

ON THE ERGODIC CAPACITY OF FREQUENCY SELECTIVE MIMO SYSTEMS EQUIPPED WITH MMSE RECEIVERS: AN ASYMPTOTIC APPROACH.

C. Artigue^{1,2}, *P. Loubaton*¹, *B. Mouhouche*²

¹ Université Paris-Est, IGM LabInfo, UMR CNRS 8049

5 Bd. Descartes, Cité Descartes, 77454 Marne la Vallée Cedex 2

² Freescale Semiconductor,

134 Avenue du General Eisenhower - B.P. 29, 31023 Toulouse cedex 1, France

ABSTRACT

This paper is devoted to the study of the ergodic capacity of frequency selective MIMO systems equipped with a MMSE receiver when the channel state information is available at the receiver side and when the second order statistics of the channel taps are known at the transmitter side. As the expression of this capacity is rather complicated and difficult to analyse, it is studied in the case where the number of transmit and receive antennas converge to $+\infty$ at the same rate. In this asymptotic regime, the main results of the paper are related to the design of an optimal precoder in the case where the transmit antennas are correlated. It is shown that the left singular eigenvectors of the optimum precoder coincide with the eigenvectors of the mean of the channel taps transmit covariance matrices, and its singular values are solution of a certain maximization problem.

1. INTRODUCTION

It is now well established that using multiple transmit and receive antennas potentially allows to increase the Shannon capacity of digital communications systems. Since the seminal work of Teletar ([8]), the ergodic Shannon capacity of block fading MIMO systems has been studied extensively. If the channel state information is available at the receiver side while the transmitter is only aware of its second order statistical properties, important questions such as the impact of channel correlations on the capacity or the design of optimal precoding schemes have been addressed by several authors (see e.g. [3] for a review).

The ergodic Shannon capacity is certainly a valuable figure of merit if the MIMO system under consideration is equipped with a maximum likelihood decoder. As this receiver may be difficult to use in practice, in particular if the MIMO channel is frequency selective, it is also quite relevant to study the potential performance of MIMO systems equipped with suboptimum linear interfaces. In this paper, we consider the MMSE receiver, and study the corre-

sponding ergodic capacity denoted C_{mmse} in the following. It is defined as the sum over the transmit antennas of the terms $\mathbb{E}(\log(1 + \text{SINR}_j))$, where SINR_j represents the output MMSE SINR associated to the stream sent by antenna j . C_{mmse} has been studied extensively in the past if the channel is static and available at the transmitter; in this context, the design of an optimum precoder has been addressed in a number of papers (see e.g. [7]). However, assuming that the transmitter is aware of the channel state information is often not realistic in the context of mobile wireless systems. In this paper, we assume that the MIMO channel is frequency selective, known at the receiver side, but that its second order statistics are available at the transmitter side. As in [2], the channel taps are modelled as independent Gaussian random matrices with possible transmit correlations. To our best knowledge, C_{mmse} seems not to have been studied extensively in the past.

As the expression of C_{mmse} is a rather complicated, we evaluate its behaviour in the case where the number of transmit and receive antennas converge to $+\infty$ at the same rate. In this asymptotic regime, C_{mmse} has the same behaviour that a simpler term \bar{C}_{mmse} . As observed in a number of situations, \bar{C}_{mmse} is a reliable approximation of C_{mmse} , even for a realistic number of transmit and receive antennas. We address the optimization of \bar{C}_{mmse} w.r.t. linear precoders, and show that the left singular vectors of the optimum precoder(s) coincide with the eigenvectors of a covariance matrix depending on the second order statistics of the channel taps. The singular values are moreover solution of a certain optimization problem.

2. ASYMPTOTIC EXPRESSION OF C_{MMSE} .

We consider a MIMO system equipped with r receive antennas and t transmit antennas. The transfer function $\mathbf{H}(z)$ of the discrete-time equivalent MIMO channel is given by $\mathbf{H}(z) = \sum_{l=0}^{L-1} \mathbf{H}_l z^{-l}$. The channel taps $(\mathbf{H}_l)_{l=0, \dots, L-1}$ are assumed to be mutually independent Gaussian random

matrices. As in [2]¹, each matrix \mathbf{H}_l is modelled as

$$\mathbf{H}_l = p^{1/2} \frac{1}{\sqrt{L}} \bar{\mathbf{H}}_l \mathbf{C}_l^{1/2} \quad (1)$$

where p represents the total received power per receive antenna and where $\bar{\mathbf{H}}_l$ is a zero mean complex Gaussian matrix with unit variance independent identically distributed (i.i.d.) entries. Matrix \mathbf{C}_l is a positive matrix modelling the impact of transmit antennas correlation on the l -th path. $p \frac{1}{L} \frac{1}{t} \text{Tr}(\mathbf{C}_l)$ represents the power of path l . If we denote \mathbf{C} the matrix

$$\mathbf{C} = \frac{1}{L} \sum_{l=0}^{L-1} \mathbf{C}_l$$

we assume that $\mathbf{C} > 0$ and that $\frac{1}{t} \text{Tr} \mathbf{C} = 1$. The normalizations ensure that the total received power is independent of L and t and thus allow to compare in a fair fashion situations in which the number of paths and the number of transmit antennas differ. Model (1) implies that $\mathbf{H}(e^{2i\pi f})$ can be written as

$$\mathbf{H}(e^{2i\pi f}) = p^{1/2} \tilde{\mathbf{H}}(e^{2i\pi f}) \mathbf{C}^{1/2} \quad (2)$$

where matrix $\tilde{\mathbf{H}}(e^{2i\pi f})$ is a zero mean Gaussian $r \times t$ random matrix with variance $\frac{1}{t}$ i.i.d. entries. This is based on the observation that the covariance matrix of $\text{vec}(\mathbf{H}(e^{2i\pi f}))$ coincides with $p/t \mathbf{C} \otimes \mathbf{I}_r$ (see [2] for more details). This means that for each f , $\mathbf{H}(e^{2i\pi f})$ can be interpreted as the channel matrix of a semi-correlated flat fading MIMO channel whose transmit correlation matrix is independent of the frequency f .

The transmitter sends t streams of symbols $((s_j(n))_{n \in \mathbb{Z}})_{j=1, \dots, t}$ (one stream per transmit antenna). In this section, we assume that the $(s_j)_{j=1, \dots, t}$ are unit variance mutually independent i.i.d. sequences. The corresponding r -variate discrete-time received signal $(\mathbf{y}(n))_{n \in \mathbb{Z}}$ is given by

$$\mathbf{y}(n) = \sum_{l=0}^{L-1} \mathbf{H}_l s(n-l) + \mathbf{v}(n)$$

where $s(n) = (s_1(n), \dots, s_t(n))^T$ and where \mathbf{v} is a white Gaussian noise with covariance matrix $\mathbb{E}(\mathbf{v}(n)\mathbf{v}(n)^H) = \sigma^2 \mathbf{I}_r$. Each symbol sequence s_j is estimated by the non causal Wiener filter whose transfer function is

$$\mathbf{h}_j(e^{2i\pi f})^H (\mathbf{H}(e^{2i\pi f}) \mathbf{H}(e^{2i\pi f})^H + \sigma^2 \mathbf{I}_r)^{-1}$$

where $\mathbf{h}_j(e^{2i\pi f})$ is the j -th column of $\mathbf{H}(e^{2i\pi f})$. It is standard that the SINR β_j provided by this linear receiver is

¹[2] addresses uplink transmissions so that the receive antennas are correlated. Here, we rather focus on downlink transmissions and the transmit antennas are correlated

given by

$$\beta_j = \frac{1}{\int_{-1/2}^{1/2} (\mathbf{Q}(e^{2i\pi f}))_{j,j} df} - 1 \quad (3)$$

where matrix $\mathbf{Q}(e^{2i\pi f})$ is defined by

$$\mathbf{Q}(e^{2i\pi f}) = \frac{\sigma^2}{p} \left(\mathbf{H}(e^{2i\pi f})^H \mathbf{H}(e^{2i\pi f}) + \frac{\sigma^2}{p} \mathbf{I} \right)^{-1}$$

The ergodic capacity \mathcal{C}_{mmse} of the MIMO system under consideration is thus equal to

$$\mathbb{E} \sum_{j=1}^t (\log_2(1 + \beta_j)) = -\mathbb{E} \sum_{j=1}^t \log_2 \left(\int_{-1/2}^{1/2} (\mathbf{Q}(e^{2i\pi f}))_{j,j} df \right) \quad (4)$$

where the mathematical expectation is over the joint probability distribution of random matrices $\mathbf{H}_0, \dots, \mathbf{H}_{L-1}$. It is clear the expression of \mathcal{C}_{mmse} is complicated and difficult to exploit.

When r and t converge to $+\infty$ at the same rate, the entries of $\mathbf{Q}(e^{2i\pi f})$ converge towards deterministic terms which only depend on the statistical properties of $\mathbf{H}(e^{2i\pi f})$. By (2), the second order statistical properties of $\mathbf{H}(e^{2i\pi f})$ are independent of the frequency f . Therefore, the large system approximation of the entries $\mathbf{Q}(e^{2i\pi f})$ are independent of f . Using the results of [4], it is possible to prove the following theorem.

Theorem 1 For each f , when t and r converge to $+\infty$ in such a way that $\frac{r}{t} \rightarrow \alpha$, then, the entries of matrix $\mathbf{Q}(e^{2i\pi f})$ converge in probability towards the entries of the deterministic matrix $(\mathbf{I}_t + \delta(\mathbf{C}) \mathbf{C})^{-1}$, where $\delta(\mathbf{C})$ is the unique strictly positive solution of the equation

$$\frac{\sigma^2}{p} \delta + \frac{1}{t} \text{Tr} \delta \mathbf{C} (\mathbf{I} + \delta \mathbf{C})^{-1} = \frac{r}{t} \quad (5)$$

Using Theorem 1, we get immediately that \mathcal{C}_{mmse} has the same asymptotic behaviour that $\bar{\mathcal{C}}_{mmse}$ defined by

$$\bar{\mathcal{C}}_{mmse} = \sum_{j=1}^t \log_2 \left(\frac{1}{(\mathbf{I} + \delta(\mathbf{C}) \mathbf{C})_{j,j}^{-1}} \right) \quad (6)$$

In other words, the relative error $\frac{|\bar{\mathcal{C}}_{mmse} - \mathcal{C}_{mmse}|}{\mathcal{C}_{mmse}}$ converges towards 0 in the above asymptotic regime.

We now present some simulation experiments which show that the number of antennas r and t for which the approximation $\bar{\mathcal{C}}_{mmse} \simeq \mathcal{C}_{mmse}$ is relevant depends on the number of paths L . In Figure 1, we represent the average of the relative error between $\bar{\mathcal{C}}_{mmse}$ and \mathcal{C}_{mmse} when $L = 1, 4, 7, 10$ for $r = t = 4$, $\frac{p}{\sigma^2} = 10$ dB, and $\mathbf{C} = \mathbf{I}$. The average of the relative error is evaluated on 200 independent realizations of the matrices $(\mathbf{H}(l))_{l=0, \dots, L-1}$. It is clear that if $L = 1$,

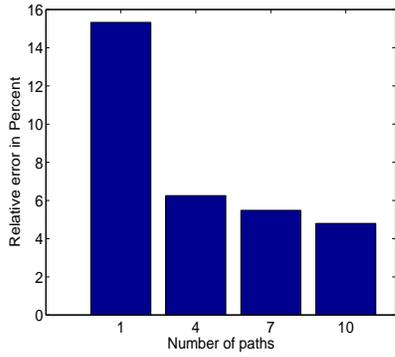


Fig. 1. Average of the relative error.

then the large system approximant is rather far from \mathcal{C}_{mmse} , but that for $L = 4, 7, 10$, the error becomes reasonable.

The above discussion shows that in rich scattering environments, our large system approximation of \mathcal{C}_{mmse} is reliable for realistic numbers of transmit and receive antennas. Therefore, it is relevant to study $\bar{\mathcal{C}}_{mmse}$ in place of \mathcal{C}_{mmse} . In the rest of the paper, we study the effect of \mathbf{C} on $\bar{\mathcal{C}}_{mmse}$, and discuss on the design of linear precoders that optimize $\bar{\mathcal{C}}_{mmse}$.

3. DESIGN OF OPTIMAL PRECODERS.

3.1. Influence of the transmit correlation matrix \mathbf{C} on $\bar{\mathcal{C}}_{mmse}$.

In order to motivate the use of linear precoders, we first establish in this paragraph that diagonal correlation matrices provide higher values of $\bar{\mathcal{C}}_{mmse}$. In the remainder of this paragraph, we mention explicitly that $\bar{\mathcal{C}}_{mmse}$ depends on $\mathbf{C} = \frac{1}{L} \sum_{l=0}^{L-1} \mathbf{C}_l$ and use the notation $\bar{\mathcal{C}}_{mmse}(\mathbf{C})$. We have the following result.

Proposition 1 *Let $\mathbf{C} = \mathbf{U}\mathbf{D}\mathbf{U}^H$ be the eigenvalues/eigenvectors decomposition of matrix \mathbf{C} , where the diagonal entries $(d_j)_{j=1,\dots,t}$ of \mathbf{D} are arranged in decreasing order. Then,*

$$\bar{\mathcal{C}}_{mmse}(\mathbf{C}) \leq \bar{\mathcal{C}}_{mmse}(\mathbf{D}) \quad (7)$$

Proof. We first note that parameter $\delta(\mathbf{C})$ defined by (5) depends on matrix \mathbf{C} , but through its eigenvalues. Therefore, $\delta(\mathbf{C})$ coincides with $\delta(\mathbf{D})$. $\bar{\mathcal{C}}_{mmse}(\mathbf{C})$ can thus be written as

$$\bar{\mathcal{C}}_{mmse}(\mathbf{C}) = \sum_{j=1}^t \log_2 \left(\frac{1}{(\mathbf{I} + \delta(\mathbf{D})\mathbf{U}\mathbf{D}\mathbf{U}^H)_{j,j}^{-1}} \right)$$

$(\mathbf{I} + \delta(\mathbf{D})\mathbf{U}\mathbf{D}\mathbf{U}^H)_{j,j}^{-1}$ is given by

$$(\mathbf{I} + \delta(\mathbf{D})\mathbf{U}\mathbf{D}\mathbf{U}^H)_{j,j}^{-1} = \sum_{k=1}^t \frac{|u_{k,j}|^2}{1 + \delta(\mathbf{D})d_k}$$

where $u_{k,j}$ is the entry (k, j) of unitary matrix \mathbf{U} . The function $y \rightarrow \log_2 \frac{1}{y}$ is convex on \mathbb{R}^+ . As $\sum_{k=1}^t |u_{k,j}|^2 = 1$ (because \mathbf{U} is unitary), we have

$$\log_2 \left[\frac{1}{\sum_{k=1}^t \frac{|u_{k,j}|^2}{1 + \delta(\mathbf{D})d_k}} \right] \leq \sum_{k=1}^t |u_{k,j}|^2 \log_2(1 + \delta(\mathbf{D})d_k)$$

Summing over j , and using that $\sum_{k=1}^t |u_{k,j}|^2 = 1$, we get that

$$\bar{\mathcal{C}}_{mmse}(\mathbf{C}) \leq \sum_{k=1}^t \log_2(1 + \delta(\mathbf{D})d_k) = \bar{\mathcal{C}}_{mmse}(\mathbf{D})$$

This result implies that the maximum of $\bar{\mathcal{C}}_{mmse}(\mathbf{C})$ over the set of all possible correlation matrices for which $\frac{1}{t}\text{Tr}(\mathbf{C}) = 1$ is reached for a diagonal correlation matrix.

3.2. Structure of linear precoders.

Proposition 1 suggests that the use of a linear precoder at the transmitter side may be fruitful. Let \mathbf{K} be a $t \times t$ matrix satisfying $\frac{1}{t}\text{Tr}(\mathbf{K}\mathbf{K}^H) = 1$, and assume that vector $\mathbf{K}\mathbf{s}_n$ is transmitted instead of vector \mathbf{s}_n at each time n . The capacity \mathcal{C}_{mmse} of this precoded system equipped with the MMSE receiver is of course obtained by replacing covariance matrix \mathbf{C} by matrix $\mathbf{K}^H\mathbf{C}\mathbf{K}$. More precisely, \mathcal{C}_{mmse} is given by (4), but in which matrix $\mathbf{Q}(e^{2i\pi f})$ is now defined by

$$\mathbf{Q}(e^{2i\pi f}) = \frac{\sigma^2}{p} \left(\mathbf{K}^H \mathbf{C}^{1/2} \tilde{\mathbf{H}}(e^{2i\pi f})^H \tilde{\mathbf{H}}(e^{2i\pi f}) \mathbf{C}^{1/2} \mathbf{K} + \frac{\sigma^2}{p} \right)_{j,j}^{-1}$$

where $\tilde{\mathbf{H}}(e^{2i\pi f})$ is a zero mean i.i.d. Gaussian random matrix with $\frac{1}{t}$ variance entries. In order to design the precoder \mathbf{K} , it is quite natural to optimize \mathcal{C}_{mmse} over the set of all precoding matrices \mathbf{K} for which $\frac{1}{t}\text{Tr}(\mathbf{K}\mathbf{K}^H) = 1$. As the expression of \mathcal{C}_{mmse} is quite complicated, we propose to study the maximization of the large system approximation $\bar{\mathcal{C}}_{mmse}$ versus \mathbf{K} . We thus denote $\bar{\mathcal{C}}_{mmse}$ by $\bar{\mathcal{C}}_{mmse}(\mathbf{K})$ in the following. $\bar{\mathcal{C}}_{mmse}(\mathbf{K})$ is given by

$$\bar{\mathcal{C}}_{mmse}(\mathbf{K}) = \sum_{j=1}^t \log_2 \left(\frac{1}{(\mathbf{I} + \delta(\mathbf{K}^H\mathbf{C}\mathbf{K})\mathbf{K}^H\mathbf{C}\mathbf{K})_{j,j}^{-1}} \right) \quad (8)$$

where $\delta(\mathbf{K}^H\mathbf{C}\mathbf{K})$ is the unique positive solution of the equation (5) in which matrix \mathbf{C} is replaced by matrix $\mathbf{K}^H\mathbf{C}\mathbf{K}$. The main result of this paper is the following theorem.

Theorem 2 *Let $\mathbf{C} = \mathbf{U}\mathbf{D}\mathbf{U}^H$ be the eigenvalues/eigenvectors decomposition of matrix \mathbf{C} . Let \mathbf{K}_{opt} be the precoding matrix \mathbf{K}_{opt} defined by*

$$\mathbf{K}_{opt} = \mathbf{U}\mathbf{D}^{-1/2}\mathbf{\Gamma}_{opt}^{1/2} \quad (9)$$

where $\mathbf{\Gamma}_{opt} = \text{Diag}(\gamma_{1,opt}, \dots, \gamma_{t,opt})$ is a positive diagonal matrix solution of the optimization problem

Problem 1 Maximize $\sum_{j=1}^t \log_2(1 + \gamma_j \delta(\mathbf{\Gamma}))$ under the constraints

$$\mathbf{\Gamma} = \text{Diag}(\gamma_1, \dots, \gamma_t) \geq 0, \quad \frac{1}{t} \text{Tr}(\mathbf{D}^{-1} \mathbf{\Gamma}) = 1 \quad (10)$$

where $\delta(\mathbf{\Gamma})$ is the unique positive solution of the equation (5) in which \mathbf{C} is replaced by $\mathbf{\Gamma}$.

Proof. In order to prove Theorem 2, we consider a precoding matrix \mathbf{K} such that $\frac{1}{t} \text{Tr}(\mathbf{K} \mathbf{K}^H) = 1$, and the eigenvalues/ eigenvectors decomposition $\mathbf{K}^H \mathbf{C} \mathbf{K} = \mathbf{W} \mathbf{\Gamma} \mathbf{W}^H$ of matrix $\mathbf{K}^H \mathbf{C} \mathbf{K}$. Then, the precoding matrix $\mathbf{K}_d = \mathbf{K} \mathbf{W}$ satisfies $\frac{1}{t} \text{Tr}(\mathbf{K}_d \mathbf{K}_d^H) = 1$ and is such that $\mathbf{K}_d^H \mathbf{C} \mathbf{K}_d$ is diagonal. More importantly, we have

$$\bar{\mathcal{C}}_{mmse}(\mathbf{K}) \leq \bar{\mathcal{C}}_{mmse}(\mathbf{K}_d). \quad (11)$$

(11) can be shown as Proposition 1. This discussion implies that it is sufficient to look for precoding matrices \mathbf{K} for which $\mathbf{K}^H \mathbf{C} \mathbf{K}$ is a positive diagonal matrix $\mathbf{\Gamma}$. This condition is satisfied if and only if \mathbf{K} can be written as $\mathbf{K} = \mathbf{U} \mathbf{D}^{-1/2} \mathbf{\Theta} \mathbf{\Gamma}^{1/2}$ where $\mathbf{\Theta}$ is a unitary matrix. As $\frac{1}{t} \text{Tr}(\mathbf{K} \mathbf{K}^H)$ is supposed to be equal to 1, matrices $\mathbf{\Gamma}$ and $\mathbf{\Theta}$ moreover satisfy $\frac{1}{t} \text{Tr} \mathbf{D}^{-1} \mathbf{\Theta} \mathbf{\Gamma} \mathbf{\Theta}^H = 1$. Each precoder \mathbf{K} defined above can thus be parameterized by the unitary matrix $\mathbf{\Theta}$ and the positive diagonal matrix $\mathbf{\Gamma}$, or equivalently by the hermitian matrix $\mathbf{R} = \mathbf{\Theta} \mathbf{\Gamma} \mathbf{\Theta}^H$ because matrix $\mathbf{\Theta} \mathbf{\Gamma}^{1/2}$ is uniquely defined from \mathbf{R} . $\mathcal{C}_{mmse}(\mathbf{K})$ is equal to $\sum_{j=1}^t \log_2(1 + \delta(\mathbf{\Gamma}) \gamma_j)$, or equivalently to $\log_2 \det(\mathbf{I} + \delta(\mathbf{R}) \mathbf{R})$ because the eigenvalues of \mathbf{R} coincide with the entries of matrix $\mathbf{\Gamma}$, which in particular imply that $\delta(\mathbf{\Gamma}) = \delta(\mathbf{R})$. Hence, the optimization of $\bar{\mathcal{C}}_{mmse}(\mathbf{K})$ is equivalent to the following problem

Problem 2 Maximize $\log_2 \det(\mathbf{I} + \delta(\mathbf{R}) \mathbf{R})$ under the constraints

$$\mathbf{R} \geq 0, \quad \frac{1}{t} \text{Tr} \mathbf{D}^{-1} \mathbf{R} = 1, \quad (12)$$

In order to complete the proof of Theorem 2, it is sufficient to prove that it exists diagonal matrices which are solutions of Problem 2. This is established in the Appendix.

3.3. Study of Problem 1.

Theorem 2 shows that the determination of an optimal precoder \mathbf{K}_{opt} needs to solve the optimization Problem 1. This problem cannot be solved in closed form, except if $\mathbf{D} = \mathbf{I}$. In this context, the results of [1] imply that the solutions of Problem 1 are the vectors $\boldsymbol{\gamma}_{opt} = (\gamma_{1,opt}, \dots, \gamma_{t,opt})$ whose s largest components coincide with $\frac{t}{s}$, and whose $t - s$ smallest ones are 0. The value of $s \leq t$ depends on the signal to noise ratio. If $s < t$, Problem 1 has therefore multiple solutions. This in particular implies that for $\mathbf{D} = \mathbf{I}$, the function $\sum_{j=1}^t \log_2(1 + \gamma_j \delta(\mathbf{\Lambda}))$ is not concave on the convex set defined by the constraints (10).

If $\mathbf{D} \neq \mathbf{I}$, it seems difficult to characterize analytically the solutions of Problem 1. We have thus to use numerical technics. We propose to parameterize γ_j by $\gamma_j = \beta_j^2$ in order to get rid of the constraint $\gamma_j \geq 0$, and to use a standard gradient algorithm with projection on the constraint $\frac{1}{t} \sum_{j=1}^t \frac{\beta_j^2}{d_j} = 1$ at each iteration. Note that the convergence of this algorithm towards a global maximum of $\sum_{j=1}^t \log_2(1 + \gamma_j \delta(\mathbf{\Lambda}))$ is not guaranteed because, as explained above, this last function is in general not concave. However, we have checked if $\mathbf{D} = \mathbf{I}$, then the gradient algorithm always converges towards one of the above mentioned global maximum.

3.4. Numerical illustration of the precoder optimization.

In order to illustrate the impact of the precoder optimization on $\bar{\mathcal{C}}_{mmse}$ and \mathcal{C}_{mmse} in a realistic context, we follow the propagation model introduced in [2] in which each path corresponds to a scatterer cluster characterized by a mean angle of departure and an angle spread. We refer the reader to [2] for more details on the corresponding correlation matrix.

In the following numerical experiment, $r = t = 4$, the number of paths is $L = 5$, they share the same power, and their mean departure angles and angle spreads, expressed in Radians, are equal to 0.72, 1.3, 0.93, 2.1, 2.9 and 0.1, 0.17, 0.03, 0.05, 0.08 respectively. In Figure 2, we have represented $\bar{\mathcal{C}}_{mmse}(\mathbf{I})$ (without precoding) as well as $\bar{\mathcal{C}}_{mmse}(\mathbf{K}_{opt})$ where \mathbf{K}_{opt} is the optimal precoder. We have also represented the true ergodic capacities $\mathcal{C}_{mmse}(\mathbf{I})$ and $\mathcal{C}_{mmse}(\mathbf{K}_{opt})$ evaluated by Monte Carlo simulations. Figure 2 confirms that \mathcal{C}_{mmse} and $\bar{\mathcal{C}}_{mmse}$ are reasonably close one from each other in both cases. In particular, optimizing $\bar{\mathcal{C}}_{mmse}$ over the precoder also allows to improve a lot the true capacity \mathcal{C}_{mmse} .

4. CONCLUDING REMARKS.

We summarize the advantages of our asymptotic analysis of \mathcal{C}_{mmse} . It first allows to prove the relevance of precoders $\mathbf{K} = \mathbf{U} \mathbf{D}^{-1/2} \mathbf{\Gamma}^{1/2}$, where $\mathbf{\Gamma}$ is a positive diagonal matrix. Second, the entries of the optimum matrix $\mathbf{\Gamma}$ are solution of an optimization problem that can be solved by a computationally attractive gradient algorithm. If in contrast, matrix $\mathbf{\Gamma}$ was designed to maximize the true ergodic capacity $\mathcal{C}_{mmse}(\mathbf{U} \mathbf{D}^{-1/2} \mathbf{\Gamma})$, the corresponding gradient algorithm would have a high computational cost. This is because this function of $\mathbf{\Gamma}$, as well as its derivatives w.r.t. the entries of $\mathbf{\Gamma}$, cannot be expressed in closed form. They have to be evaluated by Monte Carlo simulations, thus complicating a lot the maximization algorithm.

APPENDIX

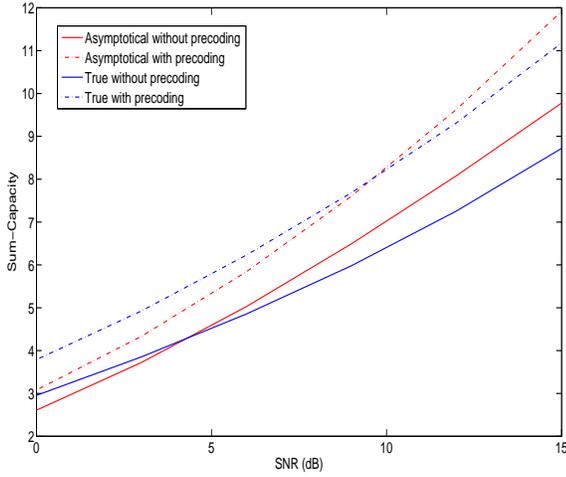


Fig. 2. Capacity with and without optimum precoding.

We prove that there exists diagonal matrices which are solutions of Problem 2. For this, we introduce a simplification of Problem 2 based on the observation that there exists a one-to-one correspondence between the set of matrices \mathbf{R} satisfying the constraints (12) and the set of positive matrices \mathbf{S} verifying

$$\mathbf{S} \geq 0, \quad \frac{\sigma^2}{p} \frac{1}{t} \text{Tr} \mathbf{D}^{-1} \mathbf{S} + \frac{1}{t} \text{Tr} \mathbf{S} (\mathbf{I} + \mathbf{S})^{-1} = \frac{r}{t} \quad (13)$$

In effect, if \mathbf{R} satisfies (12), then $\mathbf{S} = \delta(\mathbf{R})\mathbf{R}$ verifies $\frac{1}{t} \text{Tr} \mathbf{D}^{-1} \mathbf{S} = \delta(\mathbf{R}) \frac{1}{t} \text{Tr} \mathbf{D}^{-1} \mathbf{R} = \delta(\mathbf{R})$. The equation (13) follows immediately from the definition (5) of $\delta(\mathbf{R})$. Conversely, if \mathbf{S} is a positive matrix for which (13) holds, then $\mathbf{R} = \frac{1}{\frac{1}{t} \text{Tr} \mathbf{D}^{-1} \mathbf{S}} \mathbf{S}$ satisfies (12) and $\delta(\mathbf{R}) = \frac{1}{t} \text{Tr} \mathbf{D}^{-1} \mathbf{S}$. Problem 2 is eventually equivalent to

Problem 3 Maximize $\log \det (\mathbf{I} + \mathbf{S})$ under the constraints (13).

and we have to establish that there exists diagonal matrices that are solutions of Problem 3. Let \mathbf{S} be a matrix verifying the constraints (13), and let $\mathbf{S} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^H$ be the eigenvalues/ eigenvectors decomposition of matrix \mathbf{S} , where the eigenvalues $(\lambda_j)_{j=1, \dots, t}$ (the diagonal entries of diagonal matrix $\mathbf{\Lambda}$) are arranged in decreasing order. Then, we claim that it exists $\mu_d \geq 1$ such that $\mu_d \mathbf{\Lambda}$ verifies the constraints (13), and

$$\log_2 \det (\mathbf{I} + \mathbf{S}) = \log_2 \det (\mathbf{I} + \mathbf{\Lambda}) \leq \log_2 \det (\mathbf{I} + \mu_d \mathbf{\Lambda}) \quad (14)$$

To prove this, we first note that $\frac{1}{t} \text{Tr} \mathbf{S} (\mathbf{I} + \mathbf{S})^{-1} = \frac{1}{t} \text{Tr} \mathbf{\Lambda} (\mathbf{I} + \mathbf{\Lambda})^{-1}$. We recall that the eigenvalues $(d_j)_{j=1, \dots, t}$ are arranged in decreasing order. We can therefore use the useful

inequality

$$\frac{1}{t} \text{Tr} \mathbf{D}^{-1} \mathbf{V} \mathbf{\Lambda} \mathbf{V}^H \geq \frac{1}{t} \text{Tr} \mathbf{D}^{-1} \mathbf{\Lambda} \quad (15)$$

which is extensively used in [5] (see the Appendix of [5]). (15) implies that

$$\frac{\sigma^2}{p} \frac{1}{t} \text{Tr} \mathbf{D}^{-1} \mathbf{\Lambda} + \frac{1}{t} \text{Tr} \mathbf{\Lambda} (\mathbf{I} + \mathbf{\Lambda})^{-1} \leq \frac{r}{t}$$

The function $\phi(\mu)$ defined by

$$\phi(\mu) = \frac{\sigma^2}{p} \frac{1}{t} \text{Tr} \mathbf{D}^{-1} \mu \mathbf{\Lambda} + \frac{1}{t} \text{Tr} \mu \mathbf{\Lambda} (\mathbf{I} + \mu \mathbf{\Lambda})^{-1}$$

is an increasing function such that $\phi(1) \leq \frac{r}{t}$. Therefore, it exists a unique $\mu_d \geq 1$ for which $\phi(\mu_d) = 1$. This means that matrix $\mu_d \mathbf{\Lambda}$ satisfies (13). Finally, equality (14) holds because $\mu_d \geq 1$. We denote by \mathbf{S}_d the matrix $\mathbf{S}_d = \mu_d \mathbf{\Lambda}$. (14) implies that for each matrix \mathbf{S} satisfying (13), then the diagonal matrix \mathbf{S}_d verifies (13) and $\log_2 \det (\mathbf{I} + \mathbf{S}) \leq \log_2 \det (\mathbf{I} + \mathbf{S}_d)$. This, in turn, establishes that there exists diagonal matrices which are solutions of Problem 3, and therefore of Problem 2.

5. REFERENCES

- [1] C. Artigue, P. Loubaton, B. Mouhouché, "On the performance of spatial-multiplexing MIMO-CDMA downlink receivers based on chip rate equalization and despreading", In Proc. ICASSP 2008, Las Vegas, April 2008.
- [2] H. Bolcskei, D. Gesbert, A. Paulraj, "On the capacity of OFDM-based spatial multiplexing systems", IEEE Trans. on Communications, vol. 50, no. 2, pp. 225-234, February 2002.
- [3] A. Goldsmith, S.A. Jafar, N. Jindal, S. Vishwanath, "Capacity Limits of MIMO Channels", IEEE J. Sel. Areas in Comm., vol. 21, no. 5, June 2003
- [4] W. Hachem, O. Khorunzhiy, P. Loubaton, J. Najim, L. Pastur, "A new approach for capacity analysis of large dimensional multi-antenna channels, Accepted to IEEE Trans. on Information Theory, preprint arXiv:cs/0612076.
- [5] S. A. Jafar and A. Goldsmith, "Multiple-Antenna capacity in correlated Rayleigh fading with channel covariance information", IEEE Transactions on Wireless Communications, Vol. 4, No. 3, pp. 990-997, May 2005.
- [6] A.L. Moustakas, S.H. Simon, A.M. Sengupta, "MIMO Capacity Through Correlated Channels in the Presence of Correlated Interference and Noise : A (Not so) Large N Analysis", Trans. on Inf. Theo., vol. 49, no 10, pp 2545-2561, Oct. 03.
- [7] A. Scaglione, P. Stoica, S. Barbarossa, G. Giannakis, H. Sampath, "Optimal design for space-time linear precoders and decoders, IEEE Trans. on Signal Processing, vol. 50, no. 5, pp. 1051-1064, May 2002.
- [8] E. Telatar, "Capacity of Multi-antenna Gaussian Channels", Europ. Trans. Telecom., vol. 10, pp 585-595, Nov. 99