Distributed power allocation in fast fading MIMO multiple access channels

S. Lasaulce, A. Suarez, M. Debbah and E. Altman

samson.lasaulce@lss.supelec.fr, suarezr@eurecom.fr, merouane.debbah@supelec.fr, eitan.altman@sophia.inria.fr

August 15, 2007

Abstract

In this paper we design distributed power allocation algorithms for multiple access channels. Considering competing transmitting users, equipped with several antennas each and common multiple antennas at the receiver, a game theoretic framework is conducted to derive the optimum precoding matrices such that each user maximizes his own ergodic transmission rate from the sole knowledge of the overall channel statistics. Both in the scenario with no coordination and single-user decoding at the base station and that with coordination and successive decoding there exists a unique equilibrium point. The corresponding power allocation policies are obtained by exploiting random matrix theory. Numerical results validate the large system assumption in relatively small multiple input multiple output systems and show the benefits of sharing optimally the power between the transmit antennas versus the uniform scheme and also the gain in terms of sum-rate brought by an equiprobable coordination signal. Several extensions of this work are proposed (new elements on these extensions will be provided in the final version).

Index Terms

Distributed, decentralized, power allocation game, Nash equilibrium, coordination, MAC, MIMO, antenna correlation large systems.

I. Introduction

We consider the uplink of a typical cellular system, which is a multiple access channel (MAC). The MAC consists of several users, called mobile stations (MS) in cellular systems, sending independent messages[1] to a common receiver, called the base station (BS). More specifically we assume that both the base and mobile stations are equipped with –possibly correlated– multiple antennas and the different links into play are fading channels. In this context we want to investigate the optimum power allocation (PA) issue at the mobile stations when the signaling protocol overhead is absent or very reduced, which means that the base station sends no or almost no control signals to the mobile stations. This problem is commonly referred to as the distributed or decentralized power allocation problem.

From an information theoretic point of view, the optimal centralized[2] power and rate policies for the fast fading single input single output (SISO) MAC have been determined by [2][3] when channel state information at the receiver is assumed (CSIR) and by [4] when CSI is assumed at the receiver and transmitters (CSIR and CSIT), which leads to the MAC ergodic capacity region. Recently, the authors of [5][6] addressed the fast fading multiple input multiple output (MIMO) MAC with transmit antenna correlation and covariance feedback at the transmitters and determined the optimum power allocation policy in terms of ergodic sum capacity. We consider the same framework as [5][6], which considered fast fading MIMO MAC with CSIR and CDIT (channel distribution information at the transmitters), but we also assume correlation at the receiver and much more importantly we do not assume the power allocation policies to be centralized. In our context each user wants selfishly to maximize its own utility instead of a global utility function such as the sum-capacity[3]. A convenient tool to address decentralized problems turns out to be game theory (see e.g. [8][9]). In this respect the authors of [7] used a game theoretic approach to characterize the ergodic information rates of fast fading SISO and single input multiple output (SIMO) multiple access channels when perfect CSIR is assumed and each user knows his

[1]Note that the most general version of the MAC [1] can also incorporate a common message but this version is not considered here.

[2]The base station dictates the power level and rate of the individual users.

[3]In some situations the selfish behavior of the users amounts to maximizing the overall system utility e.g. the sum-rate [7].

channel and those of the other users. Although reference [7] is probably the closest work to ours we also note that other authors have worked on multiple access or interference channels from a game theoretic perspective. For example, in [10] the authors have chosen the individual mutual information (MI) as a utility function and assumed CSIR and CSIT for studying static MIMO interference channels. In [11] the authors have also considered the individual MI for studying static frequency-selective interference channels. Also some authors used different utility functions, such as those maximizing energy-efficiency (see e.g. [12][13]), in order to study the existence and uniqueness of a Nash equilibrium (NE) in MACs.

Our work can be considered as a partial extension of [7] in the sense that we address MIMO channels instead of SISO and SIMO channels but it differs from it at least in four important points. First, each user is only informed with the statistics of the different channels and not with their instantaneous knowledge. The CDIT assumption is generally considered to be more realistic in fast fading environments and in the particular case of decentralized systems it involves much less feedback signals from the base station, compared to the CSIT assumption. Second, the transmit and receive antennas can be correlated (this feature cannot be considered when assuming perfect CSI since each transmitter exploits the realization of the channel itself). Third, we exploit the theory of random matrices. Considering (moderately) large systems in terms of numbers of antennas has at least two advantages: the underlying averaging effect makes predictable certain quantities of interest, which allows each player to partially/totally predict the strategy of others, and more importantly it simplifies the derivation of distributed power allocation algorithms and the analysis of their properties. Concerning this point, random matrix theory (RMT) will be used with the same approach as the authors of [14], who studied the impact of antenna correlation on fading MIMO single-user channels. As a fourth point, we focus more on the sum-rate as a system performance criterion and present coordination schemes, complementary to (and sometimes simpler to implement than) those developed in the Stackelberg formulation of [7].

This paper is structured as follows. We first provide the signal model used to represent the fading MIMO MAC channel (Sec. II). In Sec. III we introduce four scenarios characterized by the absence/presence of coordination and the decoding scheme at the BS. Two of these scenarios are detailed in this paper: no coordination plus single-user decoding and open loop coordination plus successive decoding. The determination of the corresponding optimum input covariance matrices is conducted in Sec. IV. Numerical results are provided in Sec. V. We conclude by a very brief summary and possible extensions of this work.

II. System model

Notations: in this paper, the notations s, $\underline{v}$, $\mathbf{M}$ stand for scalar, vector and matrix respectively. The superscripts $(.)^T$ and $(.)^H$ denote transpose and transpose conjugate, respectively. The trace of the matrix $\mathbf{M}$ is denoted by $\mathrm{Tr}(\mathbf{M})$. The mathematical expectation operator is denoted by $\mathbb{E}(.)$. and $\mathcal{N}(\underline{v}, \mathbf{M})$ denotes the complex Gaussian distribution with mean $\underline{v}$ and covariance $\mathbf{M}$.

We consider the uplink of a single cell with K active users. Each mobile station is equipped with n_t antennas whereas the base station has n_r antennas (thus we assume the same number of transmitting antennas for all the users). In our analysis the flat fading channel matrices of the different links vary from symbol vector (or space-time codeword) to symbol vector. We assume that the receiver knows all the channel matrices whereas each transmitter has only access to the statistics of the different channels. The equivalent baseband signal received by the base station can be written as

$$\underline{y}(\tau) = \sum_{k=1}^{K} \boldsymbol{H}_k(\tau)\underline{x}_k(\tau) + \underline{n}(\tau) \tag{1}$$

where $\underline{x}_k(\tau)$ is the n_t-dimensional column vector of symbols transmitted by user k at time τ, $\mathbf{H}_k(\tau) \in \mathbb{C}^{n_r \times n_t}$ is the channel matrix (stationary and ergodic process) of user k and $\underline{n}(\tau)$ is a n_r-dimensional complex white Gaussian noise distributed as $\mathcal{N}(\underline{0}, \sigma^2 \mathbf{I}_r)$. For sake of clarity we will omit the time index τ from our notations. Each channel input is subject to a power constraint $\mathrm{Tr}\left[\mathbb{E}(\underline{x}_k \underline{x}_k^H)\right] \triangleq \mathrm{Tr}(\mathbf{Q}_k) \leq n_t \overline{P}_k$. In order to take into account the antenna correlation effects at the transmitters and receiver we will assume the different channel matrices to be structured according to the Kronecker propagation model [15]:

$$\forall k \in \{1, ..., K\}, \;\; \mathbf{H}_k = \mathbf{R}^{\frac{1}{2}} \mathbf{\Theta}_k \mathbf{T}_k^{\frac{1}{2}} \tag{2}$$

where $\mathbf{R}$ is the receive antenna correlation matrix, $\mathbf{T}_k$ is the transmit antenna correlation matrix for user k and $\mathbf{\Theta}_k$ is an $n_r \times n_t$ matrix whose entries are zero-mean independent and identically distributed complex Gaussian random variables with variance $\frac{1}{n_t}$. At last, note that for simplicity we will always assume $K = 2$ but all the results presented in this paper extend to $K-$user MACs, $K \geq 3$. In this respect, in some places K will be used instead of $K = 2$ and some numerical results will be provided for K arbitrary.

III. Different scenarios under investigation

In this paper four scenarios are considered. Although four scenarios are introduced in order for the reader to have a more complete picture of the possible games, we will only focus on two of them (Sec. III-A1 and Sec. III-B) in the optimization part IV. The system can be coordinated or not. In the two scenarios without coordination, once the users have acquired the knowledge of the statistics of the different channels, they do not receive any other kind of information. On the other hand, the users can be coordinated by a signal that informs them with the base station decoding rule. When no coordination is assumed the BS can either apply single-user or successive decoding. For the case of coordination, successive decoding is always assumed while it can be either an open or a closed loop scheme. In all these scenarios the users want to maximize they own ergodic transmission rates, which naturally defines the *utility functions* of the game. The *strategy* of a user consists in choosing his transmit covariance matrix/matrices.

A. No coordination

1) Single-user decoding: We assume that the BS uses single-user (SU) decoding (e.g. because the BS is neutral in the game or for limiting the receiver complexity). Each user treats the signal of the others as additive (colored) noise and wants to selfishly maximize its own transmission rate. The information rate achieved by user k equals the mutual information between $\underline{x}_k$ and $\underline{y}$ conditioned on the overall channel matrix $\mathbf{H} = [\mathbf{H}_1 \mathbf{H}_2 ... \mathbf{H}_K]$. As conditioning the mutual information by a random variable involves taking expectation over this random variable we have:

$$I(\underline{x}_k; \underline{y}|\mathbf{H}) \;=\; \mathbb{E}\left[\log_2\left|\sum_{\ell=1}^{K} \mathbf{H}_\ell \mathbf{Q}_\ell \mathbf{H}_\ell^H + \sigma^2 \mathbf{I}\right|\right] - \mathbb{E}\left[\log_2\left|\sum_{\ell \neq k} \mathbf{H}_\ell \mathbf{Q}_\ell \mathbf{H}_\ell^H + \sigma^2 \mathbf{I}\right|\right] \tag{3}$$

where $|\mathbf{A}|$ stands for the determinant of the matrix $\mathbf{A}$. We see that the second term of the EMI does not depend on $\mathbf{Q}_k$ and we therefore propose not to consider it[4] for the individual utility function of user $k \in \{1, ..., K\}$, which is chosen to be

$$u_k^{(SU)}(\mathbf{Q}_k, \mathbf{Q}_{-k}) = \mathbb{E}\left[\log_2\left|\mathbf{I} + \rho\sum_{\ell=1}^{K}\mathbf{H}_\ell\mathbf{Q}_\ell\mathbf{H}_\ell^H\right|\right], \tag{4}$$

where $\mathbf{Q}_k = (\mathbf{Q}_1, ..., \mathbf{Q}_{k-1}, \mathbf{Q}_{k+1}, ..., \mathbf{Q}_K)$ and $\rho = \frac{1}{\sigma^2}$. Clearly, the users have the same utility function but each user has to maximize it with respect to his *own* transmit covariance matrix. We see that, with the proposed choice of utility functions in the scenario where no coordination is possible and single-user decoding is assumed at the BS, the three concepts of the non-cooperative game, team problem and global optimization problem coincide. We effectively want to optimize the ergodic sum-rate of the MIMO MAC:

$$C_{sum} = \max_{\mathbf{Q}_1,...,\mathbf{Q}_K} \mathbb{E}\left[\log_2\frac{\left|\sum_{\ell=1}^{K}\mathbf{H}_\ell\mathbf{Q}_\ell\mathbf{H}_\ell^H + \sigma^2\mathbf{I}\right|}{|\sigma^2\mathbf{I}|}\right] \tag{5}$$

under the classical trace constraints. What characterizes our problem is that we only want to optimize the sum-rate over $\mathbf{Q}_k$ instead of $(\mathbf{Q}_1, ..., \mathbf{Q}_K)$. In the particular scenario under consideration we have strict concavity of the ergodic sum-rate w.r.t. $(\mathbf{Q}_1, ..., \mathbf{Q}_K)$, which is shown in Appendix A. We further note that the subset of non-negative Hermitian matrices verifying the trace constraints is convex. Therefore there exists a global maximum for the sum-rate. Now, since the players maximize the same function, we can draw the two following conclusions: (a) the global optimum is clearly a NE. This establishes the existence of a NE; (b) the strict concavity of the maximum sum-rate is equivalent to the diagonally strict concavity condition of [16], which implies that the NE is unique (see Theorem 2 of [16]). The main technical issue is to determine the user strategies at the equilibrium point i.e. $(\mathbf{Q}_1^*, ..., \mathbf{Q}_K^*)$, which is done in Sec. IV.

2) Successive decoding: From now on, we will assume a more sophisticated decoding scheme at the BS which is successive decoding or interference cancellation (SIC). In order to be completely fair the BS does not fix the decoding order. Instead, the BS chooses it randomly by flipping a fair coin at each channel realization and does not disclose this information to the users. If one assumes that users 1 and 2

[4]In general, this choice is suboptimal but considerably simplifies the PA analysis.

apply fixed strategies, $\mathbf{Q}_1$ and $\mathbf{Q}_2$, the transmission rates, which are chosen as utility functions, are

$$\left\{\begin{array}{lcl} u_1^{(SIC)}(\mathbf{Q}_1,\mathbf{Q}_2) & = & \frac{1}{2}\mathbb{E}\left[\log_2\left|\mathbf{I}+\rho\mathbf{H}_1\mathbf{Q}_1\mathbf{H}_1^H\right|\right] \\ & & +\frac{1}{2}\mathbb{E}\left[\log_2\left|\mathbf{I}+\rho\mathbf{H}_1\mathbf{Q}_1\mathbf{H}_1^H+\rho\mathbf{H}_2\mathbf{Q}_2\mathbf{H}_2^H\right|\right]-\frac{1}{2}\mathbb{E}\left[\log_2\left|\mathbf{I}+\rho\mathbf{H}_2\mathbf{Q}_2\mathbf{H}_2^H\right|\right] \\ u_2^{(SIC)}(\mathbf{Q}_1,\mathbf{Q}_2) & = & \frac{1}{2}\mathbb{E}\left[\log_2\left|\mathbf{I}+\rho\mathbf{H}_2\mathbf{Q}_2\mathbf{H}_2^H\right|\right] \\ & & +\frac{1}{2}\mathbb{E}\left[\log_2\left|\mathbf{I}+\rho\mathbf{H}_1\mathbf{Q}_1\mathbf{H}_1^H+\rho\mathbf{H}_2\mathbf{Q}_2\mathbf{H}_2^H\right|\right]-\frac{1}{2}\mathbb{E}\left[\log_2\left|\mathbf{I}+\rho\mathbf{H}_1\mathbf{Q}_1\mathbf{H}_1^H\right|\right]. \end{array}\right. \tag{6}$$

For $k \in \{1,2\}$ the function $u_k^{(SIC)}$ is concave w.r.t to $\mathbf{Q}_k$, continuous in $\mathbf{Q}_{k'}$ for $k' \neq k$ and the possible strategies $(\mathbf{Q}_1,\mathbf{Q}_2)$ are taken in a convex set. The existence of a NE in the scenario described then follows from Theorem 1 of [16]. Although we will use this scenario as a second reference in the simulation part, we will not tackle the uniqueness issue of this equilibrium, in order to allocate more space to the case of open loop coordination for which the MSs receive a coordination signal to not only improve their individual rates but also the system throughput.

B. Open loop coordination and successive decoding

In the previous scenario we were in the framework of a game with incomplete information since the players did not know the BS decoding order. For the scenario considered now we assume the existence of a coordination signal denoted by S (see [17] where the authors apply a related idea for ALOHA protocol-based MACs to obtain a correlated equilibrium [18]). It could be obtained in practice, for example, by sampling a broadcast signal (e.g. an FM signal). The realizations of this signal, which are assumed to be equiprobable, are in the finite alphabet $\mathcal{S} = \{1,...,K!\}$. For the case $K=2$ it is therefore simply binary $S \in \{1,2\}$. The difference between the scenario of the previous section and the open loop (OL) coordination scenario is that this signal is known both by the BS and the MSs and therefore provides the missing information of the previous game. Here we assume that the decoding order does not depend on the realizations of $\mathbf{H}$, which are known to the BS but not to the MSs. Thus the coordination signal sent to the users does not provide them with any additional information on the channel conditions. For this reason we call this scheme "open loop coordination". In this framework, we allow the users to apply two different strategies: $\mathbf{Q}_1^{(1)}, \mathbf{Q}_1^{(2)}$ for user 1 and $\mathbf{Q}_2^{(1)}, \mathbf{Q}_2^{(2)}$ for user 2 where the notations $(.)^{(1)}$ and $(.)^{(2)}$ correspond to the realizations of the coordination signal. When $S = 1$, user 1 is privileged since it is decoded after user 2, and conversely for $S = 2$. Thus the achieved transmission rates are given by

$$\left\{ \begin{array}{rcl} R_1^{(1)}(\mathbf{Q}_1^{(1)}, \mathbf{Q}_2^{(1)}) & = & \frac{1}{2}\mathbb{E}\left[\log_2\left|\mathbf{I} + \rho\mathbf{H}_1\mathbf{Q}_1^{(1)}\mathbf{H}_1^H\right|\right] \\ R_2^{(1)}(\mathbf{Q}_2^{(1)}, \mathbf{Q}_1^{(1)}) & = & \frac{1}{2}\mathbb{E}\left[\log_2\left|\mathbf{I} + \rho\mathbf{H}_1\mathbf{Q}_1^{(1)}\mathbf{H}_1^H + \rho\mathbf{H}_2\mathbf{Q}_2^{(1)}\mathbf{H}_2^H\right|\right] - \frac{1}{2}\mathbb{E}\left[\log_2\left|\mathbf{I} + \rho\mathbf{H}_1\mathbf{Q}_1^{(1)}\mathbf{H}_1^H\right|\right] \end{array} \right. \quad (7)$$

when $S = 1$ and by

$$\left\{ \begin{array}{rcl} R_2^{(2)}(\mathbf{Q}_2^{(2)}, \mathbf{Q}_1^{(2)}) & = & \frac{1}{2}\mathbb{E}\left[\log_2\left|\mathbf{I} + \rho\mathbf{H}_2\mathbf{Q}_2^{(2)}\mathbf{H}_2^H\right|\right] \\ R_1^{(2)}(\mathbf{Q}_1^{(2)}, \mathbf{Q}_2^{(2)}) & = & \frac{1}{2}\mathbb{E}\left[\log_2\left|\mathbf{I} + \rho\mathbf{H}_1\mathbf{Q}_1^{(2)}\mathbf{H}_1^H + \rho\mathbf{H}_2\mathbf{Q}_2^{(2)}\mathbf{H}_2^H\right|\right] - \frac{1}{2}\mathbb{E}\left[\log_2\left|\mathbf{I} + \rho\mathbf{H}_2\mathbf{Q}_2^{(2)}\mathbf{H}_2^H\right|\right], \end{array} \right. \quad (8)$$

when $S = 2$. Therefore, when $S = 1$, user 1 sees a single-user MIMO system. The optimum input covariance matrix is obtained by choosing the eigenvectors of $\mathbf{Q}_1^{(1)}$ to be the eigenvectors of $\mathbf{T}_1$ and water-filling over its eigenvalues [19][20]. User 1 has no interest in deviating from this strategy. User 2 knows it and its best strategy is to maximize the sum-rate w.r.t. $\mathbf{Q}_2$ given that $\mathbf{Q}_1 = \mathbf{Q}_1^{(1)}$. For this purpose he will choose its eigenvectors to be equal to those of $\mathbf{T}_2$ and water-fill over its eigenvalues. The same reasoning applies to the case $S = 2$. Thus, this clearly establishes the existence of a unique equilibrium. The users are thus following the coordination signal to adapt their strategies and have no interest in ignoring it. The described strategies can be checked to maximize the following utility functions:

$$\left\{ \begin{array}{rcl} v_1^{(OL)}(\mathbf{Q}_1^{(1)}, \mathbf{Q}_1^{(2)}, \mathbf{Q}_2^{(1)}, \mathbf{Q}_2^{(2)}) & = & \frac{1}{2}R_1^{(1)}(\mathbf{Q}_1^{(1)}, \mathbf{Q}_2^{(1)}) + \frac{1}{2}R_1^{(2)}(\mathbf{Q}_1^{(2)}, \mathbf{Q}_2^{(2)}) \\ v_2^{(OL)}(\mathbf{Q}_1^{(1)}, \mathbf{Q}_1^{(2)}, \mathbf{Q}_2^{(1)}, \mathbf{Q}_2^{(2)}) & = & \frac{1}{2}R_2^{(1)}(\mathbf{Q}_2^{(1)}, \mathbf{Q}_1^{(1)}) + \frac{1}{2}R_2^{(2)}(\mathbf{Q}_2^{(2)}, \mathbf{Q}_1^{(2)}). \end{array} \right. \quad (9)$$

Here we casted the considered scenario into an open loop coordination-based equilibrium, but the game can also be seen as a hierarchical decision making problem. The present problem is a hierarchical decision making problem since, for a given realization of S, the last decoded user can be seen as a leader and the other as a follower. In our case however, the leader does not care about the follower since the actions of the follower have no impact on the leader. We note that the concept of the leader and followers are also present in the Stackelberg formulation. Indeed, by introducing a utility function for the BS, the

Stackelberg formulation of [7] could be applied here but in this case the BS should be involved and send a certain amount of control signal, which is not always negligible, especially when K increases. If the BS can use the information on the channels in order to choose the decoding order, then the signal S sent to the MSs provides them with some information on the channel conditions. This allows the users to have some information on the channel conditions and therefore we can refer to this scheme as closed-loop coordination. We can then replace $\mathbf{H}_k$ by $\mathbf{H}_k^s$ which has the interpretation of the channel condition of user k given that it receives the signal s. If the decision on the decoding order is such that the statistical assumptions on $\mathbf{H}_k^s$ are those we had on $\mathbf{H}_k$ (for example eq. (1) still holds with a possible dependence of the parameters with s) then we can still use eq. (7) and (8) for the utilities except that $\mathbf{H}$ will now also depend on the coordination signal. The equilibrium policies thus derived in the open loop case extend easily to the closed-loop situation.

IV. Determining the optimum covariance matrices in the two scenarios of interest

As in [19][20][6] we distinguish two steps in the determination of the optimum covariance matrices: the optimum eigenvectors are determined in Sec. IV-A by exploiting [6][20] while the optimum eigenvalues are determined in Sec. IV-B by approximating the utility functions under the large system assumption.

A. *Optimum eigenvectors*

From now on we focus on the case with no coordination and single-user decoding at the BS and the case with coordination and successive decoding. In [6] the authors have determined the optimum structure for the transmit covariances matrices that maximizes the channel sum-rate. The proof of [6] can be reused and extended to the case where $\mathbf{R}$ is arbitrary in order to assert that there is no loss of optimality for $u_k^{(SU)}$ and $v_k^{(OL)}$ by restricting the search for the optimum covariance matrix by imposing the structure $\mathbf{Q}_k = \mathbf{U}_k \mathbf{P}_k \mathbf{U}_k^H$ where $\mathbf{T}_k = \mathbf{U}_k \mathbf{D}_k \mathbf{U}_k^H$ is the spectral decomposition of the transmit correlation matrix defined in (2) and the diagonal matrix $\mathbf{P}_k = \mathrm{Diag}(P_k(1), ..., P_k(n_t))$ represents the powers of user k

allocated to the different eigenvectors. This is what states the following theorem, which is proved in Appendix B.

Theorem 4.1 (Optimum eigenvectors): For all $k \in \{1, 2\}$, let $\mathcal{Q}_k$ be the set of $n_t \times n_t$ Hermitian matrices such that $\mathrm{Tr}(\boldsymbol{Q}_k) \leq n_t \overline{P}_k$ i.e. $\mathcal{Q}_k = \left\{\boldsymbol{Q}_k \in \mathbb{C}^{n_t \times n_t} : \boldsymbol{Q}_k = \boldsymbol{Q}_k^H, \mathrm{Tr}(\boldsymbol{Q}_k) \leq n_t \overline{P}_k\right\}$. Additionally, let $\mathcal{S}_k$ be the subset of $\mathcal{Q}_k$ such that $\boldsymbol{Q}_k = \boldsymbol{U}_k \boldsymbol{P}_k \boldsymbol{U}_k^H$ where U_k represents the eigenvectors of $\mathbf{T}_k$. Then, for any $\mathbf{Q}_{-k} \in \mathcal{Q}_{-k}$:

$$\begin{cases} \displaystyle\max_{\mathbf{Q}_k \in \mathcal{Q}_k} u_k^{(SU)}(\mathbf{Q}_k, \mathbf{Q}_{-k}) & = \displaystyle\max_{\mathbf{Q}_k \in \mathcal{S}_k} u_k^{(SU)}(\mathbf{Q}_k, \mathbf{Q}_{-k}) \\ \displaystyle\max_{(\mathbf{Q}_k^{(1)}, \mathbf{Q}_k^{(2)}) \in \mathcal{Q}_k^2} v_k^{(OL)}(\mathbf{Q}_k^{(1)}, \mathbf{Q}_k^{(2)}, \mathbf{Q}_{-k}^{(1)}, \mathbf{Q}_{-k}^{(2)}) & = \displaystyle\max_{(\mathbf{Q}_k^{(1)}, \mathbf{Q}_k^{(2)}) \in \mathcal{S}_k^2} v_k^{(OL)}(\mathbf{Q}_k^{(1)}, \mathbf{Q}_k^{(2)}, \mathbf{Q}_{-k}^{(1)}, \mathbf{Q}_{-k}^{(2)}). \end{cases} \tag{10}$$

This result is instrumental in our context: the best strategy for each user is always to choose an eigenvector basis which matches his own transmit correlation matrix and therefore does not depend on the channels of the other users. This reduces the PA game to the choice of the transmit powers only.

B. Optimum eigenvalues: a large system approach

1) Asymptotic regime: We have shown that for the two decoding schemes considered and for each user, there is no loss of optimality by choosing the eigenvectors of $\mathbf{Q}_k$ to be equal to those of $\mathbf{T}_k = \mathbf{U}_k \mathbf{D}_k \mathbf{U}_k^H$. As a consequence, one can exploit the asymptotic results of [21][14] derived for fading MIMO single-user channels with transmit and receive antenna correlation. This will lead us to simple approximations of the utility functions, which will make easier the optimization of the eigenvalues of the user transmit covariance matrices.

From now on, we assume the asymptotic regime in terms of the number of antennas, which is defined by: (a) $n_t \to \infty$; (b) $n_r \to \infty$; (c) $\displaystyle\lim_{n_t \to \infty, n_r \to \infty} \frac{n_t}{n_r} = c$ where $0 < c < \infty$. For each user $k \in \{1, ..., K\}$, we also suppose that $d_k(1), ..., d_k(n_t)$, which are the elements of the diagonal matrix $\mathbf{D}_k$ defined in Sec. IV-A, have an empirical distribution that converges to a p.d.f. $f_k(t)$ i.e. $\displaystyle\frac{1}{n_t}\sum_{i=1}^{n_t} \delta(t - d_k(i)) \to f_k(t)$.

2) No coordination and single-user decoding: Under the assumptions made above, the capacity per receive antenna $\frac{C_{sum}}{n_r}$ can be shown to converge almost surely towards a certain quantity, denoted by γ_{sum}, which can be obtained by applying Theorem 3.7 of [21]. It can be verified that:

$$\frac{C_{sum}}{n_r} \rightarrow \frac{1}{n_r}\sum_{\ell=1}^{K}\sum_{i=1}^{n_t}\log_2\left[1+K\rho P_\ell(i)d_\ell(i)\alpha\right]+\frac{1}{n_r}\sum_{j=1}^{n_r}\log_2\left[1+K\rho d^{(R)}(j)\beta\right]-\frac{n_t K^2}{n_r}\rho\alpha\beta\log_2 e \quad (11)$$

where the coefficients $d^{(R)}(j)$ correspond to the spectral decomposition of the receive correlation matrix $\mathbf{R}=\mathbf{U}_R\mathbf{D}_R\mathbf{U}_R^H$ with $\mathbf{D}_R=\mathrm{Diag}(d^{(R)}(1),...,d^{(R)}(n_r))$ and the pair (α,β) is the unique solution [22][23] of the following system of equations:

$$\begin{cases} \alpha = \dfrac{1}{Kn_t}\displaystyle\sum_{j=1}^{n_r}\frac{d^{(R)}(j)}{1+K\rho d^{(R)}(j)\beta} \\ \beta = \dfrac{1}{Kn_t}\displaystyle\sum_{\ell=1}^{K}\sum_{i=1}^{n_t}\frac{P_\ell(i)d_\ell(i)}{1+K\rho P_\ell(i)d_\ell(i)\alpha}. \end{cases} \quad (12)$$

In practice, for finite n_t, n_r the utility function $u_k^{(SU)}$ is therefore approximated by $\tilde{u}_k$ defined as $\tilde{u}_k = n_r \times \lim_{n_t\to\infty, n_r\to\infty}\frac{C_{sum}}{n_r}$. This defines an *approximate* game. For each user k, we want to determine the optimal way, in the sense of his approximated utility function $\tilde{u}_k$, to share its available power between the transmit antennas. To solve this constrained optimization problem we introduce the Lagrange multiplier λ_k and define the function

$$\mathcal{L}_{\lambda_k}(P_k(i)) \triangleq \tilde{u}_k - \lambda_k \times \left(\sum_{j=1}^{n_t}P_k(j) - n_t\overline{P}_k\right) \quad (13)$$

and search for the solution(s) $P_k^*(i)$ such that $\frac{\partial \mathcal{L}_{\lambda_k}}{\partial P_k(i)} = 0$. The solution of the corresponding optimization problem is stated through the following theorem.

Theorem 4.2 (Optimum eigenvalues for single-user decoding): Assume that the pair (α,β) *is the solution of the system of equations (12). Then the spatial power allocation maximizing the constrained approximated utility function (13) is given by the following water-filling solution:*

$$P_k^*(i) = \left[\frac{1}{n_r \ln 2\lambda_k} - \frac{1}{K\rho d_k(i)\alpha}\right]^+ \quad (14)$$

where we used the notation $[x]^+ = \max(x,0)$.

The proof of this theorem is provided in Appendix C. In the water-filling procedure the Lagrangian multiplier λ_k, for user k, is tuned in order to meet the power constraint $\sum_{i=1}^{n_t}P_k^*(i) = n_t\overline{P}_k$. Note that the power allocation for a given user k is based on the knowledge of the statistics of his channel but also others through β. We are now in position to describe the proposed (iterative) power allocation algorithm:

1) Initialize α with a value in the interval $[\alpha_{min}, \alpha_{max}]$ with $\alpha_{min} = \frac{1}{Kn_t}\sum_{j=1}^{n_r} \frac{d^{(R)}(j)}{1+K\rho d^{(R)}(j)}$ and $\alpha_{max} = \frac{1}{Kn_t}\sum_{j=1}^{n_r} d^{(R)}(j)$.
2) Apply water-filling over the $d_k(i)$ by using equation (14) in order to find $P_k(i)$ for all $i \in \{1, ..., n_t\}$ and $k \in \{1, ..., K\}$.
3) By using the powers obtained at the previous step, update the value of α by searching for the solution of the system of equations (12).
4) If α has not converged (fix an arbitrary accuracy level on α) go to step 1. Otherwise, apply for the last time step 2 and stop the iterative procedure.

A similar algorithm has been used by [24][25] in order to derive the capacity of single-user Rician MIMO channels with antenna correlation. Based on the results of [24][25] one is insured that the approximated utility function $\tilde{u}_k$ is a strictly concave function of the transmit power vectors $\{\underline{P}_1, ..., \underline{P}_K\}$, with $\forall k \in \{1, ..., K\}$, $\underline{P}_k = (P_k(1), ..., P_k(n_t))$, and if the iterative power allocation algorithm converges[5], it converges towards the global maximum. As the MSs are assumed to implement the proposed power allocation algorithm, which is based on an approximation, it is essential to have these properties since, provided the algorithm convergence, they guarantee the existence and uniqueness of an equilibrium in the proposed *approximated* game, whereas this result was only guaranteed in the *exact* game described in Sec. IV-A. Therefore, for all $\underline{P}_k$, $\tilde{u}_k(\underline{P}_k^*, \underline{P}_{-k}^*) \geq \tilde{u}_k(\underline{P}_k, \underline{P}_{-k}^*)$ where $\underline{P}_{-k}^*$ stands for the strategies of all the users except for user k.

Now we provide a modified version of the iterative power allocation algorithm described above. In this modified version we exploit the idea of asymptotic water-filling, originally introduced by [26]. This second version, which is described below, has essentially two advantages and one drawback. The asymptotic water-filling used in this version not only allows us to restrict the knowledge of the transmitters to the

[5]The authors of [24][25] observed through simulations in many different scenarios that their algorithm always converged.

p.d.f. $f_k(t), k \in \{1, ..., K\}$ instead of the knowledge of the values of $d_k(1), ..., d_k(n_t)$ but also simplify the convergence analysis of the iterative algorithm, which will not be conducted here. The drawback is that in order for the empirical distribution of the eigenvalues $d_k(1), ..., d_k(n_t)$ to be well approximated by the p.d.f. $f_k(t)$, n_t and n_r need to be relatively high. Indeed, the first version of the power allocation algorithm only relies on the approximation of the EMI, which is accurate for small values of n_t, n_r as it will be seen in the simulations.

Let us describe the algorithm in question. Only for sake of clarity we assume here that $\mathbf{R} = \mathbf{I}$. By assuming a known law f_k for the diagonal terms $d_k(i)$, so that

$$\frac{1}{n_t}\sum_{i=1}^{n_t}\frac{1}{d_k(i)} \rightarrow \int \frac{f_k(t)}{t}, \tag{15}$$

we can see that the water level $\mu_k = n_r \ln 2\lambda_k$ can be expressed analytically and only depends on the distribution of $d_k(i)$ according to the following relation, which is obtained from (14) and the power constraints:

$$\overline{P}_k = \int_0^{+\infty}\left[\frac{1}{\mu_k} - \frac{1}{K\rho t\alpha}\right]^+ f_k(t)dt = \int_{\frac{\mu_k}{K\rho\alpha}}^{+\infty}\left(\frac{1}{\mu_k} - \frac{1}{K\rho t\alpha}\right) f_k(t)dt. \tag{16}$$

Therefore μ_k can be obtained through the following fixed-point equation:

$$\mu_k = \frac{\displaystyle\int_{\frac{\mu_k}{K\rho\alpha}}^{+\infty} f_k(t)dt}{\displaystyle\overline{P}_k + K\rho\alpha\int_{\frac{\mu_k}{K\rho\alpha}}^{+\infty}\frac{f_k(t)}{t}dt}. \tag{17}$$

Now, by eliminating β in the system of equations (12), using the fact that $\frac{d_\ell(i)}{1+K\rho P_\ell(i)d_\ell(i)\alpha} = \frac{\mu_\ell}{K\rho\alpha}$ (to check this see Appendix C) when a non-zero power is allocated to the i^{th} antenna, using the expression of $P_\ell^*(i)$ (eq. (14) and using the p.d.f $f_k(t)$ instead of $d_k(i)$ in the asymptotic regime, we obtain the following equation: $\alpha + \frac{1}{K}\sum_{\ell=1}^{K}\int_{\frac{\mu_\ell}{K\rho\alpha}}^{+\infty}\left(1 - \frac{\mu_\ell}{K\rho\alpha t}\right) f_\ell(t)dt = \frac{1}{Kc}$. The second iterative algorithm follows. First, initialize the value of α as indicated in the first algorithm. Find, for every k, the asymptotic water level μ_k through eq. (17). Use these water levels to update α through the latter equation. Repeat this procedure till α has converged. Finally, apply the water-filling equation (14). We see that the convergence analysis amounts to studying the convergence of α and μ_k over iterations, which seems to be a, perhaps not trivial, but feasible task. To this end a transmit correlation profile has to be chosen and derive the corresponding p.d.f. $f_k(t)$. For instance the authors of [27] have calculated it for an exponential correlation profile: $\forall (i,j) \in \{1,...,n_t\}^2,\ T_k(i,j) = r_k^{|i-j|}$ where r_k is the correlation coefficient characterizing the correlation matrix $\mathbf{T}_k$ (this model is assumed in the simulations in Sec. V). The latter authors showed that $f_k(t) = \dfrac{1}{\pi t\sqrt{-t^2+2a_k t - 1}}$ if $\dfrac{1-r_k}{1+r_k} < t < \dfrac{1+r_k}{1-r_k}$ and 0 otherwise, with $a_k \triangleq \frac{1+r_k^2}{1-r_k^2}$. The convergence analysis could be considered as an extension of this paper.

3) Open loop coordination and successive decoding : In the case where the BS applies successive decoding in the order indicated by the coordination signal the equilibrium and the iterative algorithm analyses can be conducted by using the same reasoning as used previously. In this section we will only provide the expressions of the optimum transmit powers. Assume that $S=1$. Then the achievable transmission rates for the two users are:

$$\left\{\begin{array}{lcl} u_1^{(1)}(\mathbf{Q}_1^{(1)},\mathbf{Q}_2^{(1)}) & = & \frac{1}{2}\underbrace{\mathbb{E}\left[\log_2\left|\mathbf{I}+\rho\mathbf{H}_1\mathbf{Q}_1^{(1)}\mathbf{H}_1^H\right|\right]}_{\tau_1} \\ u_2^{(1)}(\mathbf{Q}_2^{(1)},\mathbf{Q}_1^{(1)}) & = & \frac{1}{2}\underbrace{\mathbb{E}\left[\log_2\left|\mathbf{I}+\rho\mathbf{H}_1\mathbf{Q}_1^{(1)}\mathbf{H}_1^H+\rho\mathbf{H}_2\mathbf{Q}_2^{(1)}\mathbf{H}_2^H\right|\right]}_{\tau_s} - \frac{1}{2}\underbrace{\mathbb{E}\left[\log_2\left|\mathbf{I}+\rho\mathbf{H}_1\mathbf{Q}_1^{(1)}\mathbf{H}_1^H\right|\right]}_{\tau_1}. \end{array}\right. \tag{18}$$

By exploiting the results of Sec. IV-A, Theorem 3.7 of [21] and choosing in this theorem K to be equal to the number of terms of the type $\mathbf{H}_k\mathbf{Q}_k\mathbf{H}_k^H$ present in the argument of the operator $\mathbb{E}[\log|.|]$ to be approximated, it can be checked that

$$\frac{\tau_1}{n_r} \rightarrow \frac{1}{n_r}\sum_{i=1}^{n_t}\log_2\left[1+\rho P_1^{(1)}(i)d_1(i)\alpha_1\right] + \frac{1}{n_r}\sum_{j=1}^{n_r}\log_2\left[1+\rho d^{(R)}(j)\beta_1\right] - \frac{n_t}{n_r}\rho\alpha_1\beta_1\log_2 e \tag{19}$$

where

$$\left\{\begin{array}{lcl} \alpha_1 & = & \dfrac{1}{n_t}\displaystyle\sum_{j=1}^{n_r}\frac{d^{(R)}(j)}{1+\rho d^{(R)}(j)\beta_1} \\ \beta_1 & = & \dfrac{1}{n_t}\displaystyle\sum_{i=1}^{n_t}\frac{P_1^{(1)}(i)d_1(i)}{1+\rho P_1^{(1)}(i)d_1(i)\alpha_1}. \end{array}\right. \tag{20}$$

The proof of Theorem 4.2 can be re-used here. Then, optimizing the approximated rate $\tilde{\tau}_1 = n_r \times \lim\limits_{n_t\rightarrow\infty, n_r\rightarrow\infty}\dfrac{\tau_1}{n_r}$ w.r.t. $P_1^{(1)}(i)$ leads to the following water-filling equation

$$P_1^{(1),*}(i) = \left[\frac{1}{n_r\ln 2\lambda_1} - \frac{1}{\rho d_1(i)\alpha}\right]^+. \tag{21}$$

We also know that user 2 will maximize the term $\tilde{\tau}_s = n_r \times \lim_{n_t\to\infty, n_r\to\infty} \frac{\tau_s}{n_r}$ by choosing his input covariance matrix to be structured as $\mathbf{Q}_2 = \mathbf{U}_2\mathbf{P}_2\mathbf{U}_2^H$ with

$$\frac{\tau_s}{n_r} \to \frac{1}{n_r}\sum_{\ell=1}^{2}\sum_{i=1}^{n_t}\log_2\left[1+2\rho P_\ell^{(1)}(i)d_\ell(i)\alpha_2\right] + \frac{1}{n_r}\sum_{j=1}^{n_r}\log_2\left[1+2\rho d^{(R)}(j)\beta_2\right] - \frac{4n_t}{n_r}\rho\alpha_2\beta_2\log_2 e \quad (22)$$

where

$$\begin{cases} \alpha_2 &= \dfrac{1}{2n_t}\displaystyle\sum_{j=1}^{n_r}\frac{d^{(R)}(j)}{1+2\rho d^{(R)}(j)\beta_2} \\ \beta_2 &= \dfrac{1}{2n_t}\displaystyle\sum_{\ell=1}^{2}\sum_{i=1}^{n_t}\frac{P_\ell^{(1)}(i)d_\ell(i)}{1+2\rho P_\ell^{(1)}(i)d_\ell(i)\alpha_2}. \end{cases} \quad (23)$$

Eventually the powers for user 2 can be determined by

$$P_2^{(1),*}(i) = \left[\frac{1}{n_r\ln 2\lambda_2} - \frac{1}{2\rho d_2(i)\alpha_2}\right]^+. \quad (24)$$

V. Numerical results

First we show that in order to make the large system approximation accurate the numbers of antennas do not need to be very high. This is especially true when the metric of interest is the EMI since one benefits from a double averaging effect, one from the randomness of the matrices into play and the other one from the expectation operator. Fig. 1 shows that the relative error is less than $5\ \%$ even for a 2×2 MIMO system. Fig. 2 compares the most simple decentralized power allocation scheme, which is the uniform scheme, with the optimized power allocation scheme when no coordination and single-user decoding are assumed. Since single-user decoding is used at the BS, the system performance is inteference-limited, which clearly appears in the high SNR regime. We note a significant performance gap between the uniform and optimized schemes. Fig. 3 and 4 represent the sum-rate versus K for different power allocation schemes. Here we assumed $\overline{P}_1 = ... = \overline{P}_K$. We see that coordinating the system with a (free) equiprobable random signal allows us to be quite close to the (centralized) MIMO MAC sum-capacity, which shows the interest in the proposed scheme in typical simulation scenarios.

VI. Conclusions

Our goal was to design PA algorithms in fast fading MIMO channels with correlation while minimizing the amount of control signal from the BS. To this end we only assumed CSIR and CDIT. We saw that for the four scenarios introduced there was systematically an equilibrium, which effectively allows the mobiles to choose their PA policies by themselves in order to selfishly optimize their ergodic transmission rates. Let us briefly summarize. No coordination + SU: our choice of utility functions led us to a team problem or a global optimization problem. This is thus a special case of a NE. A general NE would be obtained by choosing $u_k^{(SU)} = I(\underline{x}_k; \underline{y}|\mathbf{H})$ (extension 1). No coordination + SIC: there exists a NE. Studying this case in details would require the uniqueness analysis (ext. 2). Open loop coordination + SIC: there is a unique equilibrium. By choosing a more sophisticated coordination signal (from the BS) one can extend the open loop approach presented in this paper to a closed loop approach for MIMO channels (ext. 3). For the two scenarios detailed in this paper we obtained the optimum PA by identifying the optimum eigenvector structure and exploiting the large system assumption, which was shown to be valid for 2×2 MIMO systems. This allowed us to design simple iterative PA algorithms. These algorithms always converge in simulations but a convergence result could be obtained (ext. 4) to give more theoretical value to these algorithms. We also think that the large systems analysis, when applicable, is a relevant tool in game theoretic frameworks since the underlying averaging effects make predictable certain quantities, which is in favor of having equilibria since the players can predict the strategies of others more easily. Our results could be extended to Rician channels (ext. 5) by using the recent RMT-based results of [24] for SU channels and those of [28] for centralized MACs, whereas cumbersome and greedy optimization procedures would be needed to conduct the proposed extension.

APPENDIX A
STRICT CONCAVITY OF THE MAC MIMO ERGODIC SUM-RATE

The ergodic sum-rate writes $\mathbb{E}\left[\log_2\left|\mathbf{I} + \rho \sum_{k=1}^{K} \mathbf{H}_k \mathbf{Q}_k \mathbf{H}_k^H\right|\right] = \mathbb{E}\left[\log_2\left|\mathbf{I} + \rho \mathbf{H}\mathbf{Q}\mathbf{H}^H\right|\right] \triangleq f(\mathbf{Q})$ with $\mathbf{H} = [\mathbf{H}_1 \dots \mathbf{H}_K]$ and $\mathbf{Q} = \mathrm{Diag}(\mathbf{Q}_1, ..., \mathbf{Q}_K)$. Proving that the ergodic sum-rate is strictly concave w.r.t. $\mathbf{Q}$ is equivalent to proving that the function $\phi(\lambda) = f(\lambda \mathbf{Q}_1 + (1-\lambda)\mathbf{Q}_2)$ is strictly concave w.r.t. λ over the interval $[0,1]$ for any pair of matrices $(\mathbf{Q}_1, \mathbf{Q}_2)$. It can be checked that the second derivative of ϕ is given by:

$$\begin{aligned}\phi''(\lambda) &= -\mathbb{E}\mathrm{Tr}\left[\left(\mathbf{I}+\rho\mathbf{H}^H\mathbf{H}\mathbf{Q}\right)^{-1}\rho\mathbf{H}^H\mathbf{H}\Delta\mathbf{Q}\left(\mathbf{I}+\rho\mathbf{H}^H\mathbf{H}\mathbf{Q}\right)^{-1}\rho\mathbf{H}^H\mathbf{H}\Delta\mathbf{Q}\right] \\ &= -\mathbb{E}\mathrm{Tr}\big[\Delta\mathbf{Q}^{\frac{1}{2}}\underbrace{\left(\mathbf{I}+\rho\mathbf{H}^H\mathbf{H}\mathbf{Q}\right)^{-1}\rho\mathbf{H}^H\mathbf{H}}_{\mathbf{A}}\Delta\mathbf{Q}^{\frac{1}{2}}\Delta\mathbf{Q}^{\frac{1}{2}}\left(\mathbf{I}+\rho\mathbf{H}^H\mathbf{H}\mathbf{Q}\right)^{-1}\rho\mathbf{H}^H\mathbf{H}\Delta\mathbf{Q}^{\frac{1}{2}}\big]\end{aligned} \tag{25}$$

with $\mathbf{Q} = \lambda\mathbf{Q}_1 + (1-\lambda)\mathbf{Q}_2$ and $\Delta\mathbf{Q} = \mathbf{Q}_1 - \mathbf{Q}_2$. As the matrix $\mathbf{A}$ can be checked to be Hermitian, the matrix $\mathbf{B} = \Delta\mathbf{Q}^{\frac{1}{2}}\mathbf{A}\Delta\mathbf{Q}^{\frac{1}{2}}$ is also Hermitian and therefore $\phi''(\lambda) = -\mathbb{E}\left[\mathrm{Tr}(\mathbf{B}\mathbf{B}^H)\right] < 0$.

APPENDIX B

OPTIMUM EIGENVECTORS FOR DECENTRALIZED MIMO MAC WITH DOUBLE-SIDED CORRELATION

To prove Theorem 4.1 for $u_k^{(SU)}$ we follow exactly the same steps as [6] and use an additional argument due to the fact that the receive antenna can be correlated here. By definition $\mathbf{H}_\ell = \mathbf{R}^{\frac{1}{2}}\mathbf{\Theta}_\ell\mathbf{T}_\ell^{\frac{1}{2}} = \mathbf{U}_R\mathbf{D}_R^{\frac{1}{2}}\mathbf{U}_R^H\mathbf{\Theta}_\ell\mathbf{U}_\ell\mathbf{D}_\ell^{\frac{1}{2}}\mathbf{U}_\ell^H$, where $\mathbf{\Theta}_\ell$ is a zero-mean i.i.d. Gaussian identity covariance random matrix. Using the fact that multiplying $\mathbf{\Theta}_\ell$ by a unitary matrix does not change its joint distribution and the fact that $|\mathbf{U}\mathbf{M}\mathbf{U}^H + \mathbf{I}| = |\mathbf{M}+\mathbf{I}|$ for any unitary matrix $\mathbf{U}$ one can write:

$$\begin{aligned}&\max_{\mathbf{Q}_k} u_k^{(SU)} = \\ &\max_{\mathbf{Q}_k}\mathbb{E}\left[\log_2\left|\mathbf{I}+\rho\sum_{\ell=1}^{K}\mathbf{U}_R\mathbf{D}_R^{\frac{1}{2}}\mathbf{U}_R^H\mathbf{\Theta}_\ell\mathbf{U}_\ell\mathbf{D}_\ell^{\frac{1}{2}}\mathbf{U}_\ell^H\mathbf{Q}_\ell\mathbf{U}_\ell\mathbf{D}_\ell^{\frac{1}{2}}\mathbf{U}_\ell^H\mathbf{\Theta}_\ell^H\mathbf{U}_R\mathbf{D}_R^{\frac{1}{2}}\mathbf{U}_R^H\right|\right] = \\ &\max_{\mathbf{Q}_k}\mathbb{E}\left[\log_2\left|\mathbf{I}+\rho\sum_{\ell=1}^{K}\mathbf{D}_R^{\frac{1}{2}}\mathbf{\Theta}_\ell\mathbf{D}_\ell^{\frac{1}{2}}\mathbf{U}_\ell^H\mathbf{Q}_\ell\mathbf{U}_\ell\mathbf{D}_\ell^{\frac{1}{2}}\mathbf{\Theta}_\ell^H\mathbf{D}_R^{\frac{1}{2}}\right|\right].\end{aligned} \tag{26}$$

Then we can spectrally decompose the matrix $\mathbf{D}_\ell^{\frac{1}{2}}\mathbf{U}_\ell^H\mathbf{Q}_\ell\mathbf{U}_\ell\mathbf{D}_\ell^{\frac{1}{2}} = \tilde{\mathbf{U}}_\ell\tilde{\mathbf{D}}_\ell\tilde{\mathbf{U}}_\ell^H$ and write that

$$\begin{aligned}\max_{\mathbf{Q}_k} u_k^{(SU)} &= \max_{\mathbf{Q}_k}\mathbb{E}\left[\log_2\left|\mathbf{I}+\rho\sum_{\ell=1}^{K}\mathbf{D}_R^{\frac{1}{2}}\mathbf{\Theta}_\ell\tilde{\mathbf{U}}_\ell\tilde{\mathbf{D}}_\ell\tilde{\mathbf{U}}_\ell^H\mathbf{\Theta}_\ell^H\mathbf{D}_R^{\frac{1}{2}}\right|\right] \\ &= \max_{\mathbf{Q}_k}\mathbb{E}\left[\log_2\left|\mathbf{I}+\rho\sum_{\ell=1}^{K}\mathbf{D}_R^{\frac{1}{2}}\mathbf{\Theta}_\ell\tilde{\mathbf{D}}_\ell\mathbf{\Theta}_\ell^H\mathbf{D}_R^{\frac{1}{2}}\right|\right].\end{aligned} \tag{27}$$

We see that the function to be optimized depends on the eigenvectors $\tilde{\mathbf{U}}_k$ only through the power constraint $\mathrm{Tr}(\mathbf{Q}_\mathrm{k}) = \mathrm{Tr}(\tilde{\mathbf{U}}_\mathrm{k}^\mathrm{H}\mathbf{D}_\mathrm{k}^{-1}\tilde{\mathbf{U}}_\mathrm{k}\tilde{\mathbf{D}}_\mathrm{k}) \leq \mathrm{n_t}$. The matrix $\tilde{\mathbf{U}}_k$ can be chosen arbitrarily provided it meets the power constraint $\mathrm{Tr}(\mathbf{Q}_\mathrm{k}) \leq \mathrm{n_t}$. The choice $\tilde{\mathbf{U}}_k = \mathbf{I}$ is feasible since $\mathrm{Tr}(\mathbf{D}_\mathrm{k}^{-1}\tilde{\mathbf{D}}_\mathrm{k}) \leq \mathrm{Tr}(\tilde{\mathbf{U}}_\mathrm{k}^\mathrm{H}\mathbf{D}_\mathrm{k}^{-1}\tilde{\mathbf{U}}_\mathrm{k}\tilde{\mathbf{D}}_\mathrm{k}) \leq \mathrm{n_t}$. This shows that $\mathbf{Q}_k$ can be chosen without loss of optimality to be structured as: $\mathbf{Q}_k = \mathbf{U}_k\mathbf{D}_k^{-1}\tilde{\mathbf{D}}_k\mathbf{U}_k^H$.

For the optimization of $v_k^{(OL)}$ one has to note that for each user $k \in \{1,2\}$, $\mathbf{Q}_k^{(1)}$ and $\mathbf{Q}_k(2)$ are optimized independently. For a given realization s of S, the optimum structure of $\mathbf{Q}_k^{(s)}$ for the interference-free channel follows from [20]. For the other user re-use the derivation for $u_k^{(SU)}$ to conclude the proof.

Appendix C

Proof of Theorem 4.2

In the proof provided here we assumed for clarity $\mathbf{R} = \mathbf{I}$ but the result can be easily proved to hold in the general case. We want to derivate the utility function $\tilde{u}_k$ given by eq. (22). It turns out that the partial derivative with respect to $P_k(i)$ is the same as it would be if α and β would be assumed to be independent of these quantities. This result is useful because it allows us to cope with the convergence issue of the quantities α, β towards strict constants as the numbers of users and dimensions grow. Therefore, the main interest in the proposed derivation is that one does not need to assume α or β to be independent of $P_k(i)$. Otherwise the result can be obtained much more easily. We want to prove that the derivative of the approximated utility of user k can be expressed as: $\dfrac{\partial \gamma_{sum}}{\partial P_k(i)} = \dfrac{1}{n_r \ln 2} \dfrac{K\rho d_k(i)\alpha}{1+K\rho P_k(i) d_k(i)\alpha}$. We have:

$$n_r \gamma_{sum} = \log_2 \left\{ \prod_{\ell,j} \left[1 + K\rho P_\ell(j) d_\ell(j) \alpha(P_k(i))\right] \times (1 + K\rho\beta(P_k(i)))^{n_r} e^{-n_t K^2 \rho \alpha(P_k(i))\beta(P_k(i))} \right\}. \quad (28)$$

Define $u \triangleq \prod_{\ell,j} \left[1 + K\rho P_\ell(j) d_\ell(j) \alpha(P_k(i))\right]$ and $v \triangleq (1 + K\rho\beta(P_k(i)))^{n_r} e^{-n_t K^2 \rho \alpha(P_k(i))\beta(P_k(i))}$. With these notations: $\dfrac{\partial n_r \gamma_{sum}}{\partial P_k(i)} = \dfrac{1}{\ln 2} \dfrac{1}{uv} \dfrac{\partial uv}{\partial P_k(i)}$. It turns out that $\dfrac{\partial uv}{\partial P_k(i)} = uv \times \dfrac{K\rho d_k(i)\alpha}{1+K\rho P_k(i) d_k(i)\alpha}$. This is what we want to show. We want to derivate the function u. As u is a product of functions $u_{\ell,j}$, i.e. $u = \prod_{\ell,j} u_{\ell,j}$, its derivative u' can be written as: $u' = u \times \sum_{\ell,j} \frac{u'_{\ell,j}}{u_{\ell,j}}$. More precisely

$$u' = \underbrace{\prod_{\ell',j'} \left[1 + K\rho P_{\ell'}(j') d_{\ell'}(j')\alpha\right]}_{u} \sum_{\ell,j} \frac{N(\ell,j)}{1 + K\rho P_\ell(j) d_\ell(j)\alpha} \text{ where}$$

$$N(\ell,j) = \left| \begin{array}{ll} K\rho P_\ell(j) d_\ell(j)\alpha' & \text{if } (\ell,j) \neq (k,i) \\ K\rho d_k(i)(\alpha + P_k(i)\alpha') & \text{if } (\ell,j) = (k,i). \end{array} \right.$$

Using a similar reasoning for v one can check that $v' = v \times K\rho \left[\frac{n_r \beta'}{1+K\rho\beta} - K n_t (\alpha'\beta + \alpha\beta')\right]$. Now using the relations proved in the previous steps we have that

$$
\begin{aligned}
\frac{\partial uv}{\partial P_k(i)} &= u'v + uv' \\
&= uv \times \left\{ \sum_{\ell,j} \frac{N(\ell,j)}{1 + K\rho P_\ell(j) d_\ell(j)\alpha} + K\rho \left[\frac{n_r \beta'}{1 + K\rho\beta} - K n_t(\alpha'\beta + \alpha\beta') \right] \right\} \\
&= uv \times \left\{ \sum_{(\ell,j)\neq(k,i)} \frac{K\rho P_\ell(j) d_\ell(j)\alpha'}{1 + K\rho P_\ell(j) d_\ell(j)\alpha} + \frac{K\rho d_k(i)(\alpha + P_k(i)\alpha)'}{1 + K\rho P_k(i) d_k(i)\alpha} + K\rho \left[\frac{n_r \beta'}{1 + K\rho\beta} - K n_t(\alpha'\beta + \alpha\beta') \right] \right\}
\end{aligned}
$$

Now using the definitions of α and β (see eq. (12)) we find, after simplifications, the proposed expression for the derivative of γ_{sum}. Finally, by setting the derivative of $\mathcal{L}_{\lambda_k}(P_k(i))$ to zero we find equation (14).

REFERENCES

[1] T. S. Han, "The capacity of general multiple-access channel with certain correlated source", *in Information Control*, Vol. 40, No. 1, 1979, pp. 37–60.

[2] R. G. Gallager, "An inequality on the capacity region of multiaccess multipath channels", *MIT technical report, Laboratory for Information and Decision Systems*, Series/Report No: LIDS-P; 2240, April 1994.

[3] S. Shamai and A. D. Wyner, "Information-theoretic considerations for symmetric, cellular, multiple-access fading channels", *in IEEE Trans. on Inform. Theory*, Vol. 43, No. 6, Nov. 1997, pp. 1877–1894.

[4] D. N. C. Tse and S. Hanly, "Multi-access fading channels – Part I: polymatroid structure, optimal resource allocation and throughput capacities", *IEEE Trans. on Inform. Theory*, Vol. 44, No. 7, Nov. 1998, pp. 2796–2815.

[5] A. Soysal and S. Ulukus, "Optimum power allocation for single-user MIMO and multi-user MIMO-MAC with partial CSI", *IEEE Jour. on Selected Areas in Communications*, to appear.

[6] A. Soysal and S. Ulukus, "Optimality of beamforming in fading MIMO multiple access channels", *submitted to IEEE Trans. on Communications*, January 2007.

[7] L. Lai and H. El Gamal, "The water-filling game in fading multiple access channels", *sub. to IEEE Trans. on Inform. Theory*, 2005.

[8] D. Fudenberg and J. Tirole, "Game theory", MIT Press, 1991.

[9] E. Altman, T. Boulogne, E. El-Azouzi, T. Jiménez and, L. Wynter, "A survey on networking games in telecommunication", *Computer and operations research - Elsevier Science Ltd*, Vol. 33, No. 2, pp. 286–311, Feb. 2006.

[10] G. Arslan, M.F. Demirkol and Y. Sang, "Equilibrium efficiency improvement in MIMO interference systems: a decentralized stream control approach", *Proceedings of the IEEE International Conference on Communications*, Istanbul, Turkey, June 11-15, 2006.

[11] G. Scutari, D. P. Palomar and S. Barbarossa, "Optimal linear precoding strategies for wideband non-cooperative systems based on game theory Part I: Nash Equilibria", *to appear in IEEE Trans. on Signal Processing*, 2007.

[12] F. Meshkati, H. Poor, S. Schwartz, and N. Mandayam, "An energy-efficient approach to power control and receiver design in wireless data networks", *IEEE Trans. on Comm.*, Vol. 53, No. 11, pp. 1885–1894, Nov. 2005.

[13] F. Meshkati, M. Chiang, H. Poor, and S. Schwartz, “A game-theoretic approach to energy-efficient power control in multi-carrier CDMA systems”, *IEEE Journal on Selected Areas in Communications*, Vol. 24, No. 6, June 2006, pp. 1115–1129.

[14] A. M. Tulino, A. Lozano and S. Verdu, “Impact of antenna correlation on the capacity of multiantenna channels Tulino”, *IEEE Trans. on Inform. Theory*, Vol. 51, No. 7, July 2005, pp. 2491–2509.

[15] D. S. Shiu, G. J. Foschini, M. J. Gans and J. M. Kahn, “Fading correlation and its effects on the capacity of multielement antenna systems”, *IEEE Trans. on Comm.*, Vol. 48, No. 3, Mar. 2000, pp. 502–513.

[16] J. Rosen, “Existence and uniqueness of equilibrium points for concave n-person games”, *Econometrica*, Vol. 33, pp. 520–534, 1965.

[17] E. Altman, N. Bonneau and M. Debbah, “Correlated Equilibrium in Access Control for Wireless Communications”, *Networking 2006*, Coimbra, Portugal , May 2006.

[18] R. Aumann, “Subjectivity and correlation in randomized strategies”, Journal of Mathematical Economics 1”, 1974, pp. 67–96.

[19] S. A. Jafar and A. J. Goldsmith, “Transmitter optimization and optimality of beamforming for multiple antenna systems with imperfect feedback”, *IEEE Transactions on Wireless Communications*, July 2004, Vol. 3, No. 4, pp. 1165–1175.

[20] E. A. Jorswieck and H. Boche, “Channel capacity and capacity-range of beamforming in MIMO wireless systems under correlated fading with covariance feedback”, *IEEE Trans. on Wireless Comm.*, Vol. 3, No. 5, pp. 1543–1553, Sep. 2004.

[21] A. Tulino and S. Verdu, “Random Matrices and Wireless Communications”, *Foundations and trends in communications and information theory, NOW, The Essence of Knowledge*, 2004.

[22] Jack W. Silverstein and Z.D. Bai, “On the Empirical Distribution of Eigenvalues of a Class of Large Dimensional Random Matrices”, *Journal of Multivariate Analysis*, 1995, vol. 54, pp. 175–192.

[23] V. L. Girko, “Theory of Stochastic Canonical Equations”, Kluwer, 2001, Vol. 1 and Vol. 2, Dordrecht, The Netherlands.

[24] J. Dumont, P. Loubaton and S. Lasaulce, “On the capacity achieving transmit covariance matrices of MIMO correlated Rician channels: a large system approach”, *Proc. of Globecom*, 2006.

[25] J. Dumont, W. Hachem, S. Lasaulce, P. Loubaton and J. Najim, “Quelques propriétés d’un approximant de l’information mutuelle des canaux MIMO de Rice bi-corrélés”, *in Proc. of the Signal and Image Processing Conference (GRETSI)*, Troyes, France, Sep. 2007.

[26] C. Chuah, D. Tse, J. Kahn and R. Valenzuela, “Capacity scaling in MIMO wireless systems under correlated fading”, *IEEE Trans. on Inform. Theory*, Vol. 48, No. 3, pp. 637–650. Mar. 2002.

[27] A. Skupch, D. Seethaler and F. Hlawatsch, “Free probability based capacity calculation for MIMO channel with transmit or receive correlation”, *in IEEE Proc. of WirelessComm*, June 2005, Hawai, pp. 1041–1046.

[28] E. Riegler and G. Taricco, “On the ergodic capacity region of the separately correlated Rician fading multiple access MIMO channel”, *Proc. of Globecom*, to appear in 2007.

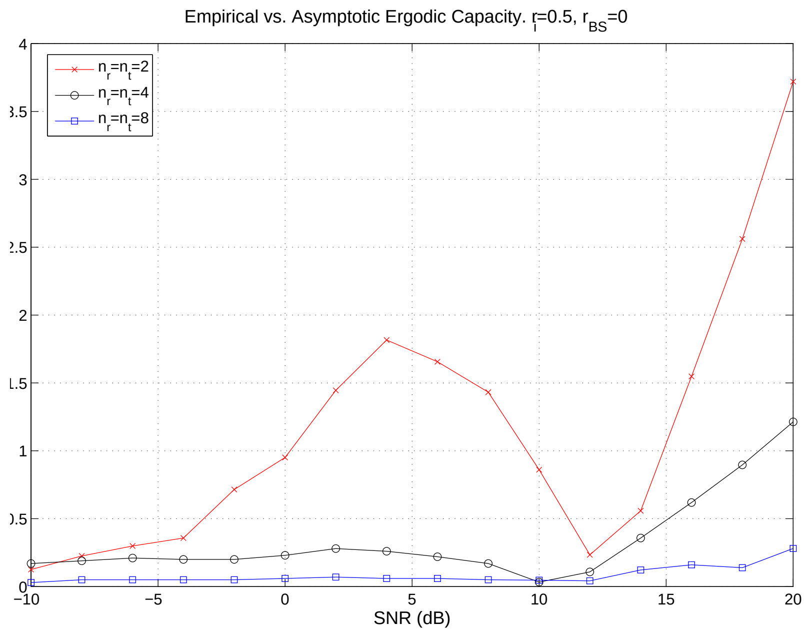

Fig. 1. Relative error [%] on the EMI as a function of SNR for different sizes of MIMO systems: $2 \times 2, 4 \times 4, 8 \times 8$ with $K = 1$, $r_1 = 0.5$, $\mathbf{R} = \mathbf{I}$.

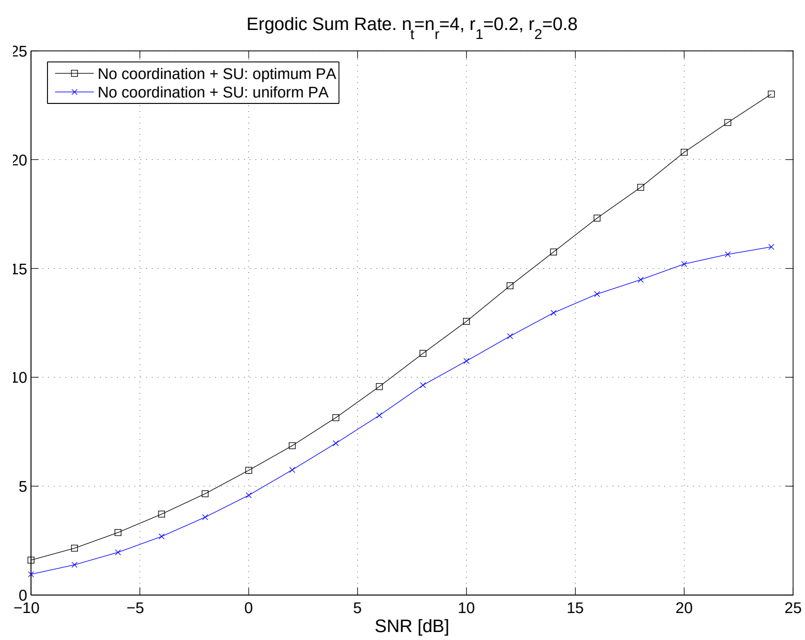

Fig. 2. Ergodic sum-rate as a function of SNR for the optimized power allocation and uniform power allocation when $K = 2, n_t = n_r = 4, r_1 = 0.2, r_2 = 0.8$ in the scenario –no coordination + single-user decoding–.

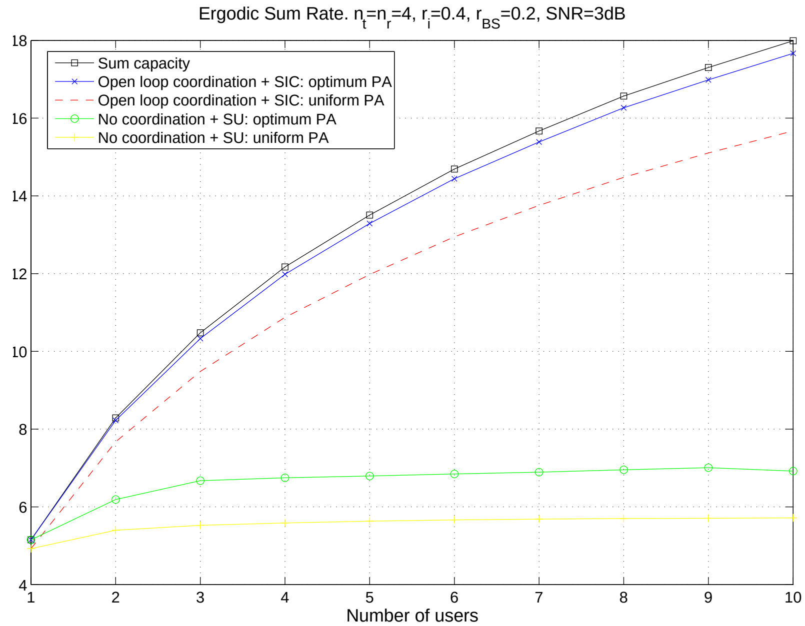

Fig. 3. Sum-rate as a function of the number or users for different power allocation schemes: 1. Team game + SIC + optimal power allocation (sum-capacity); 2. Open loop coordination + SIC + optimal power allocation; 3. Open loop coordination + SIC + uniform power allocation; 4. No coordination + SU + optimal power allocation. Setup: $n_t = n_r = 4, r_k = 0.4, r_R = 0.2, \rho = 3$ dB.

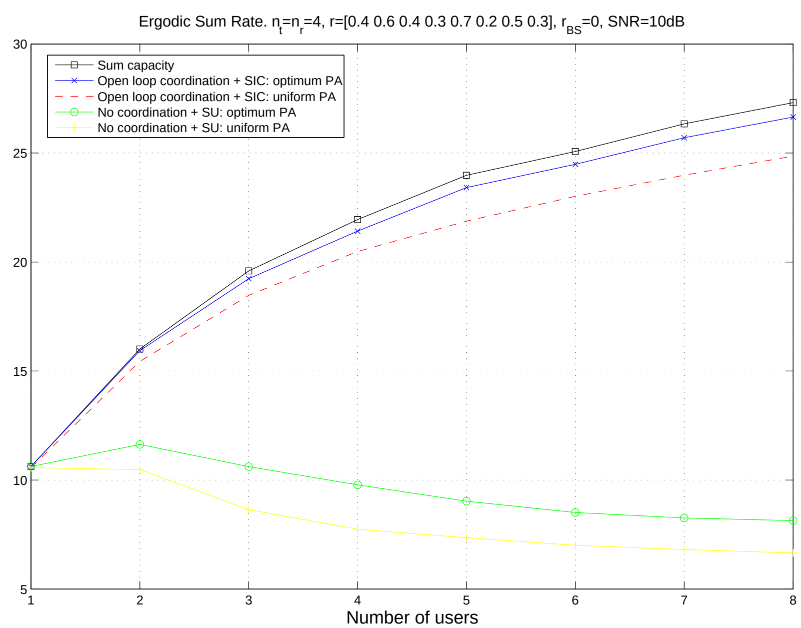

Fig. 4. Sum-rate as a function of the number or users for different power allocation schemes: 1. Team game + SIC + optimal power allocation (sum-capacity); 2. Open loop coordination + SIC + optimal power allocation; 3. Open loop coordination + SIC + uniform power allocation; 4. No coordination + SU + optimal power allocation. Setup: $n_t = n_r = 4, \underline{r} = (0.4, 0.6, 0.4, 0.3, 0.7, 0.2, 0.5, 0.3), r_R = 0, \rho = 10$ dB.