

Manuscript received June 07, 2008; revised September 29, 2010; accepted February 04, 2011. Date of current version June 22, 2011. This work was supported by the STREP project No. IST-026905 (MASCOT) within the sixth framework programme of the European Commission.

G. Taricco is with the Dipartimento di Elettronica, Politecnico di Torino, 10129, Torino, Italy (e-mail: giorgio.taricco@polito.it).

E. Riegler is with the Institute of Telecommunications, Vienna University of Technology, A-1040 Vienna, Austria (e-mail: eriegler@nt.tuwien.ac.at).

Communicated by A. J. Grant, Associate Editor for Communications.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2011.2146610

The multiple-input single-output (MISO) case was considered by Visotsky and Madhow in [44]. Their results were extended to the MIMO case by Venkatesan *et al.* in [43] under the assumption of unit-rank average channel matrix and further refined by Jafar and Goldsmith in [24] for an average channel matrix of arbitrary rank. All these works focused on uncorrelated Rician fading case and showed that the eigenvectors of the optimal input covariance matrix are the right-singular vectors of the average channel matrix.

More recently, Hoesli *et al.* [22] studied the dependence of channel capacity on the average channel matrix. Their work showed that the mutual information increases monotonically (for a fixed covariance matrix, not a fixed SNR) with the Lowner partial order of the singular values of the average channel matrix. Jayaweera and Poor [25] obtained the channel matrix singular value distribution of a Rician fading channel with receive-side correlation only. Their result was based on the use of Bessel functions of matrix argument. They showed that the mutual information decreases monotonically as the Rician factor increases for any given SNR. Tulino *et al.* [42] derived the eigenvectors of the ergodic capacity achieving covariance matrix for several MIMO channel models of interest and suggest a numerical algorithm to obtain the eigenvalues.

Finally, another effect has been shown to affect the capacity of a MIMO channel, sometimes in an unpredictable way: multiple access interference (MAI). It is known that MAI has a considerable impact on the achievable information rate and its *Gaussian approximation* (i.e., considering it as additive Gaussian noise in the derivation of the channel capacity) has been shown to produce unduly pessimistic results [2], [6], [7].

All the methods reported above are based on analytic derivations and sometimes they lead to very involved expressions that are valid in less general cases than the one we address here. In particular, it is worth noting that none of these methods allows one to obtain the ergodic capacity achieving covariance matrix for the case of separately correlated Rician fading with interference, which is the main contribution of this work.

An alternative to exact analytic calculation of the mutual information is based on the assumption that the number of transmit and receive antennas is asymptotically large. This method consists in deriving the asymptotic expansion of the probability distribution and has been applied several times in the literature. An outstanding example of this is provided by Moustakas *et al.* [31], who considered the case of separately-correlated Rayleigh fading with interference. They used the *replica method* to show that the asymptotic distribution of the mutual information is Gaussian and calculated its asymptotic mean and variance. These results have been extended in [37], [38] to the separately-correlated Rician fading case without interference and, more recently, with interference [39].

In the same asymptotic contest, Dumont *et al.* proposed in [13], [14] an asymptotic method (based on the Stieltjes transform) to derive the ergodic capacity achieving covariance matrix for the separately correlated Rician fading MIMO channel. The key ingredient is the asymptotic ergodic mutual information formula, which the authors refer to [21], is equivalent to the expression derived in [38] via the replica method. The authors compared their results against those obtained by Vu and Paulraj [45], who used an interior point with barrier optimization method and showed that the asymptotic method has a considerable advantage in terms of time efficiency. Another recent asymptotic result has been reported in [34] and concerns the case of Rayleigh fading with separate spatial correlation. The accuracy of this result when the number of antennas is finite is compared with the accuracy of our proposed method in Section IV-B.

In this work we consider the case of separately-correlated Rician fading with narrowband interference and calculate the corresponding MIMO channel ergodic capacity. Our work is an extension of previous results presented in the conference paper [39] and provides a streamlined analysis of the problem with detailed algorithm implementation and additional numerical examples. Among numerical examples, we compare (for validation purposes) our results with those obtained numerically by Vu and Paulraj [45] in the interference-free case. Then, we extend our analysis to other MIMO channels affected by interference and assess the effect of covariance optimization against iid power allocation (iid transmitted symbols).

The remainder of this paper is organized as follows. Section II describes the model of the separately-correlated Rician fading MIMO channel with a co-channel interfering multiple-antenna and multiple-access source. Key system parameters are defined (SNR, SIR). The mutual information is derived and its asymptotic approximation is given by using results from[37], [38]. Section III considers the problem of evaluating the ergodic capacity by use of the asymptotic approximation obtained. An iterative water-filling algorithm is derived, whose details are reported in Appendix A. The convergence of this algorithm is discussed in Appendix B for the Rayleigh case. Section III contains also an approximate analysis of capacity based on another approximation stemming from Jensen's inequality. Section IV presents a number of numerical examples aimed to assess the goodness of the proposed approximation and to analyze a selection of communication scenarios of interest. Finally, Section V summarizes the results presented in this paper and provides concluding remarks.

A. Notation

We denote (column-) vectors and matrices by lowercase and uppercase boldface characters, respectively. The Hermitian transpose of a matrix $\mathbf{A}$ is $\mathbf{A}^{\mathsf{H}}$. The trace of a matrix is $\mathsf{Tr}(\mathbf{A}) = \sum_a (\mathbf{A})_{aa}$. The Frobenius norm of a matrix $\mathbf{A}$ is $\|\mathbf{A}\|$; its square can be written as $\|\mathbf{A}\|^2 = \mathsf{Tr}(\mathbf{A}\mathbf{A}^{\mathsf{H}})$. The Kronecker product of two matrices $\mathbf{A}$ and $\mathbf{B}$ is $\mathbf{A} \otimes \mathbf{B}$; $\mathbf{A}^{1/2}$ is the *matrix square-root* of the Hermitian positive semidefinite matrix $\mathbf{A}$ and is defined as $\mathbf{U}\mathbf{\Lambda}^{1/2}\mathbf{U}^{\mathsf{H}}$ where $\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^{\mathsf{H}}$ is the orthogonal factorization of the positive semidefinite matrix $\mathbf{A}$ [23, p. 414]. $(\mathbf{A})_+$ denotes the element-wise positive part of the matrix $\mathbf{A}$, where the positive part of a scalar x is defined by $(x)_+ \triangleq \max\{0, x\}$. The notation $\mathbf{x} \sim \mathcal{N}_c(\boldsymbol{\mu}, \mathbf{R})$ defines a vector of complex jointly circularly-symmetric Gaussian random variables with mean value $\boldsymbol{\mu} = \mathsf{E}[\mathbf{x}]$ and covariance matrix $\mathbf{R} = \mathsf{E}[\mathbf{x}\mathbf{x}^{\mathsf{H}}] - \boldsymbol{\mu}\boldsymbol{\mu}^{\mathsf{H}}$ and its joint probability density function (pdf) is given by

$$f(\mathbf{x}) = \det(\pi\mathbf{R})^{-1} \exp[-(\mathbf{x}-\boldsymbol{\mu})^{\mathsf{H}}\mathbf{R}^{-1}(\mathbf{x}-\boldsymbol{\mu})].$$

II. System Model and Basic Results

We consider a narrowband block fading channel with n_R receive antennas, n_T transmit antennas from an intended user and n_I transmit antennas from an interfering source. The channel is specified by the following equation:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{H}_I\mathbf{x}_I + \mathbf{z}. \tag{1}$$

Here, $\mathbf{x} \in \mathbb{C}^{n_T \times 1}$ is the transmitted signal vector, $\mathbf{x}_I \in \mathbb{C}^{n_I \times 1}$ is the interfering signal vector, $\mathbf{H} \in \mathbb{C}^{n_R \times n_T}$ is the signal channel matrix, $\mathbf{H}_I \in \mathbb{C}^{n_R \times n_I}$ is the interference channel matrix, $\mathbf{z} \in \mathbb{C}^{n_R \times 1}$ is the additive noise vector and $\mathbf{y} \in \mathbb{C}^{n_R \times 1}$ is the received signal vector. Both $\mathbf{x}$ and $\mathbf{x}_I$ are assumed to have zero mean.

We assume that the additive noise vector has zero mean and covariance matrix $\mathbf{Q}_Z = \mathsf{E}[\mathbf{z}\mathbf{z}^{\mathsf{H}}]$.

The channel matrices $\mathbf{H}$ and $\mathbf{H}_I$ are assumed to be of separately (or Kronecker) correlated Rician fading type with common receive correlation matrices. Thus, they can be written as

$$\mathbf{H} = \bar{\mathbf{H}} + \mathbf{R}^{1/2}\mathbf{W}\mathbf{T}^{1/2}$$

and

$$\mathbf{H}_I = \bar{\mathbf{H}}_I + \mathbf{R}^{1/2}\mathbf{W}_I\mathbf{T}_I^{1/2}$$

where $\bar{\mathbf{H}}$ and $\bar{\mathbf{H}}_I$ represent the average channel matrices related to the presence of a line-of-sight (LOS) signal component in the multipath fading channel, the Hermitian positive definite matrices $\mathbf{R}, \mathbf{T}, \mathbf{T}_I$ are the receive and transmit (signal and interference) correlation matrices and $\mathbf{W}$ and $\mathbf{W}_I$ have iid $\mathcal{N}_c(0,1)$ entries.

We define the signal and interference covariance matrices as $\mathbf{Q}$ and $\mathbf{Q}_I$, respectively. In the following we assume that the interference covariance matrix is kept fixed and optimization is carried out on the signal covariance matrix $\mathbf{Q}$, under the power constraint

$$\mathsf{Tr}(\mathbf{Q}) \le P.$$

Following standard conventions [15], [37], we define the Rician factors as

$$K \triangleq \frac{\|\bar{\mathbf{H}}\|^2}{\mathsf{Tr}(\mathbf{T})\mathsf{Tr}(\mathbf{R})} \quad \text{and} \quad K_I \triangleq \frac{\|\bar{\mathbf{H}}_I\|^2}{\mathsf{Tr}(\mathbf{T}_I)\mathsf{Tr}(\mathbf{R})}. \tag{2}$$

We also define the signal-to-noise power ratio (SNR), the interference-to-noise power ratio (INR) and the signal-to-interference power ratio (SIR) at the receiver as

$$\begin{aligned}
\mathsf{SNR} &\triangleq \frac{(K+1)\mathsf{Tr}(\mathbf{T})\mathsf{Tr}(\mathbf{R})\mathsf{Tr}\left(\mathbf{Q}/n_T\right)}{\mathsf{Tr}(\mathbf{Q}_Z)} \\
\mathsf{INR} &\triangleq \frac{(K_I+1)\mathsf{Tr}(\mathbf{T}_I)\mathsf{Tr}(\mathbf{R})\mathsf{Tr}\left(\mathbf{Q}_I/n_I\right)}{\mathsf{Tr}(\mathbf{Q}_Z)} \\
\mathsf{SIR} &\triangleq \frac{\mathsf{SNR}}{\mathsf{INR}} = \frac{(K+1)\mathsf{Tr}(\mathbf{T})\mathsf{Tr}\left(\mathbf{Q}/n_T\right)}{(K_I+1)\mathsf{Tr}(\mathbf{T}_I)\mathsf{Tr}\left(\mathbf{Q}_I/n_I\right)}.
\end{aligned} \tag{3}$$

Remark II.1: These power ratios coincide with the corresponding *received* power ratios when $\mathbf{Q} = \left(\frac{P}{n_T}\right)\mathbf{I}_{n_T}$ (iid transmitted symbols). This assumption is required because power allocation optimization is based on a constraint on the transmitted power which would not be equivalent to a constraint on the SNR unless this is proportional to the transmitted power. Thus, we enforce this proportionality by defining it under iid power allocation (even though this is not necessarily the case considered).

A. Mutual Information

Following well known results from Cover and Thomas [10], we can write the (random) mutual information for a given channel realization as follows:

$$\begin{aligned}
I_{\mathbf{H},\mathbf{H}_I}(\mathbf{x};\mathbf{y}) &= h(\mathbf{y}) - h(\mathbf{y} \mid \mathbf{x}) \\
&= \ln\det(\mathbf{H}\mathbf{Q}\mathbf{H}^{\mathsf{H}} + \mathbf{H}_I\mathbf{Q}_I\mathbf{H}_I^{\mathsf{H}} + \mathbf{Q}_Z) \\
&\quad - \ln\det(\mathbf{H}_I\mathbf{Q}_I\mathbf{H}_I^{\mathsf{H}} + \mathbf{Q}_Z)
\end{aligned} \tag{4}$$

nat/complex dimension. Then, the ergodic capacity under the power constraint $\mathsf{Tr}(\mathbf{Q}) \le P$ is given by the following expression:

$$C = \max_{\mathsf{Tr}(\mathbf{Q}) \le P} \mathsf{E}[I_{\mathbf{H},\mathbf{H}_I}(\mathbf{x};\mathbf{y})]. \tag{5}$$

In order to calculate the capacity (5) we resort to a recent asymptotic method, allowing to approximate closely the average mutual information of a separately-correlated Rician MIMO channel when the number of transmit and receive antennas grows asymptotically large [37], [38] while their ratios approach finite constants.

Summarizing the main results from [37], [38],[1] the ergodic mutual information of a separately correlated Rician fading MIMO channel whose channel matrix is given by $\mathbf{H} = \bar{\mathbf{H}} + \mathbf{R}^{1/2}\mathbf{W}\mathbf{T}^{1/2}$ (where $\mathbf{W}$ has iid $\mathcal{N}_c(0,1)$ distributed entries) and the noise and signal covariance matrices are $\mathbf{Q}_Z$ and $\mathbf{Q}$, respectively, can be asymptotically approximated by the following expression:

$$\mathsf{E}[I_{\mathbf{H}}(\mathbf{x};\mathbf{y})] \sim \mu_I(\bar{\mathbf{H}}, \mathbf{R}, \mathbf{T}, \mathbf{Q}_Z, \mathbf{Q}) \tag{6}$$

nat/complex dimension. In this expression, we defined

$$\begin{aligned}
&\mu_I(\bar{\mathbf{H}}, \mathbf{R}, \mathbf{T}, \mathbf{Q}_Z, \mathbf{Q}) \\
&\quad \triangleq \ln\det\begin{pmatrix} \mathbf{Q}_Z + w\mathbf{R} & \tilde{\mathbf{H}} \\ -\tilde{\mathbf{H}}^{\mathsf{H}} & \mathbf{I}_{n_T} + z\tilde{\mathbf{T}} \end{pmatrix} - wz - \ln\det(\mathbf{Q}_Z) \\
&\quad = \ln\det(\mathbf{Q}_Z + w\mathbf{R}) - wz - \ln\det(\mathbf{Q}_Z) \\
&\qquad + \ln\det[\mathbf{I}_{n_T} + z\tilde{\mathbf{T}} + \tilde{\mathbf{H}}^{\mathsf{H}}(\mathbf{Q}_Z + w\mathbf{R})^{-1}\tilde{\mathbf{H}}]
\end{aligned} \tag{7}$$

where we defined

$$\tilde{\mathbf{T}} \triangleq \mathbf{Q}^{1/2}\mathbf{T}\mathbf{Q}^{1/2} \text{ and } \tilde{\mathbf{H}} \triangleq \bar{\mathbf{H}}\mathbf{Q}^{1/2}. \tag{8}$$

The parameters w, z satisfy the following equations:[2]

$$\begin{cases} w = \mathsf{Tr}\left\{[\mathbf{I}_{n_T} + z\tilde{\mathbf{T}} + \tilde{\mathbf{H}}^{\mathsf{H}}(\mathbf{Q}_Z + w\mathbf{R})^{-1}\tilde{\mathbf{H}}]^{-1}\tilde{\mathbf{T}}\right\} \\ z = \mathsf{Tr}\left\{[\mathbf{I}_{n_R} + w\mathbf{R} + \tilde{\mathbf{H}}(\mathbf{I}_{n_T} + z\tilde{\mathbf{T}})^{-1}\tilde{\mathbf{H}}^{\mathsf{H}}]^{-1}\mathbf{R}\right\}. \end{cases} \tag{9}$$

[1]Notice that here we generalize the case considered in [37], [38] by assuming a general noise covariance matrix, namely, $\mathbf{Q}_Z$.

[2]These recurrence equations are different from [37]–[39] in order to cope with the case of a singular input covariance matrix $\mathbf{Q}$. The singularity of $\mathbf{Q}$ may result as a consequence of the *water-filling* power allocation algorithm that may turn some of the channel eigenmodes off. Nevertheless, these modified equations are completely equivalent to the corresponding equations in [37]–[39].

This result can be applied to the calculation of the two differential entropy terms present in (1). The ergodic mutual information can be rewritten as

$$\mathsf{E}[I_{\mathbf{H},\mathbf{H}_I}(\mathbf{x};\mathbf{y})] = \mathcal{I}_1 - \mathcal{I}_2 \tag{10}$$

where we defined

$$\mathcal{I}_1 \triangleq \mathsf{E}\left[\ln\det(\mathbf{Q}_Z + \mathbf{H}\mathbf{Q}\mathbf{H}^{\mathsf{H}} + \mathbf{H}_I\mathbf{Q}_I\mathbf{H}_I^{\mathsf{H}}) - \ln\det(\mathbf{Q}_Z)\right] \tag{11}$$

and

$$\mathcal{I}_2 \triangleq \mathsf{E}\Big[\ln\det(\mathbf{Q}_Z + \mathbf{H}_I\mathbf{Q}_I\mathbf{H}_I^{\mathsf{H}}) - \ln\det(\mathbf{Q}_Z)\Big]. \tag{12}$$

Now, in order to find the asymptotic approximation to $\mathcal{I}_1$, we resort to the following equivalent expression of the channel (1):

$$\mathbf{y} = \left\{(\bar{\mathbf{H}}, \bar{\mathbf{H}}_I) + \mathbf{R}^{1/2}(\mathbf{W}, \mathbf{W}_I)\begin{pmatrix}\mathbf{T} & \mathbf{0}\\ \mathbf{0} & \mathbf{T}_I\end{pmatrix}^{1/2}\right\}\begin{pmatrix}\mathbf{x}\\ \mathbf{x}_I\end{pmatrix} + \mathbf{z}.$$

Then, using the asymptotic approximation (6), we obtain:

$$\mathcal{I}_1 \sim \mu_I(\bar{\mathbf{H}}_1, \mathbf{R}, \mathbf{T}_1, \mathbf{Q}_Z, \mathrm{diag}(\mathbf{Q}, \mathbf{Q}_I)) \tag{13}$$

where $\bar{\mathbf{H}}_1 \triangleq (\bar{\mathbf{H}}, \bar{\mathbf{H}}_I)$, $\mathbf{T}_1 \triangleq \mathrm{diag}(\mathbf{T}, \mathbf{T}_I)$, and

$$\mathcal{I}_2 \sim \mu_I(\bar{\mathbf{H}}_I, \mathbf{R}, \mathbf{T}_I, \mathbf{Q}_Z, \mathbf{Q}_I). \tag{14}$$

As a result, the ergodic channel capacity can asymptotically be approximated as follows:

$$\begin{aligned} C &= \max_{\mathbf{Q}\geq 0:\mathsf{Tr}(\mathbf{Q})\leq P}(\mathcal{I}_1 - \mathcal{I}_2)\\ &\sim \left\{\max_{\mathbf{Q}\geq 0,\mathsf{Tr}(\mathbf{Q})\leq P}\mu_I(\bar{\mathbf{H}}_1, \mathbf{R}, \mathbf{T}_1, \mathbf{Q}_Z, \mathrm{diag}(\mathbf{Q}, \mathbf{Q}_I))\right\}\\ &\quad - \mu_I(\bar{\mathbf{H}}_I, \mathbf{R}, \mathbf{T}_I, \mathbf{Q}_Z, \mathbf{Q}_I)\\ &\triangleq C_a. \end{aligned} \tag{15}$$

In the following section we present an algorithm to calculate the asymptotic approximation of the ergodic capacity C_a.

III. Ergodic Channel Capacity

The ergodic capacity and the corresponding optimum input covariance matrix are determined by maximizing the mutual information with respect to the input covariance matrix itself under the power constraint considered.

Following the approach presented in the previous section, we look for the asymptotic approximation C_a of the ergodic capacity C, obtained by maximizing the asymptotic approximation of $\mathcal{I}_1 - \mathcal{I}_2$ over the set of input covariance matrices satisfying a specific power constraint.

To this purpose, we notice the following identity holding for the last term in (7):

$$\begin{aligned} &\ln\det[\mathbf{I}_{n_T} + z\tilde{\mathbf{T}} + \tilde{\mathbf{H}}^{\mathsf{H}}(\mathbf{Q}_Z + w\mathbf{R})^{-1}\tilde{\mathbf{H}}]\\ &\quad = \ln\det[\mathbf{I}_{n_T} + (z\mathbf{T} + \bar{\mathbf{H}}^{\mathsf{H}}(\mathbf{Q}_Z + w\mathbf{R})^{-1}\bar{\mathbf{H}})\mathbf{Q}]. \end{aligned}$$

The derivation of C_a, defined in (15), is based on the solution of the following convex optimization problem:

$$\begin{cases}\text{minimize} & -\ln\det[\mathbf{I}_{n_T+n_I} + \boldsymbol{\Xi}\,\mathrm{diag}(\mathbf{Q}, \mathbf{Q}_I)]\\ \text{subject to} & \mathbf{Q}\geq\mathbf{0}, \mathsf{Tr}(\mathbf{Q})\leq P\end{cases} \tag{16}$$

where we defined $\boldsymbol{\Xi} \triangleq z\mathbf{T}_1 + \bar{\mathbf{H}}_1^{\mathsf{H}}(\mathbf{Q}_Z + w\mathbf{R})^{-1}\bar{\mathbf{H}}_1$ ($\mathbf{T}_1$ and $\bar{\mathbf{H}}_1$ were defined after (13)).

Here, we did not consider the term $\ln\det(\mathbf{Q}_Z + w\mathbf{R})$, which is independent of the input covariance matrix $\mathbf{Q}$ (for fixed w and z; see Appendix A and Appendix B with the description of the iterative water-filling algorithm and a discussion on its convergence).

Partitioning the matrix $\boldsymbol{\Xi}$ as

$$\boldsymbol{\Xi} = \begin{pmatrix}\boldsymbol{\Xi}_{11} & \boldsymbol{\Xi}_{12}\\ \boldsymbol{\Xi}_{21} & \boldsymbol{\Xi}_{22}\end{pmatrix}$$

where $\boldsymbol{\Xi}_{11} \in \mathbb{C}^{n_T\times n_T}$, $\boldsymbol{\Xi}_{12} \in \mathbb{C}^{n_T\times n_I}$, $\boldsymbol{\Xi}_{21} \in \mathbb{C}^{n_I\times n_T}$ and $\boldsymbol{\Xi}_{22} \in \mathbb{C}^{n_I\times n_I}$ and using the 2×2 block matrix determinant expression from [23, p.22], we obtain the following result:

$$\begin{aligned} &\ln\det[\mathbf{I}_{n_T+n_I} + \boldsymbol{\Xi}\,\mathrm{diag}(\mathbf{Q}, \mathbf{Q}_I)]\\ &\quad = \ln\det\begin{pmatrix}\mathbf{I}_{n_T} + \boldsymbol{\Xi}_{11}\mathbf{Q} & \boldsymbol{\Xi}_{12}\mathbf{Q}_I\\ \boldsymbol{\Xi}_{21}\mathbf{Q} & \mathbf{I}_{n_I} + \boldsymbol{\Xi}_{22}\mathbf{Q}_I\end{pmatrix}\\ &\quad = \ln\det(\mathbf{I}_{n_T} + \hat{\boldsymbol{\Xi}}\,\mathbf{Q}) + \ln\det(\mathbf{I}_{n_I} + \boldsymbol{\Xi}_{22}\mathbf{Q}_I) \end{aligned}$$

where we set

$$\hat{\boldsymbol{\Xi}} \triangleq \boldsymbol{\Xi}_{11} - \boldsymbol{\Xi}_{12}\mathbf{Q}_I(\mathbf{I}_{n_I} + \boldsymbol{\Xi}_{22}\mathbf{Q}_I)^{-1}\boldsymbol{\Xi}_{21} \tag{17}$$

which can be shown to be a positive semidefinite matrix. Therefore, since $\det(\mathbf{I}_{n_I} + \boldsymbol{\Xi}_{22}\mathbf{Q}_I)$ is independent of $\mathbf{Q}$, we can consider the following convex optimization problem, equivalent to (16):

$$\begin{cases}\text{minimize} & -\ln\det(\mathbf{I}_{n_T} + \hat{\boldsymbol{\Xi}}\,\mathbf{Q})\\ \text{subject to} & \mathbf{Q}\geq\mathbf{0}, \mathsf{Tr}(\mathbf{Q})\leq P.\end{cases} \tag{18}$$

Now, paralleling the approach of [47], we can see that the objective function in (18) is convex with respect to $\mathbf{Q}$ and the constraints satisfy Slater's qualification condition [3]. Thus, the Karush–Kuhn–Tucker (KKT) conditions are necessary and sufficient for the optimality of the solution [3]. In order to derive the KKT conditions, we can form the following Lagrangian function:

$$\begin{aligned} \mathcal{L}(\mathbf{Q}, \lambda, \boldsymbol{\Psi}) &= \ln\det(\mathbf{I}_{n_T} + \hat{\boldsymbol{\Xi}}\,\mathbf{Q})\\ &\quad - \lambda[\mathsf{Tr}(\mathbf{Q}) - P] + \mathsf{Tr}(\boldsymbol{\Psi}\mathbf{Q}) \end{aligned}$$

for $\lambda \geq 0, \boldsymbol{\Psi} \geq \mathbf{0}$. Under the assumption that $\hat{\boldsymbol{\Xi}}$ is positive definite (which will be released later), we calculate the gradient of $\mathcal{L}(\mathbf{Q}, \lambda, \boldsymbol{\Psi})$ with respect to $\mathbf{Q}$ and set it equal to zero. Then, we obtain the following KKT equations:

$$\begin{cases}(\hat{\boldsymbol{\Xi}}^{-1} + \mathbf{Q})^{-1} & = \lambda\mathbf{I}_{n_T} - \boldsymbol{\Psi}\\ \mathsf{Tr}(\mathbf{Q}) & = P\\ \mathsf{Tr}(\boldsymbol{\Psi}\mathbf{Q}) & = 0.\end{cases} \tag{19}$$

We can see from the first of (19) that λ must be positive. Otherwise, since $\bar{\mathbf{\Psi}}$ is positive semidefinite (by definition), the rhs could only be equal to the all-zero matrix, which is not possible.

When the matrix $\hat{\mathbf{\Xi}}$ is only positive *semidefinite* and its rank is lower than n_T, we need to modify the previous approach. Assume that $\hat{\mathbf{\Xi}} = \mathbf{U}_x^{\mathsf{H}} \mathbf{\Lambda}_x \mathbf{U}_x$ is the orthogonal decomposition of $\hat{\mathbf{\Xi}}$ with $\mathbf{\Lambda}_x = \mathrm{diag}(\tilde{\mathbf{\Lambda}}_x, \mathbf{0})$ and unitary $\mathbf{U}_x^{\mathsf{H}}$, where $\tilde{\mathbf{\Lambda}}_x$ is a positive definite diagonal matrix. Then, if $\rho_x = \mathrm{rank}(\hat{\mathbf{\Xi}}) < n_T$, we can see that

$$\ln\det(\mathbf{I}_{n_T} + \hat{\mathbf{\Xi}}\mathbf{Q}) = \ln\det(\mathbf{I}_{\rho_x} + \tilde{\mathbf{\Lambda}}_x \mathbf{Q}_1)$$

where $\mathbf{Q}_1$ is the $\rho_x \times \rho_x$ upper left sub-matrix of $\mathbf{U}_x \mathbf{Q} \mathbf{U}_x^{\mathsf{H}}$. We notice that

$$\mathsf{Tr}(\mathbf{Q}_1) \le \mathsf{Tr}(\mathbf{U}_x \mathbf{Q} \mathbf{U}_x^{\mathsf{H}}) = \mathsf{Tr}(\mathbf{Q}) \le P$$

so that the power constraint on $\mathbf{Q}_1$ remains unchanged. Then, repeating the previous argument used for positive definite $\mathbf{\Xi}$, we arrive at the following expression of the ergodic capacity achieving covariance matrix:

$$\mathbf{Q} = \mathbf{U}_x \mathrm{diag}\Big((\lambda \mathbf{I}_{\rho_x} - \tilde{\mathbf{\Lambda}}_x^{-1})_+, \mathbf{0}\Big) \mathbf{U}_x^{\mathsf{H}}. \tag{20}$$

The unknown parameter λ can be determined by solving the constraint equation $\mathsf{Tr}(\mathbf{Q}) = P$. It can be noticed that (20) can be written as

$$\mathbf{Q} = \mathbf{U}_x (\lambda \mathbf{I}_{n_T} - \mathbf{\Lambda}_x^{-1})_+ \mathbf{U}_x^{\mathsf{H}} \tag{21}$$

by assuming that matrix inversion maps the null entries on the main diagonal of the positive semidefinite matrix $\mathbf{\Lambda}_x$ to positive infinity. The unknown parameter λ is determined by solving the constraint equation:

$$\mathsf{Tr}[(\lambda \mathbf{I}_{n_T} - \mathbf{\Lambda}_x^{-1})_+] = P. \tag{22}$$

This is generally referred to as *water-filling* solution of the capacity optimization problem and can be found by following the guidelines from [10].

Remark III.1: The *water-filling* structure of the covariance matrix arises in several information theoretic optimization problems relevant to MIMO channels (see [47]). In the present case, the ergodic channel capacity can be found by solving the joint set of equations given by (9) and (21). This can be done by using an iterative water-filling algorithm that starts from an initial value of the variables $(w, z, \mathbf{Q})$ and proceeds by applying (9) and (21) iteratively until it converges. Appendix A contains a detailed description of this algorithm, whose convergence is discussed in Appendix B.

Remark III.2: Since the iterative water-filling algorithm converges to a unique solution in the covariance matrix $\mathbf{Q}$, a minor extension of [14, Prop.6] implies that this is also the optimum $\mathbf{Q}$ that maximizes $\mu_I(\bar{\mathbf{H}}, \mathbf{R}, \mathbf{T}, \mathbf{Q}_Z, \mathbf{Q})$ in (7).

IV. Numerical Results

In this section we present some numerical examples to validate the method developed in this work and to provide an assessment of the influence of several channel parameters on the iid mutual information and the channel capacity.

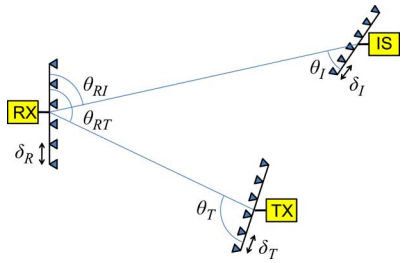

Fig. 1. Simulation scenario considered (array angles are different from the baseline scenario). It consists of three uniformly spaced antenna arrays connected to a transmitter (TX), a receiver, (RX) and an interfering source (IS).

We consider a MIMO communication system affected by co-channel interference and based on uniform linear antenna arrays (i.e., with antennas located at constant distance on straight lines). As a consequence of this assumption, the average channel matrices (accounting for the LOS propagation) have unit rank and can be expressed as follows [12], [15], [29], [41, Sec.7.2]:

$$\bar{\mathbf{H}} = \mathbf{a}_{RT} \mathbf{a}_T^{\mathsf{H}} \text{ and } \bar{\mathbf{H}}_I = \mathbf{a}_{RI} \mathbf{a}_I^{\mathsf{H}}.$$

In this expression we defined the following column vectors:

$$\begin{aligned}
\mathbf{a}_{RT} &= (\exp(-\mathrm{j}2\pi k \delta_R \cos\theta_{RT}))_{k=0}^{n_R-1} \\
\mathbf{a}_{RI} &= (\exp(-\mathrm{j}2\pi k \delta_R \cos\theta_{RI}))_{k=0}^{n_R-1} \\
\mathbf{a}_T &= (\exp(-\mathrm{j}2\pi k \delta_T \cos\theta_T))_{k=0}^{n_T-1} \\
\mathbf{a}_I &= (\exp(-\mathrm{j}2\pi k \delta_I \cos\theta_I))_{k=0}^{n_I-1}.
\end{aligned}$$

The above vectors contain the parameters $\delta_R, \delta_T, \delta_I$, which denote the constant spacing, in wavelengths, between adjacent antennas of the uniform linear arrays at the receiver, at the transmitter and at the interfering source, respectively. Additionally, $\theta_{RT}, \theta_{RI}, \theta_T, \theta_I$ denote the angles of the uniform linear arrays with respect to the direction of propagation. They refer to the angle at the receiver from the transmitter, at the receiver from the interfering source, at the transmitter and at the interfering source, respectively. The simulation scenario and its relevant parameters are represented in Fig. 1.

The transmit and receive correlation matrices are assumed to be *exponential* in the sense specified in [28]. In spite of being an approximation of the actual system correlation, this assumption has the considerable advantage of being a single-parameter characterization. The correlation matrices are defined as follows:

$$\begin{cases}
\mathbf{R} &= (\alpha_R^{|i-j|})_{i,j=1}^{n_R} \\
\mathbf{T} &= (\alpha_T^{|i-j|})_{i,j=1}^{n_T} \\
\mathbf{T}_I &= (\alpha_I^{|i-j|})_{i,j=1}^{n_I}.
\end{cases} \tag{23}$$

These matrices are positive definite for $\alpha \in [0, 1)$. When $\alpha = 0$ we have no correlation. Though antenna spacing and correlation are connected, their relationship is far from being simple, as well as evidenced, e.g., in [5] and will not be addressed in the present study.

Based on the previous assumptions, we consider a baseline reference scenario and analyze the impact of the variation of

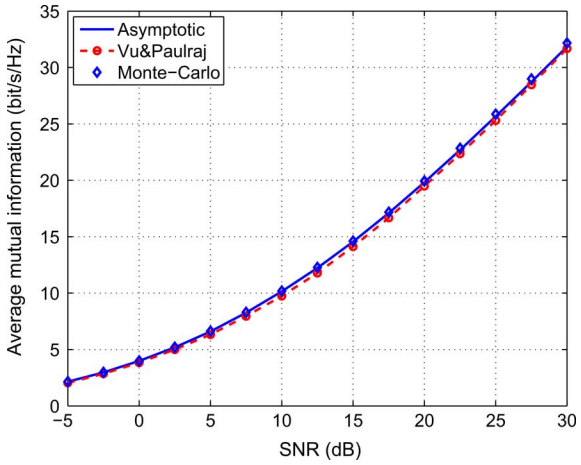

Fig. 2. Mutual information of the 4 × 4 MIMO channel proposed in [45].

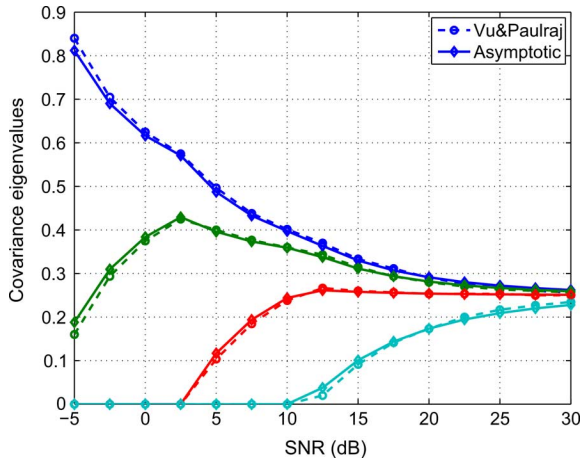

Fig. 3. Normalized eigenvalues of the optimum transmit covariance matrix for the 4 × 4 MIMO channel proposed in [45].

several parameters on its capacity performance. The baseline scenario is defined by the following parameters:

- The number of receive antennas is $n_R = 4$.
- The number of transmit antennas at the signal source is $n_T = 4$.
- The number of transmit antennas at the interfering source is $n_I = 4$.
- The antenna spacing of all arrays is equal to $\delta = 1$.
- The array angles are $\theta_{RT} = \theta_T = \theta_I = \frac{\pi}{2}$ and $\theta_{RI} = 0$ (it will be shown that the last parameter represents a worst case for system performance and maximizes the negative effect of interference).
- The Rician factors are $K = K_I = 10\,\mathrm{dB}$.
- The MIMO system is spatially uncorrelated, i.e., we assume that $\alpha_R = \alpha_T = \alpha_I = 0$ in (23).
- The covariance matrix of the interfering signal is proportional to the identity matrix (iid interference).

A. Comparison With the Results From [45]

In order to validate the method proposed in this paper, we compare the results obtained versus those presented in [45]. Our results are presented in Fig. 2 (channel capacity) and 3 (normalized covariance matrix eigenvalues). Specifically, Fig. 2 reports the capacity obtained numerically in [45, Fig.3, upper] (dotted line with circles), the capacity obtained by our asymptotic approach (solid line) and the capacity obtained by Monte-Carlo simulation and using the transmit covariance matrix obtained by our asymptotic approach (diamonds). Fig. 3 reports the normalized covariance matrix eigenvalues obtained numerically in [45, Fig.3, lower] (dotted lines with circles) and those obtained by our asymptotic approach (solid lines with diamonds). In both cases, our results are in close agreement with [45].

B. Comparison With the Results From [34]

It is worth comparing the asymptotic method proposed with another asymptotic method presented in [34]. We focus on the results presented in [34, Fig. 1] providing the mutual information scaling coefficient (in nat/s/Hz/antenna) versus the number of antennas for a separately correlated Rayleigh fading MIMO channel with correlation matrices of exponential type and base

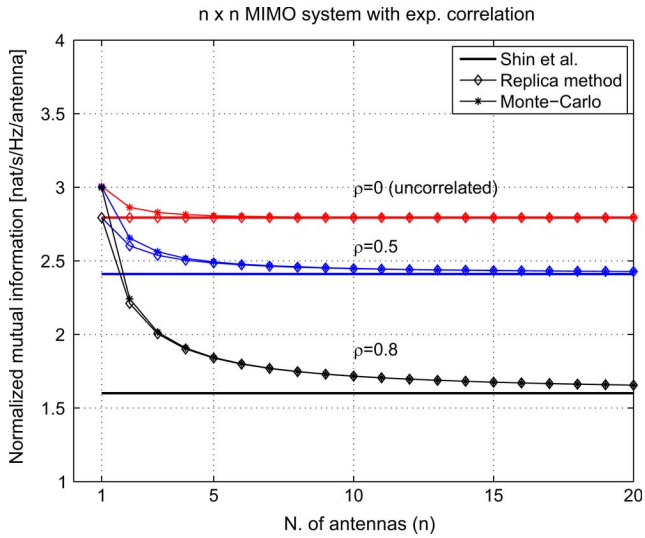

Fig. 4. Average mutual information per antenna for an $n \times n$ MIMO channel with exponential transmit and receive correlation with base $\rho = 0, 0.5, 0.8$, $\mathrm{SNR} = 15$ dB obtained by Monte-Carlo simulation, our method and the asymptotic method from [34].

0 (uncorrelated), 0.5 and 0.8. Fig. 4 compares the results from [34, Fig. 1] and from the application of our asymptotic method. The diagrams show that the proposed method (albeit equivalent to [14], [31] in this case) provides better accuracy for a finite number of antennas than the method from [34] when the level of spatial correlation ρ is greater than 0.

C. Accuracy of the Asymptotic Approximation

Here we consider the accuracy of the asymptotic approximation with respect to the number of antennas. Fig. 5 shows the accuracy of our baseline reference system with $\mathrm{SNR} = 0$ dB and $\mathrm{SIR} = -10$ dB. Here, $n_R = n_T = n_I = n$, which is in abscissa.

The diagrams report the mutual information with iid and optimum input power allocation (denoted by 'IID' and 'CAP', respectively; the latter corresponding to ergodic capacity) under the assumption that interference is dealt with as another signal

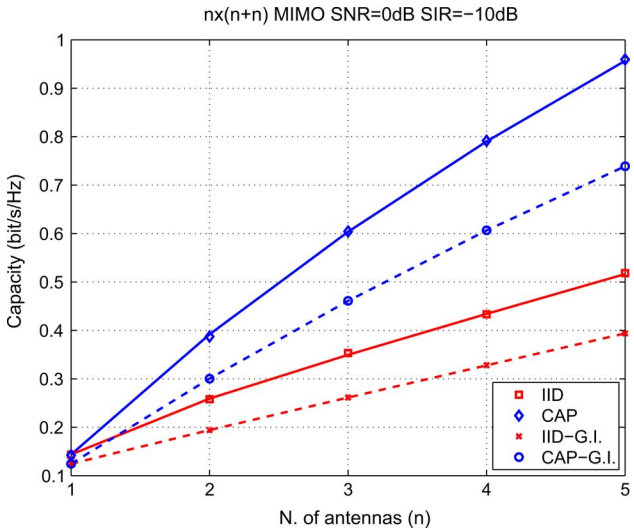

Fig. 5. Capacity versus number of antennas n for the baseline reference MIMO system with $n_R = n_T = n_I = n$, $\mathrm{SNR} = 0$ dB and $\mathrm{SIR} = -10$ dB.

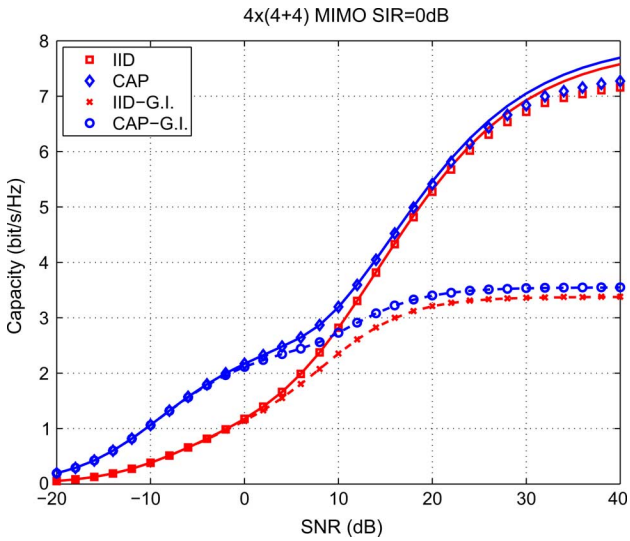

Fig. 7. Capacity versus SIR (dB) for the baseline reference MIMO system with interference and $\mathrm{SNR} = 2$ dB.

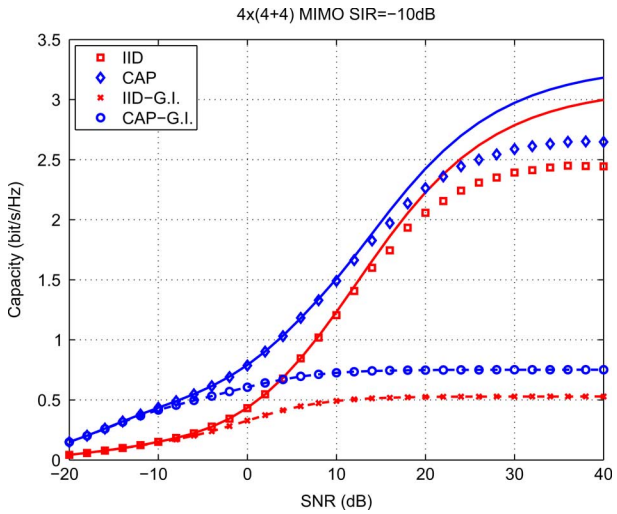

Fig. 6. Capacity versus SNR (dB) for the baseline reference MIMO system with interference and $\mathrm{SIR} = 2$ dB.

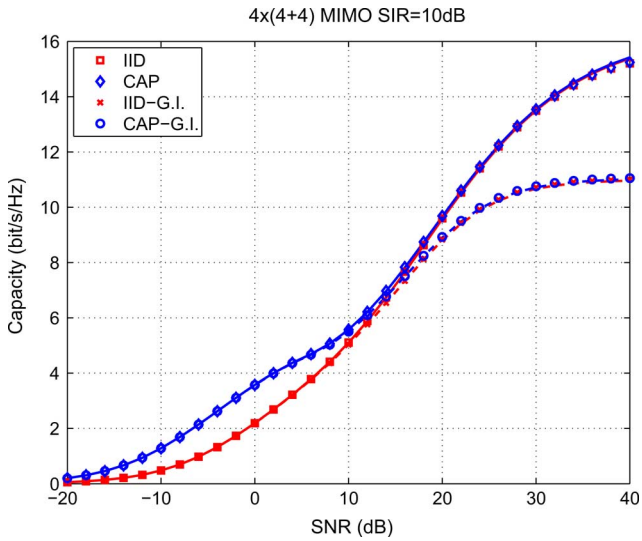

Fig. 8. Capacity versus correlation base α for the baseline reference MIMO system with interference and $\mathrm{SNR} = 10$ dB and $\mathrm{SIR} = 2$ dB.

or as Gaussian noise (denoted by 'IID-G.I.' and 'CAP-G.I.', respectively; the latter assumption is referred to as *Gaussian interference* assumption). Markers represent Monte-Carlo simulation results whereas solid and dashed lines correspond to asymptotic approximations. It can be noticed that the differences are very limited, even in the case of $n = 2$ or 1. This supports the adoption of the asymptotic approximation as a viable tool to obtain the system capacity at a moderate complexity price without sacrificing the precision of the numerical results.

D. Effect of Interference

Fig. 6 to Fig. 8 plot the system mutual information of the baseline reference MIMO system with interference and $\mathrm{SIR} = -10, 0, 10$ dB, respectively. From the diagrams, it is clear that the Gaussian interference assumption corresponds to lower values of mutual information, as was evidenced, e.g., in [6], [7]. This effect is emphasized as the SNR gets larger whereas it has more limited impact for lower SNR.

The diagrams also show the difference between ergodic capacity and iid mutual information (corresponding to iid input power allocation, i.e., input signal covariance matrix proportional to the identity matrix). As expected, the difference decreases as the SNR gets larger.

In all cases, fixing the SIR to a constant value, forces an upper bound to the achievable capacity at infinite SNR. This capacity ceiling is significantly different according to whether interference is considered as a signal or as additive Gaussian noise.

The limited accuracy of the results of Figs. 6 and 7 (in the cases of interference treated as a signal and not as additive Gaussian noise) depends on the fact that, most notably for $\mathrm{SIR} = -10$ dB but also for $\mathrm{SIR} = 0$ dB, the mutual information is obtained as the difference between two large numerical values and the results are then affected by a cancelation error that amplifies the difference between asymptotic and simulation results.

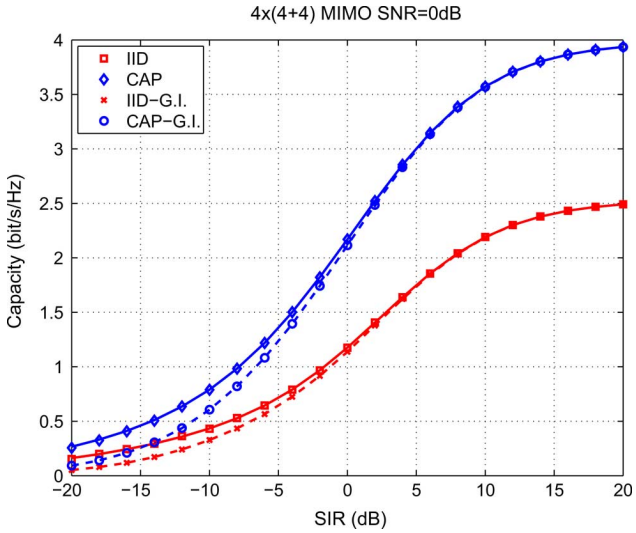

Fig. 9. Capacity versus the Rician factors $K = K_I$ (dB) for the baseline reference MIMO system with interference and $\mathrm{SNR} = 10$ dB and $\mathrm{SIR} = 2$ dB.

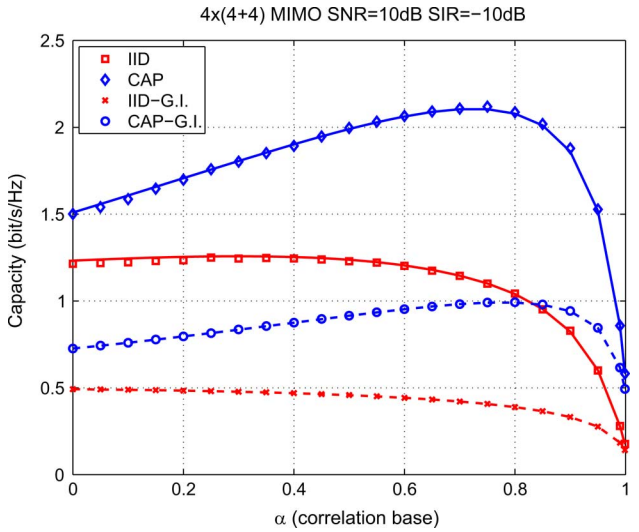

Fig. 11. Capacity versus correlation base α for the baseline reference MIMO system with interference and $\mathrm{SNR} = 10$ dB and $\mathrm{SIR} = 10$ dB.

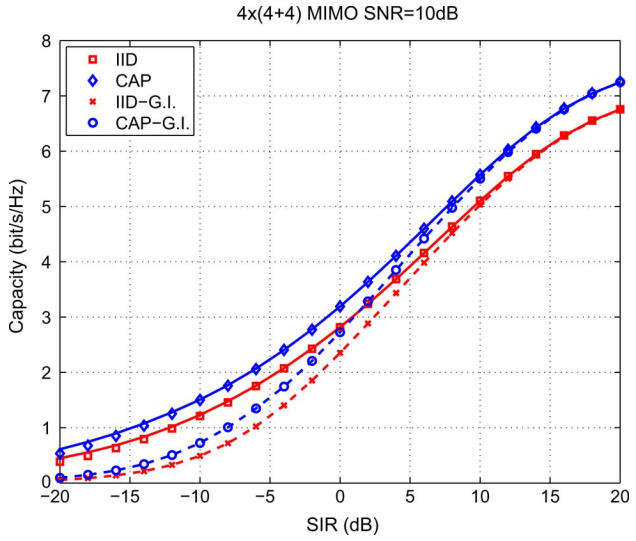

Fig. 10. Capacity versus the interference angle of arrival θ_{RI} for the baseline reference MIMO system with $\mathrm{SNR} = 10$ dB and $\mathrm{SIR} = 2$ dB.

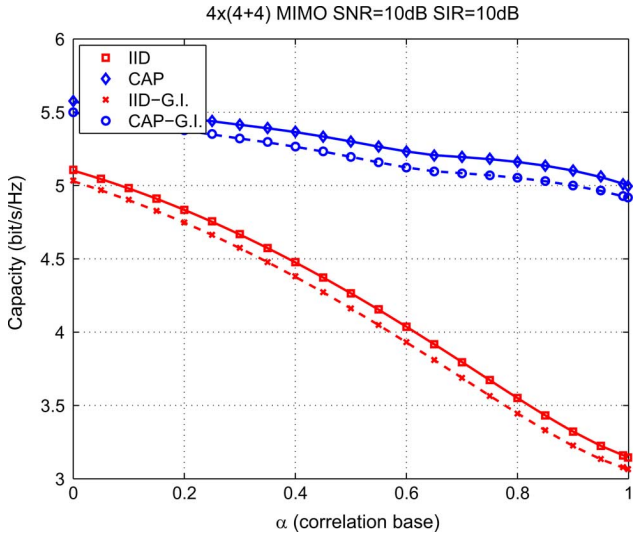

Fig. 12. Capacity versus correlation base α for the baseline reference MIMO system with interference and $\mathrm{SNR} = 10$ dB and $\mathrm{SIR} = 10$ dB.

An orthogonal viewpoint is illustrated in Fig. 9 to Fig. 10, which report the variation of the mutual information versus the SIR at fixed values of SNR (0 and 10 dB, respectively).

We can notice that the capacity behavior as the SIR increases depends very much on the SNR level considered. For low SNR, there is a wide gap between mutual information and capacity, while for high SNR the gap reduces, as a consequence of the asymptotic (with respect to the SNR) optimality of iid power allocation.

The Gaussian interference assumption has an impact if the SIR is low whereas after a certain threshold value becomes unnoticeable.

E. Effect of Correlation

In order to assess the effect of spatial correlation we plot the system mutual information versus a common correlation base α, which is assumed to determine the spatial correlation at the receiver end and at both the signal and interfering transmitter ends. In other words, we set $\alpha_R = \alpha_T = \alpha_I = \alpha$.

Figs. 11 and 12 plot the mutual information versus α for $\mathrm{SNR} = 10$ dB and $\mathrm{SIR} = -10$ and 10 dB, respectively.

We can see that the effect of increasing spatial correlation is quite unpredictable when the SIR is low whereas it provides the expected results when the SIR is high (capacity decreasing with correlation). The unreliability of the assumption that correlation decreases capacity has been observed in other works, e.g., [5]. These observations support the fact that accurate channel modeling as well as precise evaluation of channel capacity are mandatory in order to assess correctly the performance of practical MIMO systems affected by spatial correlation.

Finally, we notice from Fig. 12 that the negative effect of spatial correlation is more noticeable without power allocation optimization. In fact, the mutual information (based on iid power allocation) decreases more than the capacity (based on optimal

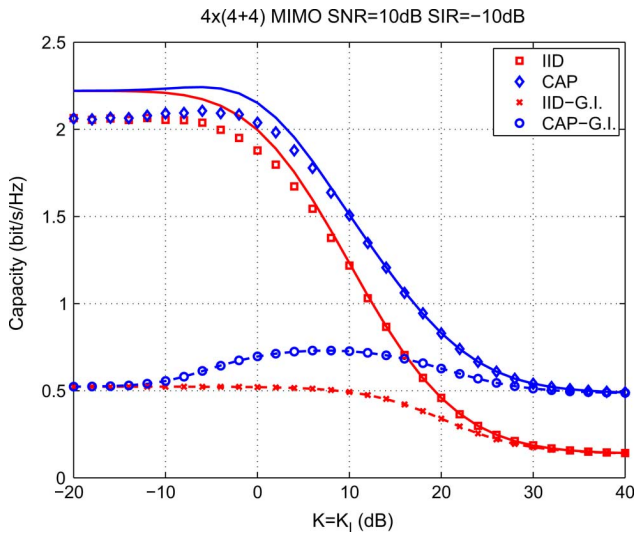

Fig. 13. Capacity versus the Rician factors $K = K_I$ (dB) for the baseline reference MIMO system with interference and $\mathrm{SNR} = 10$ dB and $\mathrm{SIR} = -10$ dB.

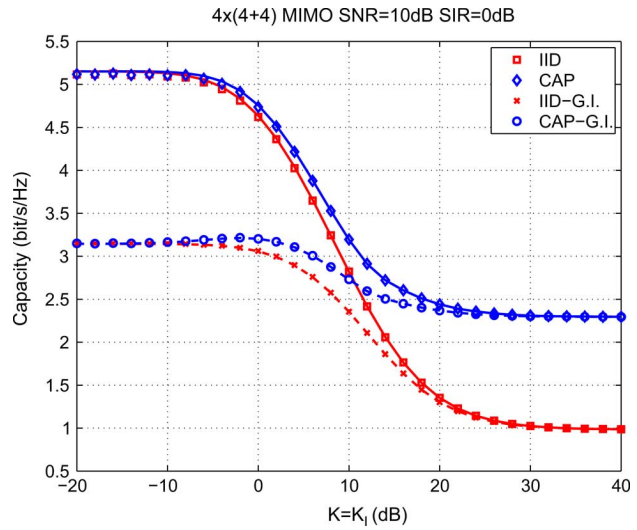

Fig. 14. Capacity versus the Rician factors $K = K_I$ (dB) for the baseline reference MIMO system with interference and $\mathrm{SNR} = 10$ dB and $\mathrm{SIR} = 0$ dB.

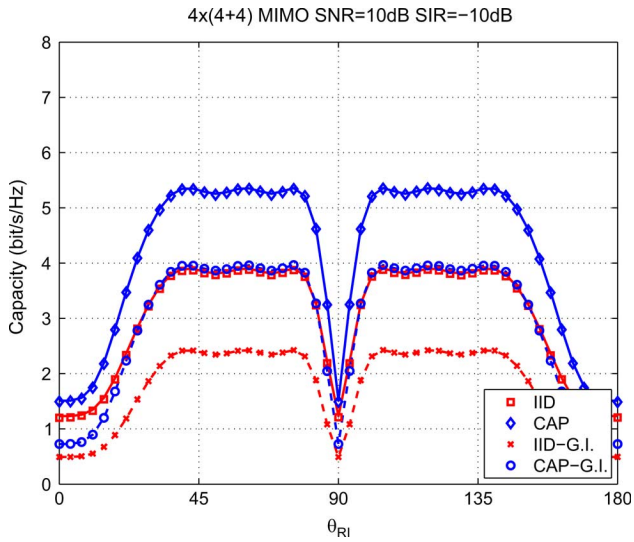

Fig. 15. Capacity versus the interference angle of arrival θ_{RI} for the baseline reference MIMO system with $\mathrm{SNR} = 10$ dB and $\mathrm{SIR} = -10$ dB.

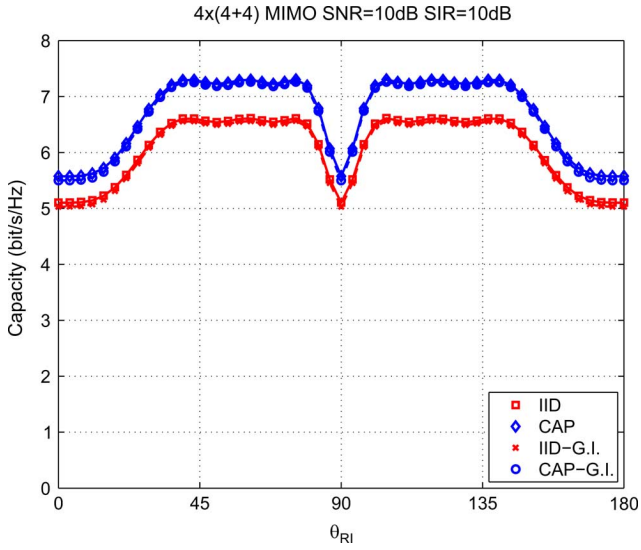

Fig. 16. Capacity versus the interference angle of arrival θ_{RI} for the baseline reference MIMO system with $\mathrm{SNR} = 10$ dB and $\mathrm{SIR} = 10$ dB.

input covariance matrix) so that we can say that power allocation optimization buys back some of the capacity loss entailed by stronger spatial correlation.

F. Effect of the Rician Factor

Figs. 13 and 14 illustrate the effect of the Rician factor K on the mutual information when the $\mathrm{SNR} = 10$ dB and the $\mathrm{SIR} = -10$ and 0 dB, respectively.

It can be noticed that all the system mutual information plots decrease monotonically as the Rician factor increases, as a consequence of having a fixed SNR. In fact, MIMO systems are affected negatively by reducing the level of scattering, which decreases as K increases.

It is noticeable that the case of strong interference ($\mathrm{SIR} = -10$ dB) exhibits also a large gap between the mutual information with and without the Gaussian interference assumption at very low K (Rayleigh fading conditions). In such situation it is vital that the interference can be recovered as a signal by successive decoding with interference cancelation because otherwise its effect can be very strong.

Again, as for Figs. 6 and 7, the limited accuracy of the results in Fig. 13 (relevant to the cases of interference treated as a signal and not as additive Gaussian noise) derives from cancelation errors.

G. Effect of the Angle of Arrival of Interference

Here we assess the effect of the angle of arrival of interference θ_{RI} on the system capacity. Figs. 15 and 16 plot the system mutual information versus θ_{RI} for $\mathrm{SNR} = 10$ dB and $\mathrm{SIR} = -10$ and 10 dB, respectively. The diagrams show that the effect of interference is stronger when the angle of arrival is a multiple of $\pi/2$ (critical angles). They also show that power allocation

optimization is more effective when the interference is not arriving from a critical angle though its power level is considerably higher than the power level of the useful signal. Again, the Gaussian interference assumption corresponds to substantially lower capacity results than the assumption that interference is a signal when the SIR is low.

V. Conclusion

In this paper, we provided an asymptotic method to derive a very good approximation to the ergodic capacity of a MIMO communication system affected by additive noise, interference, Rician fading and spatial correlation (according to the separately correlated or Kronecker model).

First, we provided an asymptotic algorithm to derive the mutual information. Next, we showed how power allocation and the input signal covariance matrix can be optimized in order to find the ergodic capacity. The convergence of this water-filling algorithm has been discussed in Appendix B.

We also analyzed the effect of considering interference as additive noise with known covariance. Though this assumption allows to simplify the receiver design, it has been shown that the results are highly suboptimal if the interference power level is relatively high compared to the signal power level.

We also assessed the impact of various system parameters on the system capacity performance. Besides the effect of the SNR and of the SIR, we also analyzed the effect of spatial correlation (through the use of a single parameter according to the exponential model), of the Rician factor and of the angle of arrival of interference. Concerning this last parameter, we showed how spatial geometry has a basic influence on the MIMO system performance and a simple analysis prompts with the existence of critical angles of arrival of the interference corresponding to negative peaks in the capacity and mutual information.

Summarizing, one of the reasons of the difficulty in analyzing MIMO communication system is the large number of parameters involved, which can affect the system performance not always independently and often in an unpredictable way. Therefore, the proposed algorithms to calculate the mutual information and the capacity represent a valuable tool for the system design and pave the way to parameter optimization in the complex MIMO environment.

Appendix A
Asymptotic Ergodic Capacity Algorithm

Here we summarize the ergodic capacity asymptotic approximation by the iterative water-filling algorithm derived in Section III. The asymptotic approximation of the ergodic capacity is given by

$$C_a = \left\{ \max_{\mathbf{Q}\geq 0, \mathsf{Tr}(\mathbf{Q})\leq P} \mu_I(\bar{\mathbf{H}}_1, \mathbf{R}, \mathbf{T}_1, \mathbf{Q}_Z, \mathrm{diag}(\mathbf{Q}, \mathbf{Q}_I)) \right\} - \mu_I(\bar{\mathbf{H}}_I, \mathbf{R}, \mathbf{T}_I, \mathbf{Q}_Z, \mathbf{Q}_I) \quad (24)$$

where $\bar{\mathbf{H}}_1 \triangleq (\bar{\mathbf{H}}, \bar{\mathbf{H}}_I)$ and $\mathbf{T}_1 \triangleq \mathrm{diag}(\mathbf{T}, \mathbf{T}_I)$.

The second term can be calculated by the following algorithm implementing the function $\mu_I(\bar{\mathbf{H}}_I, \mathbf{R}, \mathbf{T}_I, \mathbf{Q}_Z, \mathbf{Q}_I)$:

— Input: $\bar{\mathbf{H}}_I, \mathbf{R}, \mathbf{T}_I, \mathbf{Q}_Z, \mathbf{Q}_I$.

— Initialize $w_0 = z_0 = 1$, $\tilde{\mathbf{T}} = \mathbf{Q}_I^{1/2}\mathbf{T}_I\mathbf{Q}_I^{1/2}$ and $\tilde{\mathbf{H}} = \bar{\mathbf{H}}_I\mathbf{Q}_I^{1/2}$.

— Repeat for $\ell = 0, 1, 2, \ldots$ the following steps:

- Calculate:

$$\begin{cases} w_{\ell+1} = \mathsf{Tr}\left\{ [\mathbf{I}_{n_T} + z_\ell \tilde{\mathbf{T}} + \tilde{\mathbf{H}}^{\mathsf{H}}(\mathbf{Q}_Z + w_\ell \mathbf{R})^{-1}\tilde{\mathbf{H}}]^{-1}\tilde{\mathbf{T}} \right\} \\ z_{\ell+1} = \mathsf{Tr}\left\{ [\mathbf{I}_{n_R} + w_\ell \mathbf{R} + \tilde{\mathbf{H}}(\mathbf{I}_{n_T} + z_\ell \tilde{\mathbf{T}})^{-1}\tilde{\mathbf{H}}^{\mathsf{H}}]^{-1}\mathbf{R} \right\}. \end{cases}$$

- If $\max\{|w_{\ell+1}/w_\ell - 1|, |z_{\ell+1}/z_\ell - 1|\} < \varepsilon$, a prescribed error threshold, then exit the loop and denote by w, z the last calculated $w_{\ell+1}, z_{\ell+1}$

— Output:

$$\mu_I = \ln\det(\mathbf{Q}_Z + w\mathbf{R}) - wz - \ln\det(\mathbf{Q}_Z) + \ln\det[\mathbf{I}_{n_T} + z\tilde{\mathbf{T}} + \tilde{\mathbf{H}}^{\mathsf{H}}(\mathbf{Q}_Z + w\mathbf{R})^{-1}\tilde{\mathbf{H}}].$$

The first term of (24) can be calculated by using the following algorithm that incorporates the input covariance matrix optimization:

— Input: $\bar{\mathbf{H}}, \bar{\mathbf{H}}_I, \mathbf{R}, \mathbf{T}, \mathbf{T}_I, \mathbf{Q}_Z, P, \mathbf{Q}_I$.

— Initialize $w_0 = z_0 = 1$, $\bar{\mathbf{H}}_1 \triangleq (\bar{\mathbf{H}}, \bar{\mathbf{H}}_I)$, $\mathbf{T}_1 \triangleq \mathrm{diag}(\mathbf{T}, \mathbf{T}_I)$.

— Repeat for $\ell = 0, 1, 2, \ldots$ the following steps:

- Set

$$\begin{aligned} \Xi_{11} &= z_\ell \mathbf{T} + \bar{\mathbf{H}}^{\mathsf{H}}(\mathbf{Q}_Z + w_\ell \mathbf{R})^{-1}\bar{\mathbf{H}} \\ \Xi_{12} &= \bar{\mathbf{H}}^{\mathsf{H}}(\mathbf{Q}_Z + w_\ell \mathbf{R})^{-1}\bar{\mathbf{H}}_I \\ \Xi_{21} &= \bar{\mathbf{H}}_I^{\mathsf{H}}(\mathbf{Q}_Z + w_\ell \mathbf{R})^{-1}\bar{\mathbf{H}} \\ \Xi_{22} &= z_\ell \mathbf{T}_I + \bar{\mathbf{H}}_I^{\mathsf{H}}(\mathbf{Q}_Z + w_\ell \mathbf{R})^{-1}\bar{\mathbf{H}}_I. \end{aligned}$$

- Calculate

$$\hat{\Xi} = \Xi_{11} - \Xi_{12}\mathbf{Q}_I(\mathbf{I}_{n_I} + \Xi_{22}\mathbf{Q}_I)^{-1}\Xi_{21}.$$

- Calculate the orthogonal factorization $\hat{\Xi} = \mathbf{U}_x^{\mathsf{H}}\mathbf{\Lambda}_x\mathbf{U}_x$.
- Solve the *water-filling* equation $\mathsf{Tr}[(\lambda\mathbf{I}_{n_T} - \mathbf{\Lambda}_x^{-1})_+] = P$ for λ (mapping null values on the diagonal of $\mathbf{\Lambda}_x$ to $+\infty$).
- Set: $\mathbf{Q} = \mathbf{U}_x(\lambda\mathbf{I}_{n_T} - \mathbf{\Lambda}_x^{-1})_+\mathbf{U}_x^{\mathsf{H}}$, $\mathbf{Q}_1 \triangleq \mathrm{diag}(\mathbf{Q}, \mathbf{Q}_I)$.
- Calculate $\tilde{\mathbf{T}} = \mathbf{Q}_1^{1/2}\mathbf{T}_1\mathbf{Q}_1^{1/2}$, $\tilde{\mathbf{H}} = \bar{\mathbf{H}}_1\mathbf{Q}_1^{1/2}$ and

$$\begin{cases} w_{\ell+1} = \mathsf{Tr}\left\{ [\mathbf{I}_{n_T+n_I} + z_\ell \tilde{\mathbf{T}} + \tilde{\mathbf{H}}^{\mathsf{H}}(\mathbf{Q}_Z + w_\ell \mathbf{R})^{-1}\tilde{\mathbf{H}}]^{-1}\tilde{\mathbf{T}} \right\} \\ z_{\ell+1} = \mathsf{Tr}\left\{ [\mathbf{I}_{n_R} + w_\ell \mathbf{R} + \tilde{\mathbf{H}}(\mathbf{I}_{n_T+n_I} + z_\ell \tilde{\mathbf{T}})^{-1}\tilde{\mathbf{H}}^{\mathsf{H}}]^{-1}\mathbf{R} \right\}. \end{cases}$$

- If the relative differences $|w_{\ell+1}/w_\ell - 1|$ and $|z_{\ell+1}/z_\ell - 1|$ are lower than a prescribed error threshold, then exit the loop and denote by w, z the last calculated $w_{\ell+1}, z_{\ell+1}$.

— Output

$$\mu_I = \ln\det(\mathbf{Q}_Z + w\mathbf{R}) - wz - \ln\det(\mathbf{Q}_Z) + \ln\det(\mathbf{I}_{n_T} + \hat{\Xi}\,\mathbf{Q}) + \ln\det(\mathbf{I}_{n_I} + \Xi_{22}\mathbf{Q}_I).$$

APPENDIX B
ON THE CONVERGENCE OF THE ITERATIVE WATER-FILLING ALGORITHM OF APPENDIX A

Here we address the convergence of the iterative water-filling algorithm presented in Appendix A. A similar algorithm was proposed in [14], nicknamed *frozen water-filling*. The difference between the two algorithms is that

- frozen water-filling solves the fixed-point (9) completely at every iteration step; while
- the proposed water-filling algorithm performs a single iteration step in the solution of(9) (see Appendix A).

As shown in [14], the algorithm converges if the sequence $(w_\ell, z_\ell)_{\ell=0}^{\infty}$ converges.

However, we can see that in some cases the frozen water-filling algorithm does not converge but oscillates between different solutions, corresponding to different ranks of the covariance matrix $\mathbf{Q}$. An example of this oscillating behavior is given in Appendix C.

In the following we focus on our proposed iterative water-filling algorithm. Let us consider, for the sake of simplicity, an interference-free scenario ($n_I = 0$) with Rayleigh fading ($\bar{\mathbf{H}} = \mathbf{0}$) and assume, without loss of generality, $\mathbf{Q}_Z = \mathbf{I}_{n_R}$. In this case, the function defined in (7) simplifies as follows:

$$\mu(w, z, \mathbf{Q}) \triangleq \ln\det(\mathbf{I}_{n_R} + w\mathbf{R}) + \ln\det[\mathbf{I}_{n_T} + z\mathbf{T}\mathbf{Q}] - wz. \tag{25}$$

Then, using the notations of Appendix A, the first term of (24) can be calculated by the following (simplified) algorithm:

- Initialize $w_0 = z_0 = 1$.
- Calculate the orthogonal factorization $\mathbf{T} = \mathbf{U}_t\mathbf{\Lambda}_t\mathbf{U}_t^{\mathsf{H}}$. We assume that $\mathbf{\Lambda}_t = \operatorname{diag}(\lambda_1(\mathbf{T}), \ldots, \lambda_{n_T}(\mathbf{T}))$, where $\lambda_i(\mathbf{X})$ denotes the ith eigenvalue of the matrix $\mathbf{X}$, sorted in nonincreasing order ($\lambda_1(\mathbf{X}) \geq \lambda_2(\mathbf{X}) \geq \cdots$).
- Repeat for $\ell = 0, 1, 2, \ldots$ the following steps:
  - Solve $\mathsf{Tr}[(\lambda\mathbf{I}_{n_T} - z_\ell^{-1}\mathbf{\Lambda}_t^{-1})_+] = P$ for λ and let the solution be ξ_ℓ.
  - Let $\mathbf{Q}_\ell = \mathbf{U}_t(\xi_\ell\mathbf{I}_{n_T} - z_\ell^{-1}\mathbf{\Lambda}_t^{-1})_+\mathbf{U}_t^{\mathsf{H}}$.
  - Calculate

$$\begin{cases} w_{\ell+1} &= \mathsf{Tr}\Big\{[\mathbf{I}_{n_T} + z_\ell\mathbf{T}\mathbf{Q}_\ell]^{-1}\mathbf{T}\mathbf{Q}_\ell\Big\} \\ z_{\ell+1} &= \mathsf{Tr}\Big\{[\mathbf{I}_{n_R} + w_\ell\mathbf{R}]^{-1}\mathbf{R}\Big\}. \end{cases} \tag{26}$$

  - If $\max\{|w_{\ell+1}/w_\ell - 1|, |z_{\ell+1}/z_\ell - 1|\} < \varepsilon$, then exit the repeat loop.

Notice that, with *frozen* water-filling, the iteration step (26) should have been replaced by:

- Solve

$$\begin{cases} w &= \mathsf{Tr}\Big\{[\mathbf{I}_{n_T} + z\mathbf{T}\mathbf{Q}_\ell]^{-1}\mathbf{T}\mathbf{Q}_\ell\Big\} \\ z &= \mathsf{Tr}\Big\{[\mathbf{I}_{n_R} + w\mathbf{R}]^{-1}\mathbf{R}\Big\} \end{cases} \tag{27}$$

  and let the solution be $(w_{\ell+1}, z_{\ell+1})$.

In order to simplify the algorithm we set $\rho_\ell \triangleq \operatorname{rank}(\mathbf{Q}_\ell)$ for $\ell = 0, 1, \ldots$ Then, we have

$$\lambda_i(\mathbf{Q}_\ell) = \begin{cases} \xi_\ell - \frac{1}{z_\ell\lambda_i(\mathbf{T})} & 1 \leq i \leq \rho_\ell \\ 0 & i > \rho_\ell \end{cases} \tag{28}$$

and

$$\xi_\ell = \frac{1}{\rho_\ell}\left\{P + \frac{1}{z_\ell}\sum_{i=1}^{\rho_\ell}\frac{1}{\lambda_i(\mathbf{T})}\right\}.$$

The rank ρ_ℓ can be determined by the inequalities

$$\frac{1}{z_\ell\lambda_{\rho_\ell}(\mathbf{T})} < \xi_\ell \leq \frac{1}{z_\ell\lambda_{\rho_\ell+1}(\mathbf{T})} \tag{29}$$

or by noticing that

$$\rho_\ell = \max\left\{\rho : \Delta_\rho \triangleq \frac{\rho}{P\lambda_\rho(\mathbf{T})} - \sum_{i=1}^{\rho}\frac{1}{P\lambda_i(\mathbf{T})} < z_\ell\right\}. \tag{30}$$

It is worth pointing out that the sequence $\{\Delta_\rho\}_{\rho=1}^{n_T}$ is nondecreasing. In fact, by definition

$$\begin{aligned} \Delta_{\rho+1} - \Delta_\rho &= \frac{\rho+1}{P\lambda_{\rho+1}(\mathbf{T})} - \sum_{i=1}^{\rho+1}\frac{1}{P\lambda_i(\mathbf{T})} \\ &\quad - \frac{\rho}{P\lambda_\rho(\mathbf{T})} + \sum_{i=1}^{\rho}\frac{1}{P\lambda_i(\mathbf{T})} \\ &= \frac{\rho}{P\lambda_{\rho+1}(\mathbf{T})} - \frac{\rho}{P\lambda_\rho(\mathbf{T})} \\ &\geq 0 \end{aligned}$$

since the sequence of eigenvalues of $\mathbf{T}$ is nonincreasing. Moreover, since

$$\begin{aligned} \mathbf{T}\mathbf{Q}_\ell &= \mathbf{U}_t\mathbf{\Lambda}_t\mathbf{U}_t^{\mathsf{H}}\mathbf{U}_t(\xi_\ell\mathbf{I}_{n_T} - z_\ell^{-1}\mathbf{\Lambda}_t^{-1})_+\mathbf{U}_t^{\mathsf{H}} \\ &= \mathbf{U}_t(\xi_\ell\mathbf{\Lambda}_t - z_\ell^{-1}\mathbf{I}_{n_T})_+\mathbf{U}_t^{\mathsf{H}} \end{aligned}$$

the eigenvalues of the matrix $\mathbf{T}\mathbf{Q}_\ell$ are given by

$$\lambda_i(\mathbf{T}\mathbf{Q}_\ell) = (\xi_\ell\lambda_i(\mathbf{T}) - z_\ell^{-1})_+, \quad \ell = 0, 1, \ldots.$$

Then, the iterative step (26) can be simplified as follows:

$$\begin{aligned} w_{\ell+1} &= \sum_{i=1}^{\rho_\ell}\frac{\lambda_i(\mathbf{T}\mathbf{Q}_\ell)}{1 + z_\ell\lambda_i(\mathbf{T}\mathbf{Q}_\ell)} \\ &= \sum_{i=1}^{\rho_\ell}\frac{\xi_\ell\lambda_i(\mathbf{T}) - z_\ell^{-1}}{z_\ell\xi_\ell\lambda_i(\mathbf{T})} \\ &= \frac{1}{z_\ell\xi_\ell}\sum_{i=1}^{\rho_\ell}\left(\xi_\ell - \frac{1}{z_\ell\lambda_i(\mathbf{T})}\right) \\ &= \frac{P}{z_\ell\xi_\ell} = \frac{\rho_\ell}{z_\ell + \frac{1}{P}\sum_{i=1}^{\rho_\ell}\frac{1}{\lambda_i(\mathbf{T})}} \\ z_{\ell+1} &= \sum_{i=1}^{n_R}\frac{1}{w_\ell + 1/\lambda_i(\mathbf{R})}. \end{aligned}$$

After collecting the previous simplifications, we obtain a simplified version of the water-filling algorithm.

- Initialize $w_0 = z_0 = 1$.

- Factorize $\mathbf{T} = \mathbf{U}_t \mathbf{\Lambda}_t \mathbf{U}_t^{\mathsf{H}}$ and set

$$\begin{cases} \sigma_\rho \triangleq \sum_{i=1}^{\rho} \frac{1}{P\lambda_i(\mathbf{T})} \\ \Delta_\rho \triangleq \frac{\rho}{P\lambda_\rho(\mathbf{T})} - \sigma_\rho \end{cases}. \tag{31}$$

- Repeat for $\ell = 0, 1, 2, \ldots$ the following steps:
  — Calculate

$$\begin{aligned} \rho_\ell &= \max\left\{\rho : \Delta_\rho < z_\ell\right\} \\ &= \sum_{i=1}^{n_T} U(z_\ell - \Delta_i) \end{aligned} \tag{32}$$

  where $U(x) = 1$ if $x > 0$ and 0 if $x \leq 0$ is our definition of the unit step function (notice that the function is defined at $x = 0$).
  — Calculate

$$\begin{cases} w_{\ell+1} &= \frac{\rho_\ell}{z_\ell + \sigma_{\rho_\ell}} \\ &= \frac{\sum_{i=1}^{n_T} U(z_\ell - \Delta_i)}{z_\ell + \sum_{i=1}^{n_T} U(z_\ell - \Delta_i)/(P\lambda_i(\mathbf{T}))} \\ z_{\ell+1} &= \sum_{i=1}^{n_R} \frac{1}{w_\ell + 1/\lambda_i(\mathbf{R})}. \end{cases} \tag{33}$$

  — If $\max\{|w_{\ell+1}/w_\ell - 1|, |z_{\ell+1}/z_\ell - 1|\} < \varepsilon$, then exit the repeat loop.

Setting $u_\ell = \ln w_\ell$, we can see that the convergence of the iterative algorithm is equivalent to the convergence of the iterative sequence

$$u_{\ell+1} = \phi(u_{\ell-1}) \tag{34}$$

where

$$\begin{aligned} \phi(u) = \ln &\left[\sum_{i=1}^{n_T} U\left(\sum_{j=1}^{n_R} \frac{1}{\mathrm{e}^u + 1/\lambda_j(\mathbf{R})} - \Delta_i\right)\right] \\ - \ln &\left[\sum_{j=1}^{n_R} \frac{1}{\mathrm{e}^u + 1/\lambda_j(\mathbf{R})} \right. \\ &\left. + \sum_{i=1}^{n_T} \frac{1}{P\lambda_i(\mathbf{T})} U\left(\sum_{j=1}^{n_R} \frac{1}{\mathrm{e}^u + 1/\lambda_j(\mathbf{R})} - \Delta_i\right)\right]. \end{aligned} \tag{35}$$

According to the Contraction Principle [32], this iterative sequence converges provided that $\phi(x)$ is a *contraction*. This property is readily proved as follows.

1) The function $\phi(x)$ is continuous, in spite of the presence of the unit step function $U(\cdot)$. In fact, let us define $\tilde{u}_\rho$ as the unique (since the lhs is monotonically decreasing) solution of the equation[3]

$$\sum_{j=1}^{n_R} \frac{1}{\mathrm{e}^u + 1/\lambda_j(\mathbf{R})} = \Delta_\rho.$$

[3]Since $\Delta_1 = 0$, by definition, we have $\tilde{u}_1 = \infty$.

Then, we can see that

$$\begin{aligned} \phi(\tilde{u}_\rho) &= \ln(\rho - 1) - \ln\left(\Delta_\rho + \sum_{i=1}^{\rho-1} \frac{1}{P\lambda_i(\mathbf{T})}\right) \\ &= \ln[P\lambda_\rho(\mathbf{T})]. \end{aligned}$$

Moreover, for sufficiently small $\epsilon > 0$

$$\begin{aligned} \phi(\tilde{u}_\rho + \epsilon) &= \ln(\rho - 1) - \ln\left(\Delta_\rho - \epsilon' + \sum_{i=1}^{\rho-1} \frac{1}{P\lambda_i(\mathbf{T})}\right) \\ &= \ln[P\lambda_\rho(\mathbf{T})] + \epsilon'' \end{aligned}$$

for some $\epsilon', \epsilon'' > 0$. Finally, we can see that

$$\begin{aligned} \phi(\tilde{u}_\rho - \epsilon) &= \ln(\rho) - \ln\left(\Delta_\rho + \epsilon' + \sum_{i=1}^{\rho} \frac{1}{P\lambda_i(\mathbf{T})}\right) \\ &= \ln[P\lambda_\rho(\mathbf{T})] - \epsilon'' \end{aligned}$$

for some $\epsilon', \epsilon'' > 0$. It is then plain that the potential discontinuities of $\phi(u)$ at $u = \tilde{u}_\rho$ are only apparent and the function is continuous for all $u \in \mathbb{R}$.

2) A continuous function is a contraction if the absolute value of its first derivative is always smaller than 1.
To calculate the derivative $\phi'(u)$, we notice that

$$\begin{aligned} \phi(u) = \ln &\left[\sum_{i=1}^{n_T} U(\tilde{u}_i - u)\right] \\ - \ln &\left[z(u) + \sum_{i=1}^{n_T} \frac{U(\tilde{u}_i - u)}{P\lambda_i(\mathbf{T})}\right] \end{aligned} \tag{36}$$

where

$$z(u) \triangleq \sum_{j=1}^{n_R} \frac{1}{\mathrm{e}^u + 1/\lambda_j(\mathbf{R})}. \tag{37}$$

Then, we have

$$\phi'(u) = \frac{\sum_{i=1}^{n_T} \delta(u - \tilde{u}_i)}{\sum_{i=1}^{n_T} U(\tilde{u}_i - u)} - \frac{z'(u) + \sum_{i=1}^{n_T} \frac{\delta(u - \tilde{u}_i)}{P\lambda_i(\mathbf{T})}}{z(u) + \sum_{i=1}^{n_T} \frac{U(\tilde{u}_i - u)}{P\lambda_i(\mathbf{T})}}$$

where $\delta(\cdot)$ is Dirac's delta function. We notice that, whenever $u = \tilde{u}_\rho$, the diverging contributions of the Dirac's delta functions with 0 argument compensate each other since

$$\sum_{i=1}^{n_T} \frac{U(\tilde{u}_i - \tilde{u}_\rho)}{P\lambda_\rho(\mathbf{T})} = z(\tilde{u}_\rho) + \sum_{i=1}^{n_T} \frac{U(\tilde{u}_i - \tilde{u}_\rho)}{P\lambda_i(\mathbf{T})}$$

and the result is eventually finite. Therefore, we can simplify the above expression as

$$\begin{aligned} \phi'(u) &= -\frac{z'(u)}{z(u) + \sum_{i=1}^{n_T} \frac{U(\tilde{u}_i - u)}{P\lambda_i(\mathbf{T})}} \\ &= \frac{\sum_{j=1}^{n_R} [\mathrm{e}^u + 1/\lambda_j(\mathbf{R})]^{-2} \mathrm{e}^u}{\sum_{j=1}^{n_R} [\mathrm{e}^u + 1/\lambda_j(\mathbf{R})]^{-1} + \sum_{i=1}^{n_T} \frac{U(\tilde{u}_i - u)}{P\lambda_i(\mathbf{T})}} \\ &\leq \frac{\sum_{j=1}^{n_R} [\mathrm{e}^u + 1/\lambda_j(\mathbf{R})]^{-2} \mathrm{e}^u}{\sum_{j=1}^{n_R} [\mathrm{e}^u + 1/\lambda_j(\mathbf{R})]^{-1}} \\ &< 1. \end{aligned} \tag{38}$$

TABLE I
PARAMETERS FOR OSCILLATING EXAMPLE

P	n_T	n_R	$\lambda_1(\mathbf{T})$	$\lambda_2(\mathbf{T})$	$\lambda_3(\mathbf{T})$	ϵ	z_1	z_2
1	3	2	34	11	1.75	0.08	1	$1+\epsilon$

Obviously, we also have $\phi'(u) > 0$. Since the derivative is discontinuous at the points $\tilde{u}_i$, we cannot apply the Mean Value Theorem [32] directly to show that $\phi(u)$ is a contraction. However, since $0 < \phi'(u) < 1$ at all points of continuity, we can apply the Mean Value Theorem to a sequence of open interval as explained in the following. Assume that $a < \tilde{u}_m < \cdots < \tilde{u}_n < b$ are the $(n-m+1)$ points of discontinuity of $\phi'(u)$ falling in the open interval (a,b). We have

$$\begin{aligned}\phi(b)-\phi(a) &= \phi(b)-\phi(\tilde{u}_n)+\cdots+\phi(\tilde{u}_m)-\phi(a)\\ &= (b-\tilde{u}_n)\phi'(\xi_n)+\cdots+(\tilde{u}_m-a)\phi'(\xi_{m-1})\\ &< (b-\tilde{u}_n)\cdot 1+\cdots+(\tilde{u}_m-a)\cdot 1\\ &= b-a\end{aligned}$$

for some $\xi_{m-1} \in (a,\tilde{u}_m),\ldots,\xi_n \in (\xi_n,b)$. Therefore, $\phi(u)$ is a contraction.

This completes the proof of convergence of our proposed iterative water-filling algorithm in the case considered.

On the contrary, the frozen water-filling algorithm [14] does not always converge as illustrated in the Appendix C.

APPENDIX C
OSCILLATING BEHAVIOR OF THE FROZEN WATER-FILLING ALGORITHM

This appendix provides an explicit example of oscillating behavior of the frozen water-filling algorithm. Numerical results show that in this case the proposed water-filling algorithm converges, as expected from Appendix B.

In order to derive this example we consider the parameters defined in Table I.

We will see in the following that the parameters in Table I are chosen in order that:

1) z_ℓ oscillates between the two different values z_1 and z_2, as can be seen from (40);
2) the rank of $\mathbf{Q}_\ell$ oscillates as well, as can be seen from (41);
3) the eigenvalues of $\mathbf{R}$ are real positive numbers.

Now, we recall that the frozen water-filling iterative algorithm requires to solve at each step ℓ the fixed-point equations

$$\begin{cases} w_\ell &= \sum_{i=1}^{n_T} \frac{\lambda_i(\mathbf{TQ}_{\ell-1})}{1+z_\ell\lambda_i(\mathbf{TQ}_{\ell-1})} \\ z_\ell &= \sum_{i=1}^{n_R} \frac{1}{w_\ell+1/\lambda_i(\mathbf{R})} \end{cases} \tag{39}$$

where $\lambda_i(\mathbf{TQ}_\ell) = \lambda_i(\mathbf{T})\lambda_i(\mathbf{Q}_\ell)$. The eigenvalues $\lambda_i(\mathbf{Q}_\ell)$ can be obtained from (28). Setting

$$z_{2\ell+1} \triangleq z_1 \quad \text{and} \quad z_{2\ell} \triangleq z_2 \tag{40}$$

we can see from (30) that

$$\rho_{2\ell+1} = n_T - 1 \quad \text{and} \quad \rho_{2\ell} = n_T. \tag{41}$$

TABLE II
EQUIVALENT PARAMETERS FOR OSCILLATING EXAMPLE

P	n_T	n_R	$\lambda_1(\mathbf{T})$	$\lambda_2(\mathbf{T})$	$\lambda_3(\mathbf{T})$	$\lambda_1(\mathbf{R})$	$\lambda_2(\mathbf{R})$
1	3	2	34	11	1.75	806.1484	2.3988

In fact, a direct computation yields

$$\Delta_2 = 0.0615 < z_1 < \Delta_3 = 1.0225 < z_2.$$

Then, from the first of equations (39) we obtain

$$w_1 \triangleq w_{2\ell+1} = 1.8123 \quad \text{and} \quad w_2 \triangleq w_{2\ell} = 1.6658$$

so that we can rewrite the second of equations (39) as

$$\begin{cases} 1 &= \sum_{i=1}^{n_R} \frac{1}{w_1+1/\lambda_i(\mathbf{R})} \\ 1+\epsilon &= \sum_{i=1}^{n_R} \frac{1}{w_2+1/\lambda_i(\mathbf{R})}. \end{cases} \tag{42}$$

Equations (42) can be rewritten in turn as

$$\begin{cases} \tilde{w}_1 = \sum_{i=1}^{n_R} \frac{1}{1+w_1\lambda_i(\mathbf{R})} \\ \tilde{w}_2 = \sum_{i=1}^{n_R} \frac{1}{1+w_2\lambda_i(\mathbf{R})} \end{cases} \tag{43}$$

where we defined $\tilde{w}_1 \triangleq n_R - w_1$ and $\tilde{w}_2 \triangleq n_R - w_2(1+\epsilon)$. Finally, setting $\Delta \triangleq \det(\mathbf{R})$ and $\tau \triangleq \mathsf{Tr}(\mathbf{R})$ in (43), we obtain

$$\begin{cases} \tilde{w}_1(1+w_1\tau+w_1^2\Delta) = 2+w_1\tau \\ \tilde{w}_2(1+w_2\tau+w_2^2\Delta) = 2+w_2\tau. \end{cases} \tag{44}$$

These equations yield

$$\begin{aligned}\Delta &= \frac{2(w_1(\tilde{w}_1-1)-w_2(\tilde{w}_2-1))}{w_1w_2(\tilde{w}_2w_2(\tilde{w}_1-1)-\tilde{w}_1w_1(\tilde{w}_2-1))}\\ &\quad+\frac{\tilde{w}_1w_2(\tilde{w}_2-1)-\tilde{w}_2w_1(\tilde{w}_1-1)}{w_1w_2(\tilde{w}_2w_2(\tilde{w}_1-1)-\tilde{w}_1w_1(\tilde{w}_2-1))}\\ &= 1933.8\\ \tau &= \frac{2-\tilde{w}_1(1+w_1^2\Delta)}{w_1(\tilde{w}_1-1)}\\ &= 808.5472.\end{aligned}$$

The eigenvalues of $\mathbf{R}$ are then easily obtained as

$$\lambda_{1,2}(\mathbf{R}) = \frac{\tau}{2} \pm \sqrt{\frac{\tau^2}{4}-\Delta}$$

and have the explicit values

$$\lambda_1(\mathbf{R}) = 806.1484 \quad \text{and} \quad \lambda_2(\mathbf{R}) = 2.3988.$$

It is worth noting that we can start equivalently by defining the parameters in Table II. Then, we obtain an oscillating sequence (w_ℓ, z_ℓ), which are solutions of the fixed-point (39), where z_ℓ is still oscillating between $z_{2\ell+1} = 1$ and $z_{2\ell} = 1+\epsilon$.

Finally, we provide further numerical evidence about the convergence of our proposed water-filling algorithm, which was proved in Appendix B, in the specific case illustrated in Table I. Table III shows that, assuming the stopping criterion

$$|\mu_\ell/\mu_{\ell+1} - 1| < 10^{-6}$$

TABLE III
VALUES OF w_ℓ, z_ℓ AND μ_ℓ FOR CONVERGENT ALGORITHM IN APPENDIX B WITH PARAMETERS FROM TABLE II

1	w_ℓ	z_ℓ	μ_ℓ
1	2.073054	1.704540	11.485308
2	1.251936	0.883709	11.744688
3	1.991972	1.397202	11.750967
4	1.436128	0.916839	11.841094
5	1.928343	1.235381	11.850812
6	1.556719	0.944646	11.883770
7	1.877992	1.148555	11.889626
8	1.630165	0.967887	11.902284
9	1.837884	1.101479	11.905266
10	1.672961	0.987243	11.910327
11	1.805764	1.075806	11.911778
12	1.697259	1.003318	11.913843
13	1.779932	1.061762	11.914545
14	1.710853	1.016634	11.915366
15	1.759085	1.054065	11.915714
16	1.718396	1.027644	11.915993
17	1.744801	1.049843	11.916139
18	1.722561	1.035328	11.916222
19	1.737038	1.047526	11.916266
20	1.724856	1.039553	11.916291
21	1.732800	1.046254	11.916304

where μ_ℓ is the ℓth approximation to the mutual information and the initializations $w_0 = z_0 = 1$ and $\mathbf{Q} = \mathrm{diag}(1/2, 1/2)$, 21 iterations are required to achieve convergence.

REFERENCES

[1] A. Abdi and M. Kaveh, "A space-time correlation model for multielement antenna systems in mobile fading channels," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 4, pp. 550–561, Apr. 2002.

[2] D. W. Bliss, K. W. Forsythe, A. O. Hero, III, and A. F. Yegulalp, "Environmental issues for MIMO capacity," *IEEE Trans. Signal Process.*, vol. 50, no. 9, pp. 2128–2142, Sep. 2002.

[3] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York: Cambridge Univ. Press, 2004.

[4] L. E. J. Brouwer, "Über eineindeutige, stetige transformationen von flächen in sich," *Math. Ann.*, vol. 69, pp. 176–180, 1910.

[5] S. Caban and M. Rupp, "Impact of transmit antenna spacing on 2×1 Alamouti radio transmission," *IEEE Electron. Lett.*, vol. 43, no. 4, pp. 198–199, Feb. 2007.

[6] M. Chiani, M. Win, and H. Shin, "Capacity of MIMO systems in the presence of interference," presented at the IEEE Global Communications Conf., Nov. 2006.

[7] M. Chiani, M. Win, and H. Shin, "MIMO networks: The effects of interference," *IEEE Trans. Inf. Theory*, vol. 56, no. 1, pp. 336–349, Jan. 2010.

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

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

[10] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[11] S. N. Diggavi and T. M. Cover, "The worst additive noise under a covariance constraint," *IEEE Trans. Inf. Theory*, vol. 47, no. 7, pp. 3072–3081, Nov. 2001.

[12] P. Driessen and G. J. Foschini, "On the capacity formula for multiple-input multiple-output channels: A geometric interpretation," *IEEE Trans. Commun.*, vol. 47, no. 2, pp. 173–176, Feb. 1999.

[13] J. Dumont, P. Loubaton, and S. Lasaulce, "On the capacity achieving transmit covariance matrices of MIMO correlated Rician channels: A large system approach," presented at the Globecom, San Francisco, CA, Nov. 1, 2006.

[14] 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, Mar. 2010.

[15] F. R. Farrokhi, G. J. Foschini, A. Lozano, and R. A. Valenzuela, "Link-optimal space-time processing with multiple transmit and receive antennas," *IEEE Commun. Lett.*, vol. 5, no. 3, pp. 85–87, Mar. 2001.

[16] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multiple antennas," *Bell Labs Tech. J.*, vol. 1, no. 2, pp. 41–59, 1996.

[17] 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.

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

[19] A. Goldsmith, S. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Select. Areas Commun.*, vol. 21, no. 5, pp. 684–702, Jun. 2003.

[20] L. Greenstein, J. Andersen, H. Bertoni, S. Kozono, D. Michelson, and W. Tranter, "Channel and propagation models for wireless system design I and II," *IEEE J. Select. Areas Commun.*, vol. 20, no. 3/6, Apr./Aug. 2002.

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

[22] D. Hösli, Y.-H. Kim, and A. Lapidoth, "Monotonicity results for coherent MIMO Rician channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4334–4339, Dec. 2005.

[23] R. Horn and C. Johnson, *Matrix Analysis*. New York: Cambridge Univ. Press, 1985.

[24] S. A. Jafar and A. Goldsmith, "Transmitter optimization and optimality of beamforming for multiple antenna systems," *IEEE Trans. Wireless Commun.*, vol. 3, no. 4, pp. 1165–1175, Jul. 2004.

[25] S. K. Jayaweera and H. V. Poor, "On the capacity of multiple-antenna systems in Rician fading," *IEEE Trans. Wireless Commun.*, vol. 4, no. 3, pp. 1102–1111, Mar. 2005.

[26] E. Jorswieck and H. Boche, "Channel capacity and capacity-range of beamforming in MIMO wireless systems under correlated fading with covariance feedback," *IEEE Trans. Wireless Commun.*, vol. 3, no. 5, pp. 1543–1553, Sep. 2004.

[27] Y.-H. Kim and S.-J. Kim, On the Convexity of $\log\det(I + KX^{-1})$ 2006 [Online]. Available: http://arxiv.org/pdf/cs.IT/0611043

[28] S. L. Loyka, "Channel capacity of MIMO architecture using the exponential correlation matrix," *IEEE Commun. Lett.*, vol. 5, no. 9, pp. 369–371, Sep. 2001.

[29] 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, Oct. 2003.

[30] C. Martin and B. Ottersten, "Asymptotic eigenvalue distributions and capacity for MIMO channels under correlated fading," *IEEE Trans. Wireless Commun.*, vol. 3, no. 4, pp. 1350–1359, Jul. 2004.

[31] A. Moustakas, S. Simon, and A. 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.

[32] W. Rudin, *Principles of Mathematical Analysis*, 3rd ed. New York: McGraw-Hill, 1976.

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

[34] H. Shin, M. Z. Win, and M. Chiani, "Asymptotic statistics of mutual information for doubly correlated MIMO channels," *IEEE Trans. Wireless Commun.*, vol. 7, no. 2, pp. 562–573, Feb. 2008.

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

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

[37] G. Taricco, "On the capacity of separately-correlated MIMO Rician fading channels," presented at the GLOBECOM, San Francisco, CA, Dec. 1, 2006.

[38] G. Taricco, "Asymptotic mutual information statistics of separately correlated Rician fading MIMO channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3490–3504, Aug. 2008.

[39] G. Taricco and E. Riegler, "On the ergodic capacity of the asymptotic separately-correlated Rician fading MIMO channel with interference," presented at the IEEE Int. Symp. Information Theory, Nice, France, Jun. 24–30, 2007.

[40] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, Nov.–Dec. 1999.

[41] D. Tse and P. Viswanath, *Fundamentals of Wireless Communications*. New York: Cambridge Univ. Press, 2005.
[42] A. M. Tulino, A. Lozano, and S. Verdú, "Capacity-achieving input covariance for single-user multi-antenna channels," *IEEE Trans. Wireless Commun.*, vol. 5, no. 3, pp. 662–671, Mar. 2006.
[43] S. Venkatesan, S. H. Simon, and R. A. Valenzuela, "Capacity of a Gaussian MIMO channel with nonzero mean," presented at the IEEE Vehicular Technology Conf., Oct. 2003.
[44] E. Visotsky and U. Madhow, "Space-time transmit precoding with imperfect feedback," *IEEE Trans. Inf. Theory*, vol. 47, no. 6, pp. 2632–2639, Sep. 2001.
[45] M. Vu and A. Paulraj, "Capacity optimization for Rician correlated MIMO wireless channels," presented at the Asilomar Conf., Asilomar, CA, Nov. 2005.
[46] J. Winters, "On the capacity of radio communication systems with diversity in a Rayleigh fading environment," *IEEE J. Select. Areas Commun.*, vol. 5, no. 6, pp. 871–878, Jun. 1987.
[47] W. Yu, W. Rhee, S. Boyd, and J. M. Cioffi, "Iterative water-filling for Gaussian vector multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 1, pp. 145–152, Jan. 2004.

Giorgio Taricco (M'91–SM'03–F'10) was born in Torino, Italy. He received the electronic engineer degree (*cum laude*) from the Politecnico di Torino, Torino, Italy, in 1985.

He was a research associate with Italian Telecom Labs from 1985 to 1987, where he was involved in the design of the GSM mobile telephony standard (channel coding). In 1996, he was a Research Fellow at ESTEC, Noordwijk, The Netherlands. He is currently a Full Professor at the Politecnico di Torino. His research interests include error-control coding, multiuser detection, space-time coding, and MIMO communications. He has coauthored around 60 papers in international journals, around 130 international conference contributions, several book chapters, and he holds two international patents.

Prof. Taricco is an Associate Editor for the IEEE TRANSACTIONS ON INFORMATION THEORY and the *Journal on Communications and Networks*. He was the Treasurer of the conferences IEEE ISIT 2000, ISITA 2004, and IEEE ITW 2009.

Erwin Riegler (M'07) was born in Neunkirchen, Austria. He received the Dipl-Ing. degree in Technical Physics (*summa cum laude*) in 2001 and the Dr. techn. degree in Technical Physics (*summa cum laude*) in 2004 from the Vienna University of Technology, Austria.

He visited the Max Planck institute for Mathematics in the Sciences in Leipzig, Germany (September 2004–February 2005). From 2005–2006, he was a postdoctorate at the Institute for Analysis and Scientific Computing, Vienna University of Technology. From 2007–2009, he was a senior researcher at the Telecommunications Research Center Vienna (FTW). He is currently a research assistant at the Institute of Telecommunications, Vienna University of Technology, where he leads the project NOWIRE—Noncoherent Wireless Communications over Doubly Selective Channels. His research interests include noncoherent communications, interference management, large system analysis, and transceiver design.