Mutual Information Optimization and Capacity Evaluation in MIMO Systems with Magnitude Knowledge and Phase Uncertainty

Miquel Payaró[*], Antonio Pascual-Iserte[*,†], and Miguel Ángel Lagunas[*,†]
[†]Universitat Politècnica de Catalunya (UPC), Spain.
[*]Centre Tecnològic de Telecomunicacions de Catalunya (CTTC), Spain.
e-mail: mpayaro@cttc.es, tonip@gps.tsc.upc.es, m.a.lagunas@cttc.es

Abstract

This paper deals with the design of a multi-antenna transmitter in a multi-input multi-output channel to maximize the mutual information and achieve capacity. The channel has been considered as an uncertain system where the magnitude of its gains is perfectly known and there is an incomplete knowledge of the phases. This describes several practical situations, such as time division duplex systems with uncalibrated phases of the radio-frequency chains, or rapidly phase-varying mobile channels. This uncertainty is taken into account explicitly during the design from two different points of view. First, a statistical approach is taken under the objective to achieve ergodic capacity, i.e., to maximize the mutual information averaged over the phases statistics. Secondly, the objective is to achieve compound capacity, i.e., to maximize the worst-case mutual information for any possible phases realization. Since no closed-from solutions exist for these two problems, alternative approximate techniques are proposed in each case, one of them based on random matrix theory. Finally, in the simulations section, the accuracy of the proposed algorithms and approximations are evaluated, jointly with the gains provided by the optimization of the transmitter.

Index Terms

MIMO Capacity, Incomplete Channel State Information, Magnitude Knowledge, Phase uncertainty.

This work was partially supported by the Catalan government under grants 2003FI-00195, SGR2005-00690, and SGR2005-00996; by the European comission under project IST-2002-2.3.1.4 (NEWCOM); by the Spanish government under project TEC2005-08122-C03; and by the european project 2A103 MIMOWA from the MEDEA+ program.
Miquel Payaró is now with the EE Department of HKUST, Hong Kong, China.

I. Introduction

Uncertain systems are characterized by not being modeled perfectly, i.e., some characteristics of the system model are unknown, or known with some errors. The design of signal processing techniques and algorithms are directly linked to the model of the system to which they are to be applied. Consequently, the design of such techniques should take into account explicitly these uncertainties in the model in order to achieve the maximum possible performance from the system exploitation point of view, obtaining the so-called *robust designs* [6], [7], [14], [15], [30].

Although uncertainty arises in many different areas and scenarios, the focus of this paper is on communication systems. To be more specific, in this paper it is considered the communication through a multi-input multi-output (MIMO) channel where multiple antennas are available at both the transmitter and receiver ends. There are many references in the literature showing that this spatial diversity schemes are able to increase importantly the performance of the system and its capacity [24], [5]. Although the performance can be measured in terms of different parameters, in this paper an information theoretic approach is taken, i.e., the performance will be measured using the mutual information [3]. The uncertainty in this kind of scenarios is usually related to the quality and the quantity of the channel state information (CSI) which is available during the design. Note that, while usually the receiver has access to a perfect channel knowledge, this assumption is unrealistic at the transmitter side.

It is widely known that, in case of having additive white Gaussian noise (AWGN) at the receiver, the optimum statistical distribution of the transmit signals is zero-mean Gaussian if the mutual information is to be maximized. Thus, the only remaining degree of freedom when optimizing the system is the transmit covariance matrix. The design of this matrix depends on the quality of the CSI, in other words, it depends on the degree and kind of the uncertainty assumed for the knowledge of the channel at the transmitter side.

There are some works in the literature that look for the optimum transmit signalling scheme for different cases of channel uncertainty. For example, in [25] Telatar conjectured that when the channel is completely unknown to the transmitter, the optimum solution consists in transmitting independent and equal power signals from the different transmit antennas. The optimality of uniform power allocation was also proved in [18] using a game theoretic approach. On the opposite side, when perfect CSI is available, the optimum scheme consists in decomposing the channel matrix into its eigenmodes and transmitting through them independent signals with a power allocation based on the classical water-filling solution [24], [25].

Concerning intermediate cases, i.e., having partial or imperfect CSI, there are also some works that study the optimum transmission strategy to achieve capacity. A paradigmatic example is the case of knowing only the channel statistics, but not the actual realization. According to this, the transmit covariance matrix has to be designed to maximize the mutual information averaged over these statistics, i.e., the so-called *ergodic capacity*. There are several studies for the case of multi-input single-output (MISO) and MIMO channels when the channel follows a Gaussian distribution. More concisely, some works have studied independently the cases of knowing either the mean or the covariance matrix of the channel [11], [13], [17], [28]. More lately, the case of knowing jointly the mean and covariance of the channel has also been addressed in [26]. The solution to this problem is difficult, and although the eigenvectors of the optimum transmitted signals covariance matrix have been found, the eigenvalues have to be calculated using iterative numerical methods.

Another example consists in assuming that the channel is deterministic and that it belongs to a fixed uncertainty region around the channel estimate [19], [29]. In this situation, the objective is to design the transmitter so that the worst-case mutual information for any channel in this uncertainty region is maximized, achieving the so-called *compound capacity* [4].

In this paper, we take a different approach for the channel uncertainty model by considering a more engineering perspective. Here it is considered that the transmitter has perfect knowledge of the magnitude of the complex channel coefficients, but a complete lack of knowledge of the channel phases. This uncertainty model encompasses many practical situations of interest. As an illustrative example, consider the case of time division duplex (TDD) schemes, where the transmitter can estimate the channel during the uplink and use it as it was the same in the downlink thanks to the electromagnetic reciprocity principle. Note, however, that this is not true since the uplink and downlink channels are seen through different RF chains. Although calibration methods are able to compensate the gains, the phases are much more difficult to be estimated and compensated. In this work, a design of the transmitter to achieve the ergodic and compound capacities is proposed under the already mentioned uncertainty model. Unfortunately, closed-form solutions do not exist and approximations have to be applied. For both cases, a numerical optimization method is proposed based on the use of a finite set of random realizations of the channel phases, although it may require a high computational load. Additionally, for the case of the ergodic capacity, an approximate method with lower complexity is also given based on random matrix theory (RMT) [27], whereas for the case of the compound capacity, an approximate finite term Taylor expansion of the mutual information is exploited. Note that one of the major contributions of the present paper with respect to the existing literature, is the application of RMT not only for the evaluation but also for the optimization of the ergodic mutual information for the case of magnitude knowledge and phase uncertainty.

The rest of the paper is organized as follows. The system model is presented in Section II. The problems corresponding to the ergodic and compound capacities, jointly with the proposed algorithms, are addressed in Sections III and IV, respectively. Some simulations results are shown in Section V, whereas Section VI is devoted to the conclusions of this work.

Notation. Boldface upper-case letters denote matrices and boldface lower-case denote column vectors. The inequality $\mathbf{A} \succcurlyeq \mathbf{0}$ means that the matrix $\mathbf{A}$ is positive semi-definite. The superscript $(\cdot)^\dagger$ denotes transpose complex conjugate, $\mathrm{Tr}\,\cdot$ is the trace operator, and $[\cdot]_{ij}$ represents the (i,j)-th entry of a matrix. The Hadamard element-wise matrix multiplication is denoted by $\odot$ and $\mathbb{E}_{\mathbf{X}}$ is the expectation operator with respect to random variable $\mathbf{X}$. Finally, $\sup$ and $\inf$ represent the supremum and the infimum of a function, a.s. denotes almost sure convergence, and s.t. means subject to.

II. System Model

We consider a flat-fading MIMO wireless channel, where the transmitter and the receiver have n_T and n_R antennas respectively. This kind of channel is commonly represented by a matrix $\mathbf{H} \in \mathbb{C}^{n_R \times n_T}$ where the element $[\mathbf{H}]_{ij}$ represents the complex path gain between the j-th transmit and i-th receive antennas.

Let us define $\mathbf{x} \in \mathbb{C}^{n_T}$ as the transmitted signal vector, where $[\mathbf{x}]_j$ represents the signal transmitted through j-th antenna; $\mathbf{y} \in \mathbb{C}^{n_R}$ as the received signal vector, where $[\mathbf{y}]_i$ represents the signal received by i-th antenna; and $\mathbf{n} \in \mathbb{C}^{n_R}$ as the noise vector, where $[\mathbf{n}]_i$ represents the thermal noise that corrupts the received signal at i-th antenna. The entries of the noise vector are considered to be proper complex Gaussian random variables with $\mathbb{E}\mathbf{n} = \mathbf{0}$ and $\mathbb{E}\mathbf{n}\mathbf{n}^\dagger = \sigma^2 \mathbf{I}_{n_R}$.

With these definitions, the input-output relation of this communications system is given by

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

As explained in the introduction, we consider that, while the receiver is fully cognizant of the channel state, the transmitter has only knowledge about the magnitude of the entries of the channel matrix, and a complete lack of knowledge about the actual value of their phases. To separate the known from the unknown part of the channel entries we define

$$[\mathbf{H}]_{ij} = m_{ij}\mathrm{e}^{\mathrm{j}\theta_{ij}}, \quad m_{ij} \in \mathbb{R}^{+} \cup \{0\}, \quad \theta_{ij} \in [0, 2\pi), \tag{2}$$

where θ_{ij} are i.i.d. random variables uniformly distributed in $[0, 2\pi)$ as suggested by, e.g., [12].

By defining the matrices $\mathbf{M}$ and $\mathbf{P}$ such that $[\mathbf{M}]_{ij} = m_{ij}$ and $[\mathbf{P}]_{ij} = \mathrm{e}^{\mathrm{j}\theta_{ij}}$, the uncertainty model described in (2) can be compactly rewritten as

$$\mathbf{H} = \mathbf{M} \odot \mathbf{P}, \tag{3}$$

where we recall that $\odot$ represents the Hadamard element-wise matrix product.

The capacity achieving strategy in MIMO channels with AWGN is signaling using random Gaussian vectors [3]. In this case, the mutual information for a fixed transmit covariance matrix, $\mathbf{Q} = \mathbb{E}\mathbf{x}\mathbf{x}^{\dagger}$, and channel state as defined by $\{\mathbf{M}, \mathbf{P}\}$ is [25]

$$\Psi(\mathbf{Q}, \mathbf{M}, \mathbf{P}) = \log\det\left(\mathbf{I} + \sigma^{-2}(\mathbf{M} \odot \mathbf{P})\mathbf{Q}(\mathbf{M} \odot \mathbf{P})^{\dagger}\right), \tag{4}$$

where the transmitted power is given by $\mathbb{E}\mathbf{x}^{\dagger}\mathbf{x} = \operatorname{Tr}\mathbf{Q} = P$.

In the following sections, we particularize the definitions of ergodic and compound mutual information for the uncertainty model considered in this work and explain in which situations these measures become meaningful and how they can be maximized.

III. Ergodic Mutual Information and Capacity

The information theoretic community defines the ergodic mutual information as the expectation with respect to the channel state uncertainty of the instantaneous mutual information in (4). In our case, since the channel uncertainty is associated with the channel phases $\mathbf{P}$, the expression for the ergodic mutual information particularizes to

$$I_E(\mathbf{Q}, \mathbf{M}) = \mathbb{E}_{\mathbf{P}}\Psi(\mathbf{Q}, \mathbf{M}, \mathbf{P}). \tag{5}$$

Note that, by taking the expectation with respect to $\mathbf{P}$, the ergodic mutual information $I_E(\mathbf{Q}, \mathbf{M})$ does not depend on $\mathbf{P}$.

The ergodic capacity is then defined as the supremum of the ergodic mutual information with respect to the set of possible covariance matrices $\mathbf{Q}$, subject to a mean transmitted power constraint. Formally, this can be written as

$$\begin{aligned} C_E &= \sup_{\mathbf{Q}} I_E(\mathbf{Q}, \mathbf{M}) \\ &\text{s.t. } \operatorname{Tr} \mathbf{Q} \leq P, \ \mathbf{Q} \succcurlyeq \mathbf{0}. \end{aligned} \tag{6}$$

The ergodic capacity is utilized as a measure of the maximum rate that can be achieved in situations where, during the transmission of the message, the magnitude of the channel matrix, $\mathbf{M}$, remains constant while the channel phases, $\mathbf{P}$, vary sufficiently fast so that its long-term properties are revealed.

This model would correspond, for example, to a communication situation with direct line of sight, where the mobile user is moving slowly and there exists moderate to high decoherence of the channel phases as described in [20]. In this scenario, while the magnitude of the entries of $\mathbf{M}$ would not change in an appreciable way, the phases in $\mathbf{P}$ would vary very rapidly because of the relative movement. Some simulations are carried out in Section V to show the validity of this model.

In [21], the authors proved that the optimal covariance matrix that is the solution to the optimization problem described in (6) has to be a diagonal positive semi-definite matrix, which implies that independent power allocation among the transmission antennas is optimal (see Fig. 1). Thus, the problem in (6) simplifies to ⊗1

$$\begin{aligned} C_E &= \sup_{\mathbf{\Lambda}} I_E(\mathbf{\Lambda}, \mathbf{M}) \\ &\text{s.t. } \operatorname{Tr} \mathbf{\Lambda} \leq P, \ [\mathbf{\Lambda}]_{jj} \geq 0, \ [\mathbf{\Lambda}]_{jj'} = 0, \quad 1 \leq j, j' \leq n_T. \end{aligned} \tag{7}$$

Because of the difficulty of dealing with the expectation with respect to the channel phases in the expression for $I_E(\mathbf{\Lambda}, \mathbf{M})$, an expression for the optimal values of the diagonal elements of $\mathbf{\Lambda}$ is only known for the $n_T = n_R = 2$ case, as given in [21]. In the following subsections we present two different methods to numerically calculate the optimal values of the diagonal of $\mathbf{\Lambda}$ for any possible values of n_T and n_R by approximating the function $I_E(\mathbf{\Lambda}, \mathbf{M})$.

A. Approximation by Finite Sample Size

From its expression in (7), it becomes clear that the mutual information maximization problem is a convex optimization problem because it is defined as the supremum of a concave function and the constraints are linear in the design variables, [1]. However, the stochastic nature of the objective function (defined as the expectation of a random quantity) complicates the optimization by making difficult to compute the exact value of the objective function, its gradient, and its Hessian, which are needed at each optimization step.

To overcome this problem, in a practical set-up the expectation operator is approximated by taking the mean of a sufficiently large sample of realizations of the channel phases matrix $\mathbf{P}$ for a fixed magnitude matrix $\mathbf{M}$.

$$I_E(\mathbf{\Lambda}, \mathbf{M}) \approx I_E^{\text{num}}(\mathbf{\Lambda}, \mathbf{M}) = \frac{1}{|\mathcal{M}|} \sum_{\mathbf{P} \in \mathcal{M}} \Psi(\mathbf{\Lambda}, \mathbf{M}, \mathbf{P}), \tag{8}$$

where $\mathcal{M}$ is the set of realizations of the phases matrix $\mathbf{P}$. With the utilization of this approximation, we obtain the following problem

$$\begin{aligned} C_E \approx \sup_{\mathbf{\Lambda}} \; & \frac{1}{|\mathcal{M}|} \sum_{\mathbf{P} \in \mathcal{M}} \Psi(\mathbf{\Lambda}, \mathbf{M}, \mathbf{P}) \\ \text{s.t.} \; & \operatorname{Tr} \mathbf{\Lambda} \le P, \; [\mathbf{\Lambda}]_{jj} \ge 0, \; [\mathbf{\Lambda}]_{jj'} = 0, \quad 1 \le j, j' \le n_T, \end{aligned} \tag{9}$$

where the concavity of the objective function is preserved and, consequently, it is a convex optimization problem. The speed of convergence of the numerical algorithms to solve this kind of problems can be greatly increased by providing the analytical expressions of the gradient vector and the Hessian matrix of the cost function, which are given next.

Since the dependence of $I_E^{\text{num}}(\mathbf{\Lambda}, \mathbf{M})$ on $\mathbf{\Lambda}$ is only through its diagonal elements, we can calculate the gradient of $I_E^{\text{num}}(\mathbf{\Lambda}, \mathbf{M})$ with respect to the vector $\boldsymbol{\lambda}$ which is given by $[\boldsymbol{\lambda}]_j = [\mathbf{\Lambda}]_{jj}$. Defining the vectors $\mathbf{m}_j$ and $\mathbf{p}_j$, with $1 \le j \le n_T$, as the vectors such that $[\mathbf{m}_j]_i = [\mathbf{M}]_{ij}$ and that $[\mathbf{p}_j]_i = [\mathbf{P}]_{ij}$, the gradient can be expressed as [16]

$$[\nabla_{\boldsymbol{\lambda}} I_E^{\text{num}}(\mathbf{\Lambda}, \mathbf{M})]_j = \frac{1}{|\mathcal{M}|} \sum_{\mathbf{P} \in \mathcal{M}} (\mathbf{m}_j \odot \mathbf{p}_j)^\dagger \big(\mathbf{I} + \sigma^{-2} (\mathbf{M} \odot \mathbf{P}) \mathbf{\Lambda} (\mathbf{M} \odot \mathbf{P})^\dagger\big)^{-1} (\mathbf{m}_j \odot \mathbf{p}_j). \tag{10}$$

Similarly, the Hessian matrix $\nabla^2_{\boldsymbol{\lambda}} I_E^{\text{num}}(\mathbf{\Lambda}, \mathbf{M})$ is defined as the matrix whose (j, j')-th element fulfills

$$[\nabla^2_{\boldsymbol{\lambda}} I_E^{\rm num}(\boldsymbol{\Lambda},\mathbf{M})]_{jj'} = \frac{\partial^2 I_E^{\rm num}(\boldsymbol{\Lambda},\mathbf{M})}{\partial\lambda_j\partial\lambda_{j'}} =$$

$$= -\frac{1}{|\mathcal{M}|}\sum_{\mathbf{P}\in\mathcal{M}} |(\mathbf{m}_j\odot\mathbf{p}_j)^\dagger\big(\mathbf{I}+\sigma^{-2}(\mathbf{M}\odot\mathbf{P})\boldsymbol{\Lambda}(\mathbf{M}\odot\mathbf{P})^\dagger\big)^{-1}(\mathbf{m}_{j'}\odot\mathbf{p}_{j'})|^2. \quad (11)$$

In the simulations section we show some ergodic capacity results, which have been obtained by solving the convex optimization problem in (9). We highlight that, although being a convex optimization problem, for large values of n_T and n_R the solution is computationally hard to obtain. In the following section, we present an alternative method to overcome this problem.

B. Optimization Utilizing Results from Random Matrix Theory

Although the method presented in the last section to compute the ergodic mutual information is generic, in the sense that the values of n_T and n_R are arbitrary, and it can achieve any desired accuracy, its main drawback is that it is computationally hard to obtain a solution, which makes difficult its implementation in real-time systems.

We recall that in the previous section we approximated the expectation with respect to the channel phases of the instantaneous mutual information by a mean over a sample of phases realizations, as shown in (8). An alternative approximation for this expectation can be obtained within the framework of RMT. See further [27] for an excellent and exhaustive introduction to RMT with applications to the analysis of wireless communications systems.

One of the objects of study of the theory of random matrices is the eigenvalue probability density of random matrices such as $\mathbf{X}\mathbf{X}^\dagger$, with $\mathbf{X}\in\mathbb{C}^{n_R\times n_T}$ when the number of columns, n_T, and rows, n_R, of the matrix grows without bound but keeping the ratio n_T/n_R held constant, as

$$\lim_{\substack{n_T\to\infty\\ n_R\to\infty}} \frac{n_T}{n_R} = \beta. \quad (12)$$

An interesting result in RMT is that the asymptotic eigenvalue density of $\mathbf{X}\mathbf{X}^\dagger$ is independent of the particular distribution of the entries of $\mathbf{X}$ as long as they are independently distributed. In particular, the entries of $\mathbf{X}$ have to fulfill

$$[\mathbf{X}]_{ij} = \sqrt{\frac{d_{ij}}{n_T}} z_{ij}, \quad (13)$$

with z_{ij} being i.i.d. random variables with $\mathbb{E}z_{11} = 0$ and $\mathbb{E}|z_{11}|^2 = 1$, and some other technical requirements which are out of the scope of the present paper, but are fulfilled by the presented model.

The matrix definition in (13) fits exactly our uncertainty model as we explain in the following. We recall that we are interested in the quantity

$$I_E(\mathbf{\Lambda}, \mathbf{M}) = \mathbb{E}_{\mathbf{P}} \log\det\left(\mathbf{I} + \sigma^{-2}(\mathbf{M} \odot \mathbf{P})\mathbf{\Lambda}(\mathbf{M} \odot \mathbf{P})^\dagger\right). \tag{14}$$

Utilizing the spectral decomposition of the matrix $(\mathbf{M} \odot \mathbf{P})\mathbf{\Lambda}(\mathbf{M} \odot \mathbf{P})^\dagger$ we obtain an equivalent expression for the ergodic mutual information.

$$I_E(\mathbf{\Lambda}, \mathbf{M}) = \mathbb{E}_{\mathbf{P}} \sum_{i=1}^{n_R} \log\left(1 + \sigma^{-2}\zeta_i\right), \tag{15}$$

where ζ_i, with $1 \le i \le n_R$, represents the eigenvalues of the matrix $(\mathbf{M} \odot \mathbf{P})\mathbf{\Lambda}(\mathbf{M} \odot \mathbf{P})^\dagger$. Note that the eigenvalues ζ_i are random quantities that depend on the current realization of the channel phases $\mathbf{P}$. By introducing the spectral measure of $(\mathbf{M} \odot \mathbf{P})\mathbf{\Lambda}(\mathbf{M} \odot \mathbf{P})^\dagger$ as the following random probability measure

$$\mu(\zeta) = \frac{1}{n_R} \sum_{i=1}^{n_R} \delta(\zeta - \zeta_i), \tag{16}$$

the mutual information expression in (15) can be reformulated as

$$I_E(\mathbf{\Lambda}, \mathbf{M}) = \mathbb{E}_{\mathbf{P}} \int_0^\infty \log\left(1 + \sigma^{-2}\zeta\right)\mu(\zeta)\mathrm{d}\zeta, \tag{17}$$

where the expectation with respect to the channel phases now acts upon the probability measure $\mu(\zeta)$. As commented above, the probability measure $\mu(\zeta)$ is the eigenvalue density of the matrix $(\mathbf{M} \odot \mathbf{P})\mathbf{\Lambda}(\mathbf{M} \odot \mathbf{P})^\dagger$. Defining $\mathbf{X} = (\mathbf{M} \odot \mathbf{P})\mathbf{\Lambda}^{1/2}$ we can readily see that

$$\mathbf{X}\mathbf{X}^\dagger = (\mathbf{M} \odot \mathbf{P})\mathbf{\Lambda}(\mathbf{M} \odot \mathbf{P})^\dagger. \tag{18}$$

Now, the entries of $\mathbf{X}$ fulfill that

$$[\mathbf{X}]_{ij} = [\mathbf{M}]_{ij}[\mathbf{\Lambda}^{1/2}]_{jj}\mathrm{e}^{\mathrm{j}\theta_{ij}}, \tag{19}$$

which is the same model as in (13) by identifying

$$\sqrt{\frac{d_{ij}}{n_T}} = [\mathbf{M}]_{ij}[\mathbf{\Lambda}^{1/2}]_{jj}, \tag{20}$$

$$z_{ij} = \mathrm{e}^{\mathrm{j}\theta_{ij}}, \tag{21}$$

and noting that $\mathbb{E}\mathrm{e}^{\mathrm{j}\theta_{ij}} = 0$ and $\mathbb{E}|\mathrm{e}^{\mathrm{j}\theta_{ij}}|^2 = 1$. Consequently the asymptotical limit of the distribution $\mu(\zeta)$ can be computed with the results of RMT. However, we are not interested in the asymptotical limit of the distribution $\mu(\zeta)$ but rather, on a deterministic approximation of it, for finite values of n_T and n_R. Fortunately, such approximation of $\mu(\zeta)$ has been developed in, e.g., [27] and [10], and it is reproduced in the following, utilizing the guidelines described in [9].

First of all, we begin by differentiating and integrating the expression (14) with respect to σ^2 and we obtain

$$I_E(\mathbf{\Lambda}, \mathbf{M}) = \int_{\sigma^2}^{\infty} \left(\frac{1}{\xi} - \mathbb{E}_{\mathbf{P}} \frac{1}{n_R} \operatorname{Tr}\left((\mathbf{M} \odot \mathbf{P}) \mathbf{\Lambda} (\mathbf{M} \odot \mathbf{P})^{\dagger} + \xi \mathbf{I} \right)^{-1} \right) \mathrm{d}\xi \tag{22}$$

We readily see that the second term inside the integral can be equivalently expressed as

$$\frac{1}{n_R} \operatorname{Tr}\left((\mathbf{M} \odot \mathbf{P}) \mathbf{\Lambda} (\mathbf{M} \odot \mathbf{P})^{\dagger} + \xi \mathbf{I} \right)^{-1} = \frac{1}{n_R} \sum_{i=1}^{n_R} \frac{1}{\zeta_i + \xi} = \int \frac{1}{\zeta + \xi} \mu(\zeta) \mathrm{d}\zeta = m_{\mu}(-\xi), \tag{23}$$

where $m_{\mu}(z)$ is the Stieltjes transform of the random measure $\mu(\zeta)$, [27]. Having introduced this transformation, the mutual information expression in (22) becomes

$$I_E(\mathbf{\Lambda}, \mathbf{M}) = \int_{\sigma^2}^{\infty} \left(\frac{1}{\xi} - \mathbb{E}_{\mathbf{P}} m_{\mu}(-\xi) \right) \mathrm{d}\xi. \tag{24}$$

As commented above, the key idea is to replace the expectation of the Stieltjes transform of the random measure $\mu(\zeta)$ by the Stieltjes transform of a deterministic measure $\nu(\zeta)$ in such a way that

$$\lim_{\substack{n_T \to \infty \\ n_R \to \infty}} \mu(\zeta) - \nu(\zeta) = 0, \text{ a.s.,} \quad \text{with } \lim_{\substack{n_T \to \infty \\ n_R \to \infty}} \frac{n_T}{n_R} = \beta. \tag{25}$$

Consequently, we obtain the following approximation:

$$I_E(\mathbf{\Lambda}, \mathbf{M}) \approx \bar{I}_E(\mathbf{\Lambda}, \mathbf{M}) = \int_{\sigma^2}^{\infty} \left(\frac{1}{\xi} - m_{\nu}(-\xi) \right) \mathrm{d}\xi, \tag{26}$$

which fulfills that

$$\lim_{\substack{n_T \to \infty \\ n_R \to \infty}} I_E(\mathbf{\Lambda}, \mathbf{M}) - \bar{I}_E(\mathbf{\Lambda}, \mathbf{M}) = 0, \text{ a.s.,} \quad \text{with } \lim_{\substack{n_T \to \infty \\ n_R \to \infty}} \frac{n_T}{n_R} = \beta. \tag{27}$$

and for any finite values of n_T and n_R, $\bar{I}_E(\mathbf{\Lambda}, \mathbf{M})$ approximates $I_E(\mathbf{\Lambda}, \mathbf{M})$. Unfortunately, the mean and variance of this estimator are not fully characterized yet and its study is a subject of ongoing research.

Now we only need to obtain a procedure to calculate $m_\nu(z)$ as a function of $\mathbf{\Lambda}$ and $\mathbf{M}$. In [27] and [10], the authors found that the Stieltjes transform of the deterministic probability measure $\nu(\zeta)$, fulfilling the property in (25)) is given by

$$m_\nu(z) = \frac{1}{n_R} 1^\dagger \mathbf{t}(z), \tag{28}$$

where $1^\dagger$ is the all-one row vector of the appropriate dimension, and $\mathbf{t}(z)$ is the solution of the following system of equations

$$[\mathbf{t}(z)]_i = \frac{1}{-z\left(1 + n_T^{-1}\tilde{\mathbf{d}}_i^\dagger \tilde{\mathbf{t}}(z)\right)}, \quad \text{for } 1 \le i \le n_R \tag{29}$$

$$[\tilde{\mathbf{t}}(z)]_j = \frac{1}{-z\left(1 + n_T^{-1}\mathbf{d}_j^\dagger \mathbf{t}(z)\right)}, \quad \text{for } 1 \le j \le n_T, \tag{30}$$

where $\mathbf{d}_j \in \mathbb{R}^{n_R}$, with $1 \le j \le n_T$, and $\tilde{\mathbf{d}}_i \in \mathbb{R}^{n_T}$, with $1 \le i \le n_R$ are column vectors with real non-negative entries such that

$$[\mathbf{d}_j]_i = d_{ij} = n_T[\mathbf{M}]_{ij}^2 [\mathbf{\Lambda}]_{jj} \tag{31}$$

$$[\tilde{\mathbf{d}}_i]_j = d_{ij} = n_T[\mathbf{M}]_{ij}^2 [\mathbf{\Lambda}]_{jj}. \tag{32}$$

With these definitions and integrating (26) it is now possible to obtain an expression for the deterministic approximation of the mutual information as

$$\begin{aligned}\bar{I}_E(\mathbf{\Lambda}, \mathbf{M}) = -\sum_{i=1}^{n_R} \log(\sigma^2[\mathbf{t}(-\sigma^2)]_i) - \sum_{j=1}^{n_T} \log(\sigma^2[\tilde{\mathbf{t}}(-\sigma^2)]_j) - \\ - \sigma^2 \sum_{i=1}^{n_R}\sum_{j=1}^{n_T} [\mathbf{t}(-\sigma^2)]_i [\tilde{\mathbf{t}}(-\sigma^2)]_j [\mathbf{M}]_{ij}^2 [\mathbf{\Lambda}]_{jj}. \end{aligned} \tag{33}$$

Although this expression for $\bar{I}_E(\mathbf{\Lambda})$ is rather complicated, it can be numerically evaluated in a very efficient manner. The computationally heavier part is the calculation of $\mathbf{t}$ and $\tilde{\mathbf{t}}$ which has to be done with numerical methods (we utilized the fixed-point technique, which, for our case of concern, is a very efficient and reliable method).

Once we have obtained a method to compute $\bar{I}_E(\mathbf{\Lambda}, \mathbf{M})$, we can now approximate the ergodic capacity as $C_E \approx \bar{C}_E$, where $\bar{C}_E$ is the solution to the following optimization problem

$$\begin{aligned} \bar{C}_E = \sup_{\mathbf{\Lambda}} \bar{I}_E(\mathbf{\Lambda}, \mathbf{M}) \\ \text{s.t. } \operatorname{Tr} \mathbf{\Lambda} \le P, \ [\mathbf{\Lambda}]_{jj} \ge 0, \quad 1 \le j \le n_T, \end{aligned} \tag{34}$$

which is solved numerically utilizing standard methods. We highlight again that the proposed method to approximate C_E is the result from the application of RMT. As a key contribution of this paper, we remark that, differently from previous works where RMT was only utilized to *evaluate* mutual information, we also utilize it to *optimize* the mutual information.

In the simulations section, the goodness of this approximation is validated experimentally, and we also try to give some insight into the implications of RMT in our problem.

IV. Compound Mutual Information and Capacity

The compound mutual information is defined in [4] as the infimum, with respect to the channel state uncertainty, of the mutual information expression in (4). In our case, it particularizes to

$$I_C(\mathbf{Q}, \mathbf{M}) = \inf_{\mathbf{P} \in \mathcal{P}} \Psi(\mathbf{Q}, \mathbf{M}, \mathbf{P}), \tag{35}$$

where $\mathcal{P} \equiv \{\mathbf{X} \in \mathbb{C}^{n_R \times n_T} \mid |[\mathbf{X}]_{ij}| = 1\}$ defines the set of all possible channel phases compatible with the incomplete knowledge about $\mathbf{H}$. It is noteworthy that the compound mutual information does not depend on the statistical properties of the uncertainty in the channel, because it only considers the worst-case scenario.

The compound capacity is then naturally defined as the maximum with respect to the transmit covariance matrix of the compound mutual information, subject to a total mean transmitted power constraint.

$$\begin{aligned} C_C &= \sup_{\mathbf{Q}} I_C(\mathbf{Q}, \mathbf{M}) \\ &\text{s. t. } \operatorname{Tr} \mathbf{Q} \leq P, \ \mathbf{Q} \succcurlyeq \mathbf{0}. \end{aligned} \tag{36}$$

The compound mutual information is a measure of the worst-case achievable rates in situations where no significant channel variability may occur during the transmission of the message and the transmitter is only informed of (or is only able to estimate accurately) the magnitude matrix $\mathbf{M}$. This may be the case of a static communication between the transmitter and the receiver, or when communicating in a slow fading environment.

As formally proven in [23], the compound capacity can also be achieved by a diagonal transmit covariance matrix, i.e., transmitting independent data streams through different antennas with an appropriate power allocation (see Fig. 1). Therefore, the optimization problem can now be written as follows, where the optimization field has been reduced and simplified from a full matrix $\mathbf{Q}$ in (36) to a diagonal matrix $\mathbf{\Lambda}$:

$$\begin{aligned} \bar{C}_C &= \sup_{\mathbf{\Lambda}} I_C(\mathbf{\Lambda}, \mathbf{M}) \\ &\text{s.t. } \operatorname{Tr} \mathbf{\Lambda} \leq P,\ [\mathbf{\Lambda}]_{jj} \geq 0, \quad 1 \leq j \leq n_T. \end{aligned} \tag{37}$$

In [22], a closed-form solution was given to the above problem for the case $n_T = 2$. For other values of n_T we need to utilize the numerical methods described in the following two subsections.

A. Approximation by Finite Sample Size

Similarly as in the ergodic channel capacity presented in the previous section, in this case, the compound mutual information is proposed to be calculated approximately by using a large sample of realizations of the channel phases matrix $\mathbf{P}$ for a fixed magnitude matrix $\mathbf{M}$. Using the same notation as before, let $\mathcal{M}$ be the set of samples of random phases realizations, i.e., $\mathcal{M} = \{\mathbf{P}_1, \ldots, \mathbf{P}_{|\mathcal{M}|}\}$, where k will be used as the index corresponding to the k-th element, i.e., $\mathbf{P}_k$. Using this, the compound mutual information can be approximated as

$$I_C(\mathbf{\Lambda}, \mathbf{M}) \simeq I_C^{\text{num}}(\mathbf{\Lambda}, \mathbf{M}) = \inf_{\mathbf{P} \in \mathcal{M}} \Psi(\mathbf{\Lambda}, \mathbf{M}, \mathbf{P}) = \Psi(\mathbf{\Lambda}, \mathbf{M}, \mathbf{P}_{k_{\min}}), \tag{38}$$

where $k_{\min}$ is the index corresponding to the phase matrix at which the minimum in the expression above is achieved, which depends on $\mathbf{\Lambda}$ for a fixed $\mathbf{M}$, i.e., $k_{\min} = k_{\min}(\mathbf{\Lambda})$. The optimization problem, corresponding to the maximization of $I_C^{\text{num}}(\mathbf{\Lambda}, \mathbf{M})$, can now be written as the following convex optimization problem:

$$\begin{aligned} \bar{C}_C &\simeq \sup_{\mathbf{\Lambda}} \Psi(\mathbf{\Lambda}, \mathbf{M}, \mathbf{P}_{k_{\min}(\mathbf{\Lambda})}) \\ &\text{s.t. } \operatorname{Tr} \mathbf{\Lambda} \leq P,\ [\mathbf{\Lambda}]_{jj} \geq 0, \quad 1 \leq j \leq n_T. \end{aligned} \tag{39}$$

The numerical algorithms for convex problems based, for example, on interior point methods [1], can be applied to obtain the solution of the above problem. The speed of convergence of these kind of procedures can be greatly increased by providing to the algorithm the analytical expressions of the gradient vector and the Hessian matrix of the cost function. In this case, the gradient and Hessian can be directly found by adapting the expressions given in Section III-A to the compound case. This results in the following expressions:

$$[\nabla_{\boldsymbol{\lambda}} I_C^{\text{num}}(\mathbf{\Lambda}, \mathbf{M})]_j = (\mathbf{m}_j \odot \mathbf{p}_j)^\dagger \big(\mathbf{I} + \sigma^{-2}(\mathbf{M} \odot \mathbf{P})\mathbf{\Lambda}(\mathbf{M} \odot \mathbf{P})^\dagger\big)^{-1} (\mathbf{m}_j \odot \mathbf{p}_j)\Big|_{\mathbf{P} = \mathbf{P}_{k_{\min}(\mathbf{\Lambda})}}. \tag{40}$$

Similarly, the Hessian matrix $\nabla^2_{\boldsymbol{\lambda}} I_C^{\text{num}}(\boldsymbol{\Lambda}, \mathbf{M})$ is defined as the matrix whose (j, j')-th element fulfills

$$[\nabla^2_{\boldsymbol{\lambda}} I_C^{\text{num}}(\boldsymbol{\Lambda}, \mathbf{M})]_{jj'} = \\ = - \left|(\mathbf{m}_j \odot \mathbf{p}_j)^\dagger \left(\mathbf{I} + \sigma^{-2}(\mathbf{M} \odot \mathbf{P})\boldsymbol{\Lambda}(\mathbf{M} \odot \mathbf{P})^\dagger\right)^{-1} (\mathbf{m}_{j'} \odot \mathbf{p}_{j'})\right|^2 \Big|_{\mathbf{P}=\mathbf{P}_{k_{\min}(\boldsymbol{\Lambda})}}. \quad (41)$$

B. Approximated Approach Inspired by Taylor Expansion

The previous approach based on the approximation of the minimum of the mutual information over a finite set of realizations of the phase matrix $\mathbf{P}$ may have an unaffordable computational load in order to obtain accurate approximations since the sample size may be required to be too high, specially for moderate and high values of n_T and n_R.

In this subsection a different approach is taken. First, the second order Taylor expansion of the determinant in the expression of the mutual information is calculated and, afterwards, the minimum value of this Taylor expansion over all the possible phases realizations is derived. Once this expression has been found, a convex problem is formulated inspired by the previous result in order to calculate an approximation of the maximum value of the worst-case original mutual information, i.e., the compound capacity. In the cases where the Taylor expansion coincides with the original exact determinant, this convex optimization problem is able to find the compound capacity exactly. This happens when either the number of transmit or receive antennas is equal to 2 or less, i.e., $\min\{n_T, n_R\} \leq 2$. In a general case, the approximation will improve as the terms of order equal to or greater than 3 in the Taylor expansion have a low contribution, which happens, for example, in low SNR conditions. The goodness of this approximation will be evaluated in the simulations section.

Defining $\lambda_j = [\boldsymbol{\Lambda}]_{jj}$, it can be proved that the second order Taylor expansion of the determinant $\det(\mathbf{I} + \sigma^{-2}\mathbf{H}\boldsymbol{\Lambda}\mathbf{H}^\dagger)$, formulated as $\det\left(\mathbf{I} + \sigma^{-2}\mathbf{H}\boldsymbol{\Lambda}\mathbf{H}^\dagger\right) = \mathcal{D}(\boldsymbol{\Lambda}, \mathbf{H}) + o\left(\sum_{j,j'} \lambda_j \lambda_{j'}\right)$, is given by the following function:

$$\mathcal{D}(\boldsymbol{\Lambda}, \mathbf{H}) = 1 + \sigma^{-2} \sum_j \lambda_j \alpha_{jj} + \sigma^{-4} \sum_{j,j'>1} \lambda_j \lambda_{j'} (\alpha_{jj}\alpha_{j'j'} - |\alpha_{jj'}|^2), \quad (42)$$

where $\alpha_{jj'} = \mathbf{h}_j^\dagger \mathbf{h}_{j'}$ and $\mathbf{h}_j$ represents the j-th column of the matrix $\mathbf{H}$. The objective now is to find a lower bound of the function $\mathcal{D}(\boldsymbol{\Lambda}, \mathbf{H})$ over all the possible phases realizations which model the uncertainty in the channel knowledge. In order to do it, note that $|\alpha_{jj'}|^2$ can be upper bounded as

$$\begin{aligned} |\alpha_{jj'}|^2 &= |\mathbf{h}_j^\dagger \mathbf{h}_{j'}|^2 \left(\sum_k h_{kj}^* h_{kj'}\right)\left(\sum_l h_{lj}^* h_{lj'}\right) = \sum_k |h_{kj}|^2 |h_{kj'}|^2 + \sum_{k,l\neq k} h_{kj}^* h_{kj'} h_{lj} h_{lj'}^* \quad (43) \\ &= \sum_k m_{kj}^2 m_{kj'}^2 + \sum_{k,l>k} 2 m_{kj} m_{kj'} m_{lj} m_{lj'} \cos(\theta_{kj} - \theta_{kj'} - \theta_{lj} + \theta_{lj'}) \quad (44) \\ &\leq \sum_k m_{kj}^2 m_{kj'}^2 + \sum_{k,l>k} 2 m_{kj} m_{kj'} m_{lj} m_{lj'} = |\mathbf{m}_j^\dagger \mathbf{m}_{j'}|^2, \quad (45) \end{aligned}$$

where $\mathbf{m}_j$ is the j-th column of the matrix $\mathbf{M}$. Noticing that $\alpha_{jj} = \|\mathbf{h}_j\|^2 = \|\mathbf{m}_j\|^2$, and using the upper bound above, it can be easily seen that

$$\mathcal{D}(\mathbf{\Lambda}, \mathbf{H}) \geq \mathcal{D}(\mathbf{\Lambda}, \mathbf{M}), \quad (46)$$

i.e., the minimum of the second order Taylor expansion of the determinant over all the possible phases corresponds to the evaluation of this Taylor expansion for the concrete case of having zero-phases.

Inspired by this result, in the following it will be assumed that

$$\det(\mathbf{I} + \sigma^{-2}\mathbf{H}\mathbf{\Lambda}\mathbf{H}^\dagger) \gtrsim \det(\mathbf{I} + \sigma^{-2}\mathbf{M}\mathbf{\Lambda}\mathbf{M}^\dagger), \quad (47)$$

which will be more accurate as the terms of order greater than 2 in the Taylor expansion are more negligible. Using this approximation, the compound capacity will now be found numerically as the solution to the following convex optimization problem:

$$\begin{aligned} \bar{C}_C \simeq \sup_{\mathbf{\Lambda}} \ & \log\det(\mathbf{I} + \sigma^{-2}\mathbf{M}\mathbf{\Lambda}\mathbf{M}^\dagger) \\ \text{s.t. } & \operatorname{Tr}\mathbf{\Lambda} \leq P,\ [\mathbf{\Lambda}]_{jj} \geq 0, \quad 1 \leq j \leq n_T. \end{aligned} \quad (48)$$

Similarly as presented above, this convex problem can be solved easily using efficient numerical algorithms based, for example, on the interior point method [1]. Also as noted before, the speed of convergence of these kind of procedures can be greatly increased by providing to the algorithm the analytical expressions of the gradient vector and the Hessian matrix of the cost function. In this case, the gradient and Hessian can be directly found by adapting the expressions given in Section IV-A to the objective function of the problem in (48). This results in the following expressions:

$$[\nabla_{\boldsymbol{\lambda}} \log\det(\mathbf{I} + \sigma^{-2}\mathbf{M}\mathbf{\Lambda}\mathbf{M}^\dagger)]_j = \mathbf{m}_j^\dagger \big(\mathbf{I} + \sigma^{-2}\mathbf{M}\mathbf{\Lambda}\mathbf{M}^\dagger\big)^{-1}\mathbf{m}_j, \quad 1 \leq j \leq n_T. \quad (49)$$

Similarly, the Hessian matrix $\nabla^2_{\boldsymbol{\lambda}} I_C^{\mathrm{num}}(\boldsymbol{\Lambda}, \mathbf{M})$ is defined as the matrix whose (j, j')-th element fulfills

$$[\nabla^2_{\boldsymbol{\lambda}} \log\det(\mathbf{I} + \sigma^{-2}\mathbf{M}\boldsymbol{\Lambda}\mathbf{M}^\dagger)]_{jj'} = -|\mathbf{m}_j^\dagger \left(\mathbf{I} + \sigma^{-2}\mathbf{M}\boldsymbol{\Lambda}\mathbf{M}^\dagger\right)^{-1} \mathbf{m}_{j'}|^2, \quad 1 \le j, j' \le n_T. \quad (50)$$

V. Simulations

In this section, some numerical and simulation results are presented in order to illustrate the concepts that have been presented previously and to evaluate the proposed solutions in terms of achieved performance.

First, in Fig. 2, the accuracy of the approximation of the ergodic mutual information using a sample set of random realizations of the phases matrix is evaluated. This evaluation is performed in terms of the sample variance of the estimation vs. the size of the sample set and for different numbers of transmit and receive antennas. As a general conclusion, it is observed that when the number of random samples and antennas increases, the sample variance decreases, i.e., the quality of the approximates improves, as expected. On the other hand, a qualitative evaluation of the accuracy of the technique based on random matrix theory is shown in Fig. 3. There, the estimated histogram of the eigenvalues of a random MIMO channel (i.e., the eigenvalues of the expression $\mathbf{H}\mathbf{H}^\dagger$) for different numbers of transmit and receive antennas is presented jointly with the asymptotic non-random probability density function $\nu(\zeta)$ derived from RMT. As expected, as the size of the MIMO channel increases, the accuracy gets better. Note, however, that even in the case of a low number of antennas, the asymptotic distribution is able to reproduce the mean behavior of the random distribution. ⊗2 ⊗3

Fig. 4 shows again the random histogram of the eigenvalues and the asymptotic non-random probability density function $\nu(\zeta)$ derived from RMT for two different situations. First, the situation corresponding to a transmitter in which a uniform power allocation is used, i.e., a situation in which the knowledge of the modulus of the channel gains is not exploited (eigenvalues of $\mathbf{H}\mathbf{I}\mathbf{H}^\dagger$). Secondly, the case where the power allocation is derived so that the ergodic mutual information is maximized, i.e., the ergodic capacity is achieved (eigenvalues of $\mathbf{H}\boldsymbol{\Lambda}\mathbf{H}^\dagger$, with $\boldsymbol{\Lambda}$ containing the optimized power allocation). As a conclusion, it is shown that thanks to the optimization of the power allocation, the distribution of the eigenvalues is shifted to the right and, therefore, the mutual information increases. ⊗4

A more concrete numerical evaluation of the ergodic capacity is shown in Figs. 5 and 6. ⊗5,6

From the first one it is concluded that both the numerical algorithm based on the sample set of random realizations of the phases matrix and the algorithm derived from the application of random matrix theory achieve almost the same ergodic capacity. This suggests to choose the second technique, since it has a much lower computational load than the first one. In the second figure, i.e., Fig. 6, the optimized power allocation is compared with the uniform power allocation, that is, that corresponding to the complete uncertainty of both the magnitude and the phases of the channel gains. As seen in the figure, the gain provided by the optimized power allocation is specially important at low SNR, since at high SNR, the optimized power allocation tends to be the uniform power allocation, as expected.

As it has been explained in Section III, one of the situations that motivate the ergodic capacity study corresponds to a communication scenario with a line of sight where the mobile user is moving slowly. In this case, if perfect wave propagation is assumed, the entries of the phases matrix $\mathbf{P}$ are not independent. However, in the ergodic capacity results presented above, it has been assumed that the channel phases are independently distributed. Our assumption is justified by noting that in actual scenarios, perfect wave propagation is just an idealization and, in practice, there exist wave propagation phenomena which lead to the loss of phase coherence and consequently also lead to phase independence. The phase decoherence model presented in [2], [20] is the one that has been utilized to validate the results obtained for the ergodic capacity. In this mentioned model the decoherence among channel phases is controlled with the parameter σ_ϕ^2. If $\sigma_\phi^2 = 0$, then we recover the perfect wave propagation model. In Fig. 7, we have depicted $\otimes$7
the ergodic capacity as a function of the system mean SNR, utilizing the wave propagation model described in [20] for different values of the decoherence parameter σ_ϕ^2. It can be seen that, for $\sigma_\phi^2 = 0$ there is a mismatch between the actual ergodic capacity and the one obtained with the independent phases assumption. As the decoherence parameter σ_ϕ^2 increases the mismatch is greatly reduced. As an indicator, a value of $\sigma_\phi^2 = 0.25$ is appropriate to describe practical line of sight scenarios as given by [8]. In Fig. 7 it can be seen that for $\sigma_\phi^2 = 0.25$ there is a good match between the capacity results obtained utilizing the wave propagation model and assuming independent phases.

Concerning the simulations corresponding to the compound capacity, some results are given in the remaining figures. Fig. 8 also evaluates the sample variance of the compound capacity $\otimes$8
corresponding to the application of a sample set of random realizations of the phases matrix.

The same conclusion as in Fig. 2 can be obtained. Note, however, that the sample variances obtained in the case of the compound capacity are approximately an order of magnitude higher than in the case of the ergodic capacity. Finally, the same conclusions as before are derived from the observation of Figs. 9 and 10. Note that in Fig. 9 the techniques that are compared $\otimes$9,10 are the one based on the application of the sample set of random realizations of the phases matrix and the algorithm inspired by the second order Taylor expansion of the determinant in the expression of the mutual information. As previously, both algorithms provide almost the same result and, therefore, it is suggested to use the second one, since it has a much lower computational complexity.

VI. Conclusions

In this paper, the problem of the design of a transmitter in a MIMO channel to achieve capacity has been addressed. More concisely, the MIMO channel has been considered as a system with uncertainty in the channel knowledge in terms of known magnitude of the channel gains and completely unknown phases. The design problem has been derived from two different perspectives that deal with the uncertainty in the system through two different approaches. On the one hand, in the ergodic capacity case, the uncertainty has been addressed from a statistical point of view, i.e., the capacity has been calculated as the average of the mutual information over the statistics of the unknown phases. On the other hand, in the compound capacity case, the capacity has been found as the infimum of the mutual information over all the possible phases realizations.

The main problem when dealing with the two designs of the transmit covariance matrix to achieve ergodic and compound capacities is that closed-form expressions do not exist. This paper has then focused the attention on the application of numerical methods to find the optimal transmission signaling schemes that, in both cases, can be implemented by transmitting independent data streams through different antennas with an appropriate power allocation. In both cases, two different numerical approaches have been proposed. One of them is based on a set of random samples of the phases matrix, which are used to approximate the ergodic and compound mutual informations. For the case of the ergodic capacity, an additional technique based on the tools provided by RMT is also presented, which reduces drastically the computational complexity. This reduction in the computational load can also be achieved in the compound capacity problem by using an approximation of the worst-case mutual information inspired by the minimization of the second order Taylor expansion of the determinant in expression of the mutual information.

Finally, the simulations have been useful to show two different aspects. First, the accuracy of the proposed approximated solutions has been tested, showing that the low complexity techniques provide nearly the same results as the high load algorithms. Finally, the ergodic and compound capacities themselves have been evaluated for different scenarios, leading to the conclusion that the optimization of the power allocation with only the knowledge of the magnitude of the channel gains allows to improve the system performance importantly.

REFERENCES

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

[2] W. Carey, "Measurement of down-slope sound propagation from a shallow source to a deep ocean receiver," *The Journal of the Acoustical Society of America*, vol. 79, p. 49, 1986.

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

[4] I. Csiszar and J. Korner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. New York: Academic Press, 1981.

[5] G. Foschini and M. Gans, "On limits of wireless communications in a fading environtment when using multiple antennas," *Wireless Personal Communications*, vol. 6, no. 3, pp. 311–335, Mar. 1998.

[6] H. Gao, J. Lam, and C. Wang, "Robust energy-to-peak filter design for stochastic time-delay systems," *Systems & control letters*, vol. 55, no. 2, pp. 101–111, 2006.

[7] H. Gao, J. Lam, L. Xie, and C. Wang, "New approach to mixed H_2/H_∞ filtering for polytopic discrete-time systems," *IEEE Trans. on Signal Processing*, vol. 53, no. 8, pp. 3183–3192, 2005.

[8] A. B. Gershman, C. F. Mecklenbräuker, and J. F. Bohme, "Matrix fitting approach to direction of arrival estimation with imperfect spatial coherence of wavefronts," *IEEE Trans. on Signal Processing*, vol. 45, no. 7, pp. 1894–1899, 1997.

[9] W. Hachem, "Random matrices and the Stieltjes transform. MIMO channels capacity approximation," *Lecture notes of the NEWCOM Summer School on Random Matrix Theory for Wireless Communications*, July 2006.

[10] W. Hachem, P. Loubaton, and J. Najim, "Deterministic equivalents for certain functionals of large random matrices," *http://arxiv.org/abs/math.PR/0507172*, 2005.

[11] S. Jafar, S. Vishwanath, and A. Goldsmith, "Channel capacity and beamforming for multiple transmit and receive antennas with covariance feedback," in *Proc. IEEE Intl. Conf. Commun. (ICC)*, June 2001.

[12] W. Jakes, Jr., *Microwave Mobile Communications*. New York: John Wiley & Sons, Inc., 1974.

[13] E. Jorswieck and H. Boche, "Optimal transmission with imperfect channel state information at the transmit antenna array," *Wireless Personal Communications*, vol. 27, pp. 33–56, 2003.

[14] S. A. Kassam and H. V. Poor, "Robust techniques for signal processing: A survey," *Proceedings of the IEEE*, vol. 73, no. 3, pp. 433–481, Mar. 1985.

[15] H. Li and M. Fu, "A linear matrix inequality approach to robust H_∞ filtering," *IEEE Trans. on Signal Processing*, vol. 45, no. 9, pp. 2338–2350, 1997.

[16] J. R. Magnus and H. Neudecker, *Matrix Differential Calculus with Applications in Statistics and Econometrics*. John Wiley & Sons, 1999.

[17] A. Narula, M. Lopez, M. Trott, and G. Wornell, "Efficient use of side information in multiple-antenna data transmission over fading channels," *IEEE Journal on Selected Areas in Communications*, vol. 16, pp. 1423–1436, Oct. 1998.

[18] D. P. Palomar, J. M. Cioffi, and M. A. Lagunas Hernández, "Uniform power allocation in MIMO channels: A game-teoretic approach," *IEEE Trans. on Information Theory*, vol. 49, no. 7, pp. 1707–1727, July 2003.

[19] A. Pascual Iserte, D. P. Palomar, A. I. Pérez Neira, and M. A. Lagunas Hernández, "A robust maximin approach for MIMO communications with imperfect channel state information based on convex optimization," *IEEE Trans. on Signal Processing*, vol. 54, no. 1, pp. 346–360, Jan. 2006.

[20] A. Paulraj and T. Kailath, "Direction of arrival estimation by eigenstructure methods with imperfect spatial coherence of wave fronts," *The Journal of the Acoustical Society of America*, vol. 83, p. 1034, 1988.

[21] M. Payaró, X. Mestre, and M. A. Lagunas, "Optimum transmit architecture of a MIMO system under modulus channel knowledge at the transmitter," in *Proc. IEEE Information Theory Workshop (ITW'04)*, Oct. 2004.

[22] M. Payaró, X. Mestre, A. Pérez-Neira, and M. A. Lagunas, "Robust power allocation techniques for MIMO systems under modulus channel knowledge at the transmitter," in *Proc. IEEE Sig. Proc. Advances in Wireless Commun. (SPAWC'05)*, June 2005.

[23] M. Payaró, A. Wiesel, J. Yuan, and M. A. Lagunas, "On the capacity of linear vector Gaussian channels with magnitude knowledge and phase uncertainty," in *Proc. IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing (ICASSP'06)*, May 2006.

[24] G. G. Raleigh and J. M. Cioffi, "Spatio-temporal coding for wireless communication," *IEEE Trans. on Communications*, vol. 46, no. 3, pp. 357–366, Mar. 1998.

[25] I. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. Telecommunications*, vol. 10, no. 6, pp. 585–595, Nov.-Dec. 1999.

[26] A. M. Tulino and A. Lozano and S. Verdú, "Capacity-achieving input covariance for single-user multi-antenna channels," *IEEE Trans. on Wireless Communications*, vol. 5, no. 3, pp. 662–671, Mar. 2005.

[27] A. Tulino and S. Verdú, *Random Matrices and Wireless Communications. Foundations and Trends in Communications and Information Theory*, 1st ed. NOW Publishers, 2004.

[28] E. Visotsky and U. Madhow, "Space-time transmit precoding with imperfect feedback," *IEEE Trans. on Information Theory*, vol. 47, pp. 2632–2639, Sep. 2001.

[29] A. Wiesel, Y. C. Eldar, and S. Shamai, "Beamforming maximizes the rank one Ricean MIMO compound capacity," in *Proc. IEEE Signal Processing Advances in Wireless Communications (SPAWC'05)*, June 2005.

[30] K. Zhou, J. C. Doyle, and K. Glover, *Robust and Optimal Control*. Upper Saddle River, NJ: Prentice Hall International, 1996.

List of Figures

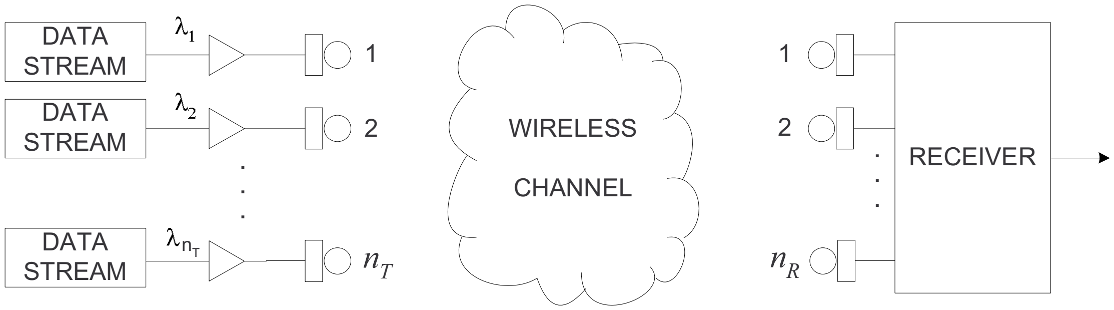

Fig. 1. Optimum transmitter architecture for the case of complete phase unknowledge of the channel gains. Both for the cases of ergodic and compound capacities, the optimum signalling consists in transmitting independent signals through different transmit antennas with an appropriate power allocation.

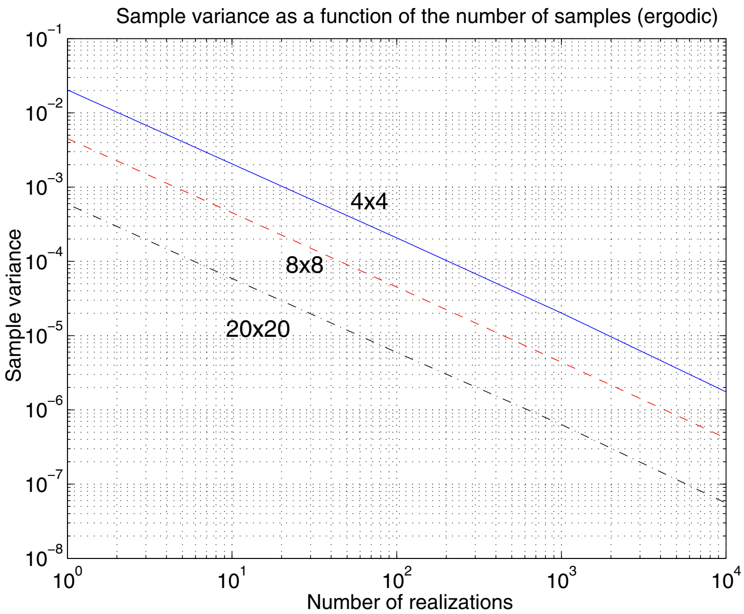

Fig. 2. Sample variance evaluation of the ergodic mutual information when using a set of random realizations of the phases matrix. Different numbers of random realizations of the phases matrix and numbers of transmit and receive antennas have been evaluated.

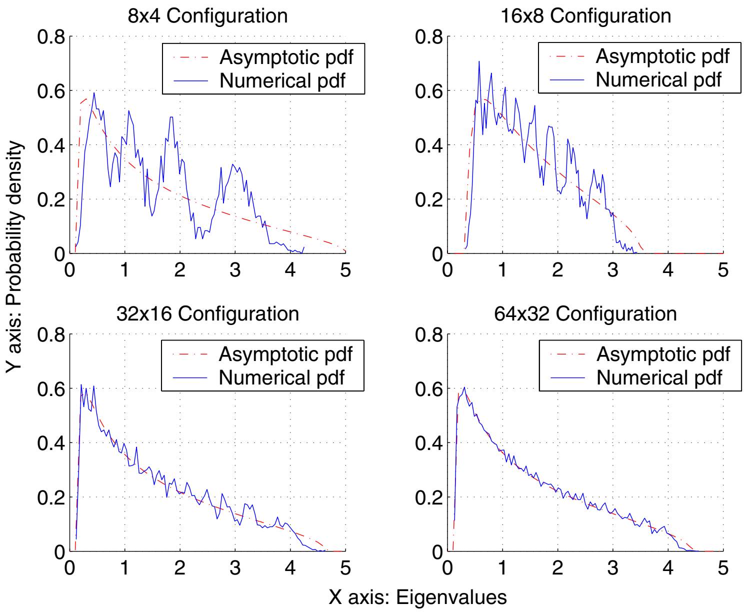

Fig. 3. Comparison between the numerical and, therefore, random histogram of the eigenvalues of a random channel, i.e., of the expression $\mathbf{H}\mathbf{H}^{\dagger}$, and the asymptotic non-random probability density function $\nu(\zeta)$ given by random matrix theory for different numbers of transmit and receive antennas.

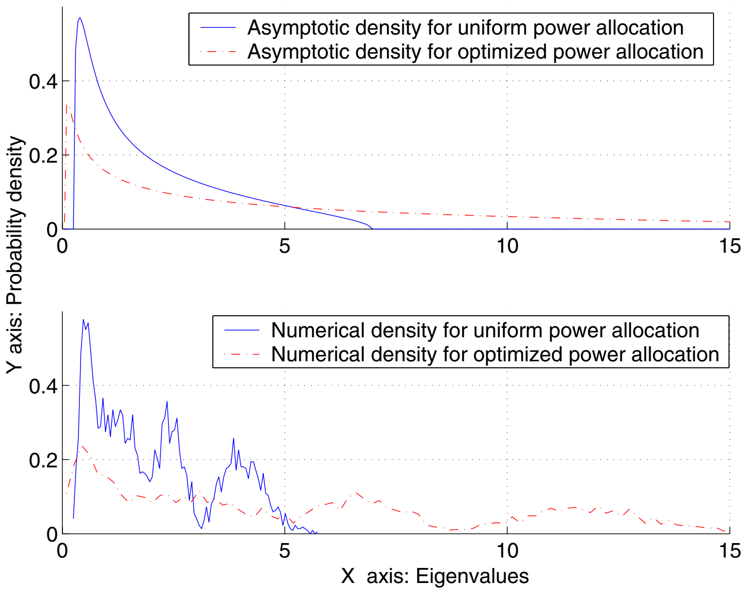

Fig. 4. Comparison between the numerical and the asymptotic probability density functions of the eigenvalues resulting from uniform power allocation (i.e., of the expression $\mathbf{H}\mathbf{I}\mathbf{H}^{\dagger}$) and the optimized power allocation to achieve ergodic capacity (i.e., of the expression $\mathbf{H}\mathbf{\Lambda}\mathbf{H}^{\dagger}$).

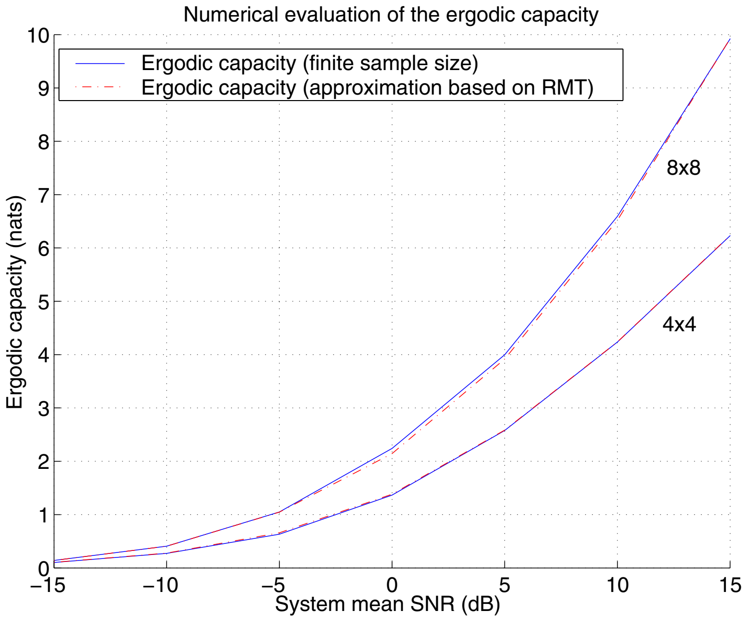

Fig. 5. Evaluation of the optimized power allocation in terms of ergodic mutual information using the approximated algorithm with a finite sample set of random realization of the phases matrix and the algorithm based on random matrix theory.

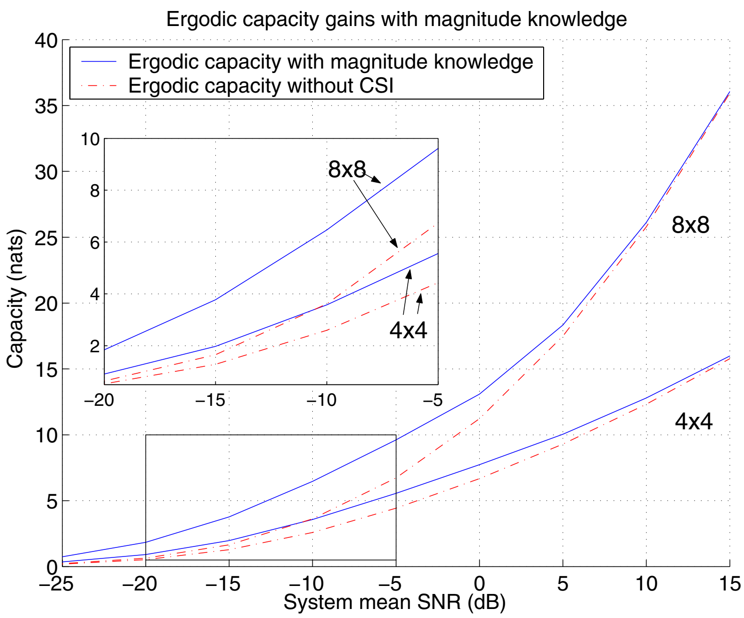

Fig. 6. Comparison of the ergodic capacity that can be achieved when the transmitter has magnitude knowledge and phase uncertainty using the optimized power allocation with the capacity that can be achieved when the transmitter has no CSI, i.e., with uniform power allocation.

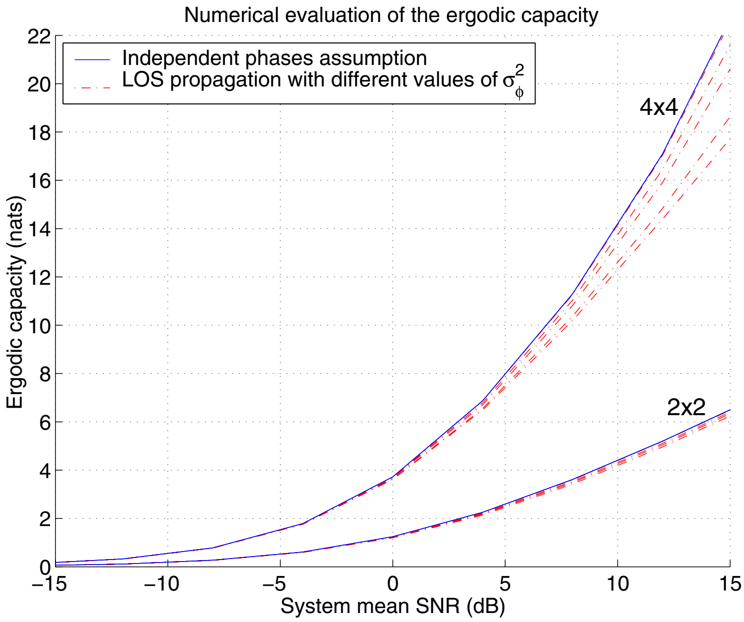

Fig. 7. Numerical evaluation of the validity of the independent phases assumption. In blue, we have plotted the ergodic capacity assuming that the channel phases are independently distributed. In dashed red we have plotted the ergodic capacity utilizing the model described in [20] for the variance values $\sigma_\phi^2 = [0, 0.05, 0.1, 0.25, 0.5]$. The higher the values of σ_ϕ^2, the closer to the blue plot.

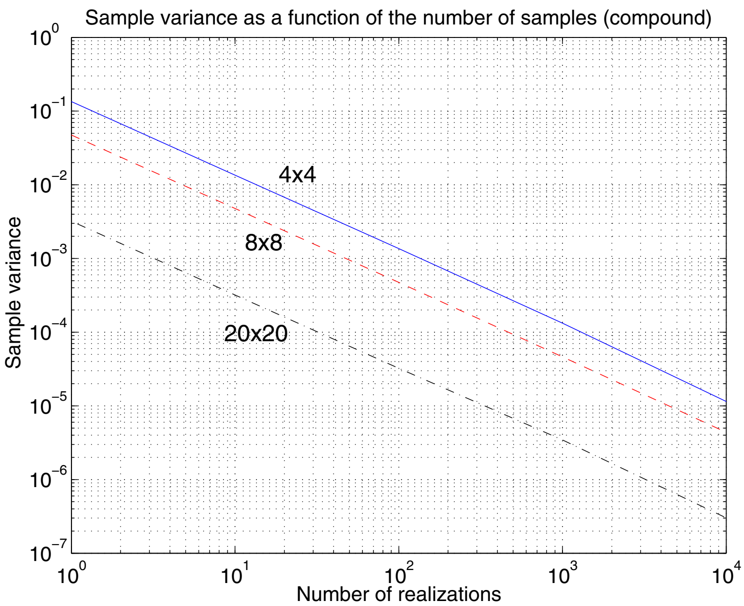

Fig. 8. Sample variance evaluation of the compound mutual information when using a set of random realizations of the phases matrix. Different numbers of random realizations of the phases matrix and numbers of transmit and receive antennas have been evaluated.

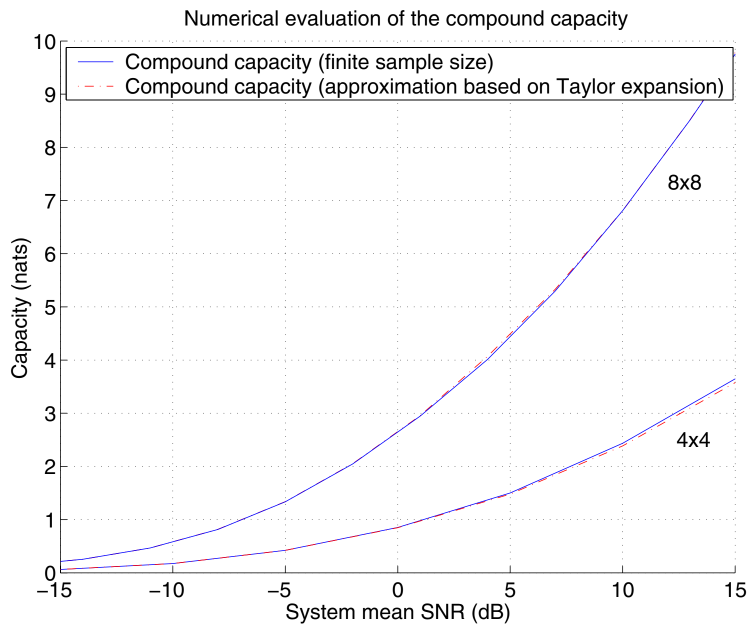

Fig. 9. Evaluation of the optimized power allocation in terms of compound mutual information using the approximated algorithm with a finite sample set of random realization of the phases matrix and the algorithm inspired by the second order Taylor expansion of the determinant in the expression of the mutual information.

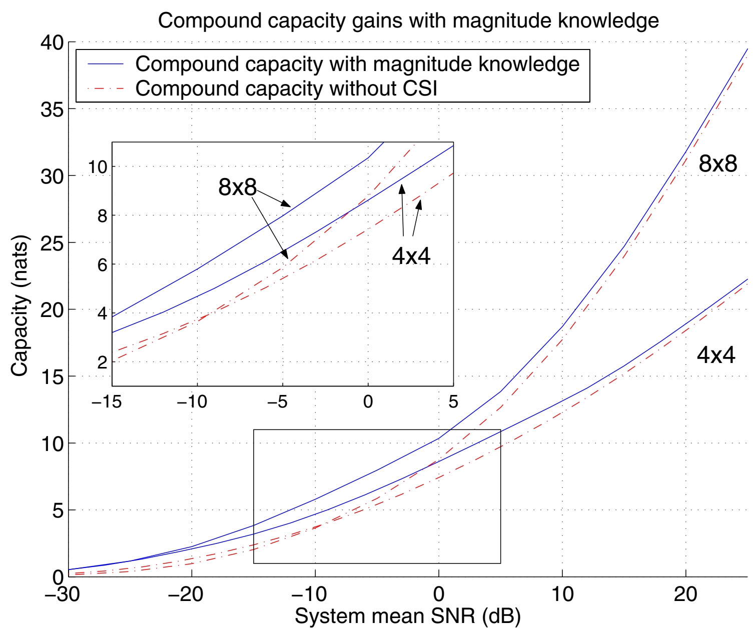

Fig. 10. Comparison of the ergodic capacity that can be achieved when the transmitter has magnitude knowledge and phase uncertainty using the optimized power allocation with the capacity that can be achieved when the transmitter has no CSI, i.e., with uniform power allocation.