Power Allocation Games for MIMO Multiple Access Channels with Coordination

Elena Veronica Belmega, Samson Lasaulce, Merouane Debbah

Power Allocation Games for MIMO Multiple Access Channels with Coordination

Elena-Veronica Belmega, *Student Member, IEEE,* Samson Lasaulce, *Member, IEEE,* and Merouane Debbah, *Senior Member, IEEE*

Abstract

A game theoretic approach is used to derive the optimal decentralized power allocation (PA) in fast fading multiple access channels where the transmitters and receiver are equipped with multiple antennas. The players (the mobile terminals) are free to choose their PA in order to maximize their individual transmission rates (in particular they can ignore some specified centralized policies). A simple coordination mechanism between users is introduced. The nature and influence of this mechanism is studied in detail. The coordination signal indicates to the users the order in which the receiver applies successive interference cancellation and the frequency at which this order is used. Two different games are investigated: the users can either adapt their temporal PA to their decoding rank at the receiver or optimize their spatial PA between their transmit antennas. For both games a thorough analysis of the existence, uniqueness and sum-rate efficiency of the network Nash equilibrium is conducted. Analytical and simulation results are provided to assess the gap between the decentralized network performance and its equivalent virtual multiple input multiple output system, which is shown to be zero in some cases and relatively small in general.

Index Terms

Game theory, large systems, MAC, MIMO, Nash equilibrium, power allocation games, random matrix theory.

The material in this paper was presented in part at the 6th IEEE/ACM Intl. Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks and Workshops (WiOpt), Berlin, Germany, 4 April 2008 [1].

E. V. Belmega and S. Lasaulce are with LSS (joint lab of CNRS, Supélec, Paris 11), Supélec, Plateau du Moulon, 91192 Gif-sur-Yvette, France, {belmega,lasaulce}@lss.supelec.fr; M. Debbah is with the "Chaire Alcatel-Lucent" at Supélec, merouane.debbah@supelec.fr.

I. Introduction

We consider a special case of decentralized or distributed wireless networks, the decentralized multiple access channel (MAC). In this context, the MAC consists of a network of several mobile stations (MS) and one base station (BS). In the present work, the network is said to be decentralized in the sense that each user can choose freely his power allocation (PA) policy in order to selfishly maximize a certain individual performance criterion. This means that, even if the the BS broadcasts some specified policies, every (possibly cognitive) user is free to ignore the policy intended for him if the latter does not maximize his performance criterion.

The problem of decentralized PA in wireless networks is not new and has been properly formalized for the first time in [2], [3]. Interestingly, this problem can be formulated quite naturally as a non-cooperative game with different performance criteria (utilities) such as the carrier-to-interference ratio [4], aggregate throughput [5] or energy efficiency [6], [7]. In this paper, we assume that the users want to maximize information-theoretic utilities and more precisely their Shannon transmission rates. Many reasons why this kind of utilities is often considered are provided in the literature related to the problem under investigation (some references are provided further). Here we will just mention three of them. First, Shannon transmission rates allow one to characterize the performance limits of a communication system and study the behavior of (selfish) users in a network where good coding schemes are implemented. As there is a direct relationship between the achievable transmission rate of a user and his signal-to-interference plus noise ratio (SINR), they also allow one to optimize performance metrics like the SINR or related quantities of the same type (e.g., the carrier-to-interference ratio). From the mathematical point of view, Shannon rates have many desirable properties (e.g., concavity properties), which allows one to conduct deep performance analyses. Therefore they provide useful insights and concepts that are exploitable for a practical design of decentralized networks. Indeed, the point of view adopted here is close to the one proposed by the authors of [8] for DSL (digital subscriber lines) systems, which are modeled as a parallel interference channel; [9] for the single input single output (SISO) and single input multiple output (SIMO) fast fading MACs with global CSIR and global CSIT (Channel State Information at the Receiver/Transmitters); [10] for MIMO (Multiple Input Multiple Output) MACs with global CSIR, channel distribution information at the transmitters (global CDIT) and single-user decoding (SUD) at the receivers; [11], [12] for Gaussian MIMO interference channels with global CSIR and local CSIT and, by definition of the conventional interference channel [13], SUD at the receivers. Note that reference [14] where the authors considered Gaussian MIMO MACs with neither CSIT nor CDIT differs from our approach and that of [8], [9], [10], [11], [12] because in [14] the MIMO MAC is seen as a two-player zero-sum game where the first player is the group of transmitters

and the second player is the set of MIMO sub-channels. In the list of the aforementioned references, [9] seems to be the closest work to ours. However, our approach differs from [9] on several technical key points. First of all, not only the BS but also the MSs can be equipped with multiple antennas. This is an important technical difference since the power control problem of [9] becomes a PA problem for which the precoding matrix of each user has to be determined. Also the issues regarding the existence and uniqueness of the network equilibrium are more complicated to be dealt with, as it will be seen. Specifically, random matrix theory will be exploited to determine the optimum eigenvalues of the precoding matrices. In [9], several assumptions made, especially the one involving the knowledge of all the instantaneous channels at each MS can be argued in some contexts. One of our objectives is to decrease the amount of signaling needed from the BS. This is why we assume that the BS can only send to the users sufficient training signals for them to know the statistics of the different channels and a simple and common coordination signal. The underlying coordination mechanism is simple because it consists in periodically sending the realization of a $K!$-state random signal, where K is the number of active users. As it will be seen in detail, such a mechanism is mandatory because, in contrast with [10], we assume here successive interference cancellation (SIC) at the BS. Thus each user needs to know his decoding rank in order to adapt his PA policy to maximize the transmission rate. The coordination signal precisely indicates to all the users the decoding order employed by the receiver. Therefore the proposed formulation can be seen from two different standpoints. If the distribution of the coordination signal is fixed, then the addressed problem can be regarded as a non-cooperative game where the BS is imposed to follow the realizations of the random coordination signal. In this case the respective signal can be generated by any device (and not necessarily by the BS), in order to select the decoding order. On the other hand, if the distribution of the coordination signal can be optimized, the problem can be addressed as a Stackelberg game. Here the BS is the game leader and chooses his best mixed strategy (namely a distribution over the possible decoding orders) in order to maximize a certain utility, which will be chosen to be the network uplink sum-rate.

In the described framework, one of our objectives is to know how well a non-cooperative but weakly coordinated system performs in terms of overall sum-rate w.r.t. its centralized counterpart (by "centralized" we mean that the users are imposed to follow the BS PA policies) when SIC is used at the BS. In this setting, several interesting questions arise. When the users' utility functions are chosen to be their individual transmission rates, is there a Nash equilibrium (NE) in the corresponding game and is it unique? What is the optimum way for a selfish user to allocate (spatially or temporally) his transmit power? How to choose the coordination signal that maximizes the network sum-rate? What is the performance loss of the decentralized

network w.r.t. the equivalent virtual MIMO network?

This paper is structured as follows. After presenting the system model (Sec. II), we study in detail two PA games. In the first case (Sec. III), each MS is imposed to share his power uniformly between his transmit antennas but can freely allocate his power over time. In the second case (Sec. IV), we assume that the temporal PA is uniform and thus our objective is to derive the best spatial PA scheme. For each of these frameworks the existence, uniqueness, determination and sum-rate efficiency of the NE is investigated. Numerical results are provided in Sec. V to illustrate our theoretical analysis and in particular to better assess the sum-rate efficiency of the different games considered. We conclude the paper by several remarks and possible extensions of our work in Sec. VI.

II. System Model

Throughout the paper $\underline{v}$, $\mathbf{M}$, $(.)^T$ and $(.)^H$ will stand for vector, matrix, transpose and transpose conjugate, respectively. For simplicity and without loss of generality, we will assume a MAC with $K = 2$ users. Note that the type of multiple access technique assumed corresponds to the one considered in the standard definition of the Gaussian MAC by [15],[16]: all transmitters send at once and at different rates over the entire bandwidth. In this (information theoretic) context, very long codewords can be used and the receiver is not limited in terms of complexity. Thus the codewords of the different transmitters can be decoded jointly using a maximum likelihood decoding procedure (see [16] for more details). Interestingly, the transmission rates of the capacity region corresponding to the coding-decoding procedure just mentioned, can also be achieved, as discussed in [16], by using perfect SIC at the receiver. In this paper we also adopt this decoding scheme, which means that not only the different channel matrices are perfectly known to the receiver but also that the codewords of all the users are decoded reliably. The case of imperfect CSIR and error propagation in the SIC procedure is thus seen as a useful extension of this paper. Since we assume SIC at the BS and that the users want to maximize their individual transmission rates, it is necessary for them to know the decoding order used by the BS. This is why we assume the existence of a source broadcasting a discrete coordination signal to all the terminals in presence. If this source is the BS itself, this induces a certain cost in terms of downlink signaling but the distribution of the coordination signal can then be optimized. On the other hand, if the coordination signal comes from an external source, e.g., an FM transmitter, the MSs can acquire their coordination signal for free in terms of downlink signaling. However this generally involves a certain sub-optimality in terms of uplink rate. Analyzing this kind of tradeoffs is precisely one of the goals of this paper. In both cases, the coordination signal will be represented by a Bernouilli random variable denoted with $S \in \mathcal{S}$. Since we study the $2-$user

MAC, $\mathcal{S}=\{1,2\}$ is a binary alphabet and S is distributed as $\Pr[S=1]=p$, $\Pr[S=2]=1-p\triangleq\overline{p}$. Without loss of generality we assume that when the realization of S is 1, user 1 is decoded in the second place and therefore sees no multiple access interference; in a real wireless system the frequency at which the realizations would be drawn is roughly proportional to the reciprocal of the channel coherence time (T_{coh}). Note that the proposed coordination mechanism is suboptimal in the sense that the coordination signal does not depend on the realizations of the channel matrices. We will see that the corresponding performance loss is in fact very small.

We will further consider that each MS is equipped with n_t antennas whereas the BS has n_r antennas. In our analysis, the flat fading channel matrices of the different links vary from symbol vector to symbol vector. We assume that the receiver knows all the channel matrices whereas the transmitters have only access to the statistics of the different channels. At this point, the authors would like to re-emphasize their point of view:

- On the one hand, we think that in some contexts our approach can be interesting in terms of signaling cost. We have seen that S lies in a $K-$element alphabet and the realizations are drawn approximatively at $\frac{1}{T_{\mathrm{coh}}}[\mathrm{Hz}]$, therefore the coordination mechanism requires at most $\frac{\log_2(K!)}{T_{\mathrm{coh}}}$ bps from the BS and 0 bps if it is built from an external source. Another source of signaling cost is the acquisition of the knowledge of the statistics of the uplink channels at the MSs. For example, in the context of coherent communications where the BS regularly sends some data to the MSs and channel reciprocity assumption is valid (e.g., in time division duplex systems) the corresponding cost can be reasonable. In general, this cost will have to be compared to the cost of the centralized system where the BS has to send accurate enough quantized versions of the (possibly large) precoding matrices at a certain frequency.
- On the other hand, even if our approach is not interesting in terms of signaling, it can be very useful in contexts where terminals are autonomous and may have some selfish reasons to deviate from the centralized policies. In such scenarios, the concept of network equilibrium is of high importance.

The equivalent baseband signal received by the BS can be written as:

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

where $\underline{x}_k^{(s)}(\tau)$ is the n_t-dimensional column vector of symbols transmitted by user k at time τ for the realization $s\in\mathcal{S}$ of the coordination signal, $\mathbf{H}_k(\tau)\in\mathbb{C}^{n_r\times n_t}$ is the channel matrix (stationary and ergodic process) of user k and $\underline{z}^{(s)}(\tau)$ is an n_r-dimensional complex white Gaussian noise distributed as $\mathcal{N}(\underline{0},\sigma^2\mathbf{I}_{n_r})$; for sake of clarity we will omit the time index τ from our notations. As [17] we assume that, for each $s\in\mathcal{S}$, the data streams of

user k are multiplexed in the eigen-directions of the matrix $\mathbf{Q}_k^{(s)} = \mathbb{E}\left[\underline{x}_k^{(s)}\underline{x}_k^{(s),H}\right] \triangleq \mathbf{V}_k^{(s)}\mathbf{P}_k^{(s)}\mathbf{V}_k^{(s),H}$. Finding the optimal eigen-values $\mathbf{P}_k^{(s)}$ and coordinate systems $\mathbf{V}_k^{(s)}$ that maximize the transmission rate of user k is one of the main issues we will solve in the next two sections. In order to take into account the antenna correlation effects at the transmitters and receiver, we assume the different channel matrices to be structured according to the Kronecker propagation model [18] with common receive correlation [19]:

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

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 $\frac{1}{n_t}$. The motivation for assuming a channel model with common receive correlation is twofold. First, there exist some situations where this MIMO MAC model is realistic, the most simple situation being the case of no receive correlation i.e., $\mathbf{R} = \mathbf{I}$ (see e.g., [20]). Although it is not explicitly stated in [19] the second feature of this model is that the overall channel matrix $\mathbf{H} = [\mathbf{H}_1 ... \mathbf{H}_K]$ can also be factorized as a Kronecker model, which will allow us to re-exploit existing results from the random matrix theory literature. Therefore the case where the overall channel matrix is not separable can be seen as a possible extension of this paper that can be dealt with by using the results in [21].

In this paper we study in detail two special but useful cases of decentralized PA problems. In the first case (Game 1), we assume (for instance because of practical technical/complexity constraints) that each user is imposed to share his power uniformly between his transmit antennas but can freely allocate his power over time; this problem will be referred to as temporal PA game (Sec. III). In the second case (Game 2), for every realization of the coordination signal, each user is assumed to transmit with the same total power (denoted by P_k) but can freely share it between his antennas; this problem will be referred to as spatial PA game (Sec. IV). For both games the *strategy* of user $k \in \{1, 2\}$ consists in choosing the distribution of $\underline{x}_k^{(s)}$, for each $s \in \mathcal{S}$ in order to maximize his *utility* function which is given by:

$$u_k(\mathbf{Q}_1^{(1)}, \mathbf{Q}_1^{(2)}, \mathbf{Q}_2^{(1)}, \mathbf{Q}_2^{(2)}) \quad = \quad \sum_{s=1}^{2} \Pr[S = s] R_k^{(s)}(\mathbf{Q}_1^{(s)}, \mathbf{Q}_2^{(s)}) \tag{3}$$

where

$$R_k^{(s)}(\mathbf{Q}_1^{(s)}, \mathbf{Q}_2^{(s)}) = \left| \begin{array}{ll} \mathbb{E}\log|\mathbf{I} + \eta\mathbf{H}_k\mathbf{Q}_k^{(s)}\mathbf{H}_k^H| & \text{if } k = s \\ \mathbb{E}\log|\mathbf{I} + \eta\sum_{k=1}^{2}\mathbf{H}_k\mathbf{Q}_k^{(s)}\mathbf{H}_k^H| - \mathbb{E}\log|\mathbf{I} + \eta\mathbf{H}_{-k}\mathbf{Q}_{-k}^{(s)}\mathbf{H}_{-k}^H| & \text{if } k \neq s \end{array} \right. \tag{4}$$

with $\eta \triangleq \frac{1}{\sigma^2}$ and the usual notation for $-k$, which stands for the other user than k. Note that we implicitly assume Gaussian codebooks for the two users since this choice is optimum in terms of their individual Shannon

transmission rates (see e.g., [22]). This is why the strategy of a user boils down to choosing the best pair of covariance matrices $(\mathbf{Q}_k^{(1)}, \mathbf{Q}_k^{(2)})$. The corresponding maximization is performed under the following transmit power constraint for each MS: $\mathrm{Tr}\left(\sum_{s=1}^{2} \Pr[S=s]\mathbf{Q}_k^{(s)}\right) \leq n_t P_k$. The main difference between Games 1 and 2 relies precisely on how this general power constraint is specialized. In Game 1, the precoding matrices are imposed to have the following structure: $\forall k \in \{1,2\}, \forall s \in \{1,2\}, \mathbf{Q}_k^{(s)} = \alpha_k^{(s)} P_k \mathbf{I}_{n_t}$, which amounts to rewriting the total power constraint as follows

$$\sum_{s=1}^{2} \Pr[S=s]\alpha_k^{(s)} \leq 1. \tag{5}$$

On the other hand, in Game 2, the power constraint expresses as

$$\forall k \in \{1,2\}, \forall s \in \{1,2\}, \mathrm{Tr}(\mathbf{Q}_k^{(s)}) \leq n_t P_k. \tag{6}$$

In both game frameworks, an important issue for a wireless network designer/owner is to know whether by leaving the users decide their PA by themselves, the network is going to operate at a given and predictable state. This precisely corresponds to the notion of a network equilibrium, a state from which no user has interest to deviate. The main issue is to know if there exists an equilibrium point, whether it is unique, how to determine the corresponding strategies and characterize the efficiency of this equilibrium in terms of network sum-rate.

III. Temporal power allocation game

As mentioned above, in the temporal power allocation (TPA) game, the strategy of user $k \in \{1,2\}$ merely consists in choosing the best pair $(\alpha_k^{(1)}, \alpha_k^{(2)})$. Since each transmission rate is a concave and non-decreasing function of the $\alpha_k^{(s)}$'s, each user will saturate the power constraint (5) i.e., $\sum_{s=1}^{2} \Pr[S=s]\alpha_k^{(s)} = 1$, which leads to optimizing a single parameter $\alpha_k^{(1)}$ or $\alpha_k^{(2)}$. From now on, for sake of clarity we will use the notations $\alpha_1^{(1)} = \alpha_1$, $\alpha_2^{(2)} = \alpha_2$. Indeed, it is easy to verify that the power constraints are characterized completely, for the first user by $\alpha_1^{(2)} = \frac{1-p\alpha_1}{1-p}$ with $\alpha_1 \in \mathcal{A}_1^{\mathrm{TPA}} \triangleq \left[0, \frac{1}{p}\right]$, and for the second user by $\alpha_2^{(1)} = \frac{1-(1-p)\alpha_2}{p}$ with $\alpha_2 \in \mathcal{A}_2^{\mathrm{TPA}} \triangleq \left[0, \frac{1}{1-p}\right]$. Thus the strategy of user $k \in \{1,2\}$ consists in choosing the best fraction α_k from the action set $\mathcal{A}_k^{\mathrm{TPA}}$. Our main goal is to investigate if there exists an NE and determine the corresponding profile of strategies $(\alpha_1^{\mathrm{NE}}, \alpha_2^{\mathrm{NE}})$. It turns out that the issues of the existence and uniqueness of an NE can be properly dealt with by applying Theorems 1 and 2 of [23] in our context. For making this paper sufficiently self-contained, we review here these two theorems (Theorem 2 is given for the $2-$user case for simplicity and because it is sufficient under our assumptions).

Theorem 1: [23] *Let $\mathcal{G} = (\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{u_k\}_{k\in\mathcal{K}})$ be a game where $\mathcal{K} = \{1, ..., K\}$ is the set of players, $\mathcal{A}_1, ..., \mathcal{A}_K$ the corresponding sets of strategies and $u_1, ..., u_k$ the utilities of the different players. If the following*

three conditions are satisfied: (i) each u_k is continuous in the vector of strategies $(a_1, ..., a_K) \in \prod_{k=1}^{K} \mathcal{A}_k$; (ii) each u_k is concave in $a_k \in \mathcal{A}_k$; (iii) $\mathcal{A}_1, ..., \mathcal{A}_K$ are compact and convex sets; then $\mathcal{G}$ has at least one NE.

Theorem 2: [23] *Consider the K-player concave game of Theorem 1 with $K = 2$. If the following (diagonally strict concavity) condition is met: for all $(a_1', a_1'') \in \mathcal{A}_1^2$ and $(a_2', a_2'') \in \mathcal{A}_2^2$ such that $(a_1', a_2') \neq (a_1'', a_2''), (a_1'' - a_1') \left[\frac{\partial u_1}{\partial a_1}(a_1', a_2') - \frac{\partial u_1}{\partial a_1}(a_1'', a_2'')\right] + (a_2'' - a_2') \left[\frac{\partial u_2}{\partial a_2}(a_1', a_2') - \frac{\partial u_2}{\partial a_2}(a_1'', a_2'')\right] > 0$; then the uniqueness of the NE is insured.*

At this point we can state the first main result of this paper, which is provided in the following theorem. For sake of clarity we will also use the notations: $p_k \triangleq p$ if $k = 1$ or $p_k \triangleq \overline{p}$ if $k = 2$.

Theorem 3 (Existence and uniqueness of an NE in Game 1): the temporal PA game described by: the set of players $\mathcal{K} = \{1, 2\}$; the sets of actions $\mathcal{A}_k^{\mathrm{TPA}} = \left[0, \frac{1}{p_k}\right]$ and utilities $u_k(\alpha_k, \alpha_{-k}) = pR_k^{(1)}(\alpha_k, \alpha_{-k}) + \overline{p}R_k^{(2)}(\alpha_k, \alpha_{-k})$, where the rates $R_k^{(s)}$ follow from Eq. (4) has a unique NE.

Proof:

Existence of an NE. It is guaranteed by the geometrical and topological properties of the utility functions and the strategy sets of the users (over which the maximization is performed). Indeed, we can apply [23] in our matrix case. Without loss of generality, let us consider user 1. The utility of user 1 comprises two terms corresponding to the two coordination signal realizations: $u_1(\alpha_1, \alpha_2) = pR_1^{(1)}(\cdot, \cdot) + \overline{p}R_1^{(2)}(\cdot, \cdot)$. Using the fact that $\frac{d^2}{dt^2} \log |\mathbf{X} + t\mathbf{Y}\mathbf{Y}^H| = -\mathrm{Tr}\left[\mathbf{Y}^H(\mathbf{X} + t\mathbf{Y}\mathbf{Y}^H)^{-1}\mathbf{Y}\mathbf{Y}^H(\mathbf{X} + t\mathbf{Y}\mathbf{Y}^H)^{-1}\mathbf{Y}\right]$ it is easy to verify that $\frac{\partial^2 R_1^{(1)}}{\partial \alpha_1^2}(\alpha_1, \alpha_2) = -\mathbb{E}\mathrm{Tr}[\mathbf{B}\mathbf{B}^H] < 0$ and $\frac{\partial^2 R_1^{(2)}}{\partial \alpha_1^2}(\alpha_1, \alpha_2) = -\mathbb{E}\mathrm{Tr}[\mathbf{C}\mathbf{C}^H] < 0$ where $\mathbf{B} = \rho_1 \mathbf{H}_1^H (\mathbf{I} + \rho_1 \alpha_1 \mathbf{H}_1 \mathbf{H}_1^H)^{-1} \mathbf{H}_1$, $\mathbf{C} = \frac{p}{\overline{p}} \rho_1 \mathbf{H}_1^H \left(\mathbf{I} + \rho_1 \frac{1 - p\alpha_1}{\overline{p}} \mathbf{H}_1 \mathbf{H}_1^H + \rho_2 \alpha_2 \mathbf{H}_2 \mathbf{H}_2^H\right)^{-1} \mathbf{H}_1$ and $\rho_1 = \eta P_1$, $\rho_2 = \eta P_2$ correspond to the signal-to-noise ratios of the users. Thus for every user k, the utility u_k is strictly concave w.r.t. to α_k. Also it is continuous in (α_1, α_2) over the convex and compact strategy sets $\mathcal{A}_k^{\mathrm{TPA}}$. Therefore the existence of at least one NE is guaranteed. Interestingly, we observe that for a fixed game rule, which is the value of the parameter p, there will always be an equilibrium. The users adapt their strategies to the rule of the game in order to optimize their individual transmission rates.

Uniqueness of the NE. We always apply [23] in our matrix case (see Appendix A) and prove that the diagonally strict concavity condition is actually met. The key of the proof is the following Lemma which is proven in Appendix B.

Lemma 1: Let $\mathbf{A}'$, $\mathbf{A}''$, $\mathbf{B}'$ and $\mathbf{B}''$ be Hermitian and non-negative matrices such that either $\mathbf{A}' \neq \mathbf{A}''$ or $\mathbf{B}' \neq \mathbf{B}''$. Assume that the classical matrix order $\succeq$ is total for each of the pairs of matrices $(\mathbf{A}', \mathbf{A}'')$ and $(\mathbf{B}', \mathbf{B}'')$ i.e., either $\mathbf{A}' \succeq \mathbf{A}''$ (resp. $\mathbf{B}' \succeq \mathbf{B}''$) or $\mathbf{A}'' \succ \mathbf{A}'$ (resp. $\mathbf{B}'' \succ \mathbf{B}'$). Then we have $\mathrm{Tr}(\mathbf{M} + \mathbf{N}) \geq 0$ with $\mathbf{M} = (\mathbf{A}'' - \mathbf{A}') \left[(\mathbf{I} + \mathbf{A}')^{-1} - (\mathbf{I} + \mathbf{A}'')^{-1}\right]$, $\mathbf{N} = (\mathbf{B}'' - \mathbf{B}') \left[(\mathbf{I} + \mathbf{B}' + \mathbf{A}')^{-1} - (\mathbf{I} + \mathbf{B}'' + \mathbf{A}'')^{-1}\right]$.

It can be shown (see Appendix A for more details) that the diagonally strict concavity condition writes in our setup as $p\mathcal{T}^{(1)}+\overline{p}\mathcal{T}^{(2)}>0$ where $\forall s\in\{1,2\}$, $\mathcal{T}^{(s)}$ is defined by $\mathcal{T}^{(s)}=\mathrm{Tr}(\mathbf{M}^{(s)}+\mathbf{N}^{(s)})$ where the matrices $\mathbf{M}^{(s)}$, $\mathbf{N}^{(s)}$ have exactly the same structure as $\mathbf{M}$, $\mathbf{N}$ in the above Lemma. For example, if we consider two pairs of parameters $(\alpha_1',\alpha_1'')\in\left(\mathcal{A}_1^{\mathrm{TPA}}\right)^2$ and $(\alpha_2',\alpha_2'')\in\left(\mathcal{A}_2^{\mathrm{TPA}}\right)^2$ such that either $\alpha_1'\neq\alpha_1''$ or $\alpha_2'\neq\alpha_2''$ as in Theorem 2, $\mathcal{T}^{(1)}$ can be obtained by using the following matrices $\mathbf{A}'=\rho_1\alpha_1'\mathbf{H}_1\mathbf{H}_1^H$, $\mathbf{A}''=\rho_1\alpha_1''\mathbf{H}_1\mathbf{H}_1^H$, $\mathbf{B}'=\rho_2\frac{1-\overline{p}\alpha_2'}{p}\mathbf{H}_2\mathbf{H}_2^H$, $\mathbf{B}''=\rho_2\frac{1-\overline{p}\alpha_2''}{p}\mathbf{H}_2\mathbf{H}_2^H$. The term $\mathcal{T}^{(2)}$ has a similar form as $\mathcal{T}^{(1)}$ thus, applying Lemma 1 twice and considering the special structure of the four matrices ($\mathbf{A}'$, $\mathbf{A}''$, $\mathbf{B}'$, $\mathbf{B}''$), one can prove that the term $p\mathcal{T}^{(1)}+\overline{p}\mathcal{T}^{(2)}$ is strictly positive. Therefore the unconditional uniqueness of the NE is guaranteed. ■

Determination of the NE. In order to determine the selfish PA of the users at the NE, we now exploit the large system approach derived in [24] for single-user fading MIMO channels. This will lead us to simple approximations of the utility functions which are much easier to optimize. From now on, we assume the asymptotic regime in terms of the number of antennas: $n_t\longrightarrow\infty$, $n_r\longrightarrow\infty$, and $\lim_{n_t\to\infty,n_r\to\infty}\frac{n_t}{n_r}=c<\infty$. In this asymptotic regime, references [24], [25], [26] provide an equivalent of the ergodic capacity of single-user MIMO channels, which corresponds exactly to the situation seen by user 1 (resp. 2) when $S=1$ (resp. $S=2$); this gives directly the approximation of the rates $R_1^{(1)}$ and $R_2^{(2)}$; see Eq. (4). From Eq. (4) we also see that the rates $R_1^{(2)}$ and $R_2^{(1)}$ correspond to the difference between the sum-rate of the equivalent $Kn_t\times n_r$ virtual MIMO system and an $n_t\times n_r$ single-user MIMO system, therefore the results of [24], [25], [26] can also be applied directly. The corresponding approximates can then be easily checked to be:

$$\begin{aligned}
\tilde{R}_1^{(1)}(\alpha_1,\alpha_2) &= \sum_{i=1}^{n_t}\log_2\left[1+\eta\alpha_1P_1d_1^{(\mathrm{T})}(i)\gamma_1\right]+\sum_{j=1}^{n_r}\log_2\left[1+\eta d^{(\mathrm{R})}(j)\delta_1\right]-n_t\eta\gamma_1\delta_1\log_2 e\\
\tilde{R}_2^{(1)}(\alpha_1,\alpha_2) &= \sum_{i=1}^{n_t}\log_2\left[1+2\eta\alpha_1P_1d_2^{(\mathrm{T})}(i)\gamma_2\right]+\sum_{i=1}^{n_t}\log_2\left[1+2\eta\frac{1-\overline{p}\alpha_2}{p}P_2d_2^{(\mathrm{T})}(i)\gamma_2\right]\\
&\quad+\sum_{j=1}^{n_r}\log_2\left[1+2\eta d^{(\mathrm{R})}(j)\delta_2\right]-4n_t\eta\gamma_2\delta_2\log_2 e-\tilde{R}_1^{(1)}(\alpha_1,\alpha_2).
\end{aligned}\tag{7}$$

where $\forall k\in\{1,2\}, d_k^{(\mathrm{T})}(i)$, $i\in\{1,...,n_t\}$ are the eigenvalues of the transmit correlation matrices $\mathbf{T}_k$ (see Eq. (2)), $d^{(\mathrm{R})}(j)$, $j\in\{1,...,n_r\}$, are the eigenvalues of the receive correlation matrix $\mathbf{R}$ and the parameters γ_i, δ_j are the unique solutions of the following systems of $2-$degree equations:

$$\left\{\begin{array}{rcl}\gamma_1 &=& \dfrac{1}{n_t}\displaystyle\sum_{j=1}^{n_r}\frac{d^{(\mathrm{R})}(j)}{1+\eta d^{(\mathrm{R})}(j)\delta_1}\\ \delta_1 &=& \dfrac{1}{n_t}\displaystyle\sum_{i=1}^{n_t}\frac{\alpha_1P_1d_1^{(\mathrm{T})}(i)}{1+\eta\alpha_1P_1d_1^{(\mathrm{T})}(i)\gamma_1}\end{array}\right.\tag{8}$$

$$\left\{\begin{array}{rcl}\gamma_2 & = & \dfrac{1}{2n_t}\displaystyle\sum_{j=1}^{n_r}\frac{d^{(\mathrm{R})}(j)}{1+2\eta d^{(\mathrm{R})}(j)\delta_2}\\ \delta_2 & = & \dfrac{1}{2n_t}\left[\displaystyle\sum_{i=1}^{n_t}\frac{\alpha_1 P_1 d_1^{(\mathrm{T})}(i)}{1+2\eta\alpha_1 P_1 d_1^{(\mathrm{T})}(i)\gamma_2}+\sum_{i=1}^{n_t}\frac{\frac{1-\overline{p}\alpha_2}{p}P_2 d_2^{(\mathrm{T})}(i)}{1+2\eta\frac{1-\overline{p}\alpha_2}{p}P_2 d_2^{(\mathrm{T})}(i)\gamma_2}\right].\end{array}\right. \tag{9}$$

The approximate functions $\tilde{R}_1^{(2)}(\cdot,\cdot)$ and $\tilde{R}_2^{(2)}(\cdot)$ can be obtained in a similar way and the approximated utility of user $k\in\{1,2\}$ follows: $\tilde{u}_k(\alpha_1,\alpha_2)=p\tilde{R}_k^{(1)}(\alpha_1,\alpha_2)+\overline{p}\tilde{R}_k^{(2)}(\alpha_1,\alpha_2)$. Now, in order to solve the constrained optimization problem, we introduce the Lagrange multipliers $(\lambda_{11},\lambda_{12},\lambda_{21},\lambda_{22})\in[0,+\infty)^4$ and define for $k\in\{1,2\}$ the function $\mathcal{L}_k(\alpha_1,\alpha_2,\lambda_{k1},\lambda_{k2})=-\tilde{u}_k(\alpha_1,\alpha_2)+\lambda_{k1}\left(\alpha_k-\frac{1}{p_k}\right)-\lambda_{k2}\alpha_k$. The Kuhn-Tucker optimality conditions follow. Therefore, the optimum selfish PAs, $(\alpha_1^{\mathrm{NE}},\alpha_2^{\mathrm{NE}})$, can be obtained by using a fixed-point method and an iterative algorithm, following the same idea as in [10] for non-coordinated MIMO MACs with single-user decoding. At this point we have to make an important technical comment. Our proof for the existence and uniqueness of the NE holds for the exact game. For the approximated game, we need the approximated utilities to have the same properties as their exact counterparts. It turns out that the large system approximation of the ergodic mutual information can be shown to have the desired properties [27]. In particular, the results of [27] show that the approximated utilities are strictly concave and that if the iterative PA algorithm converges, it converges towards the global maximum.

Sum-rate efficiency of the NE. Now, let us focus on the sum-rate of the decentralized network and compare it with the optimal sum-rate of its centralized counterpart. The centralized network sum-rate, denoted by $R_{\mathrm{sum}}^{(\mathrm{C})}$, is by definition obtained by jointly maximizing the sum-rate over all the pairs of power fractions $(\alpha_1,\alpha_2)\in[0,1]^2$: $R_{\mathrm{sum}}^{(\mathrm{C})}\triangleq\max_{(\alpha_1,\alpha_2)}u_1(\alpha_1,\alpha_2)+u_2(\alpha_1,\alpha_2)$. Knowing that $\log|\cdot|$ is a concave function, one can easily verify that the maximum is obtained for $(\alpha_1^*,\alpha_2^*)=(1,1)$ and that $R_{\mathrm{sum}}^{(C)}=\mathbb{E}\log\left|\mathbf{I}+\rho_1\mathbf{H}_1\mathbf{H}_1^H+\rho_2\mathbf{H}_2\mathbf{H}_2^H\right|$. As the optimum precoding matrices are proportional to the identity matrix, it can be checked that the network sum-rate at the NE (denoted by $R_{\mathrm{sum}}^{\mathrm{NE}}$) is equal to the centralized network sum-rate for $p=0$ and $p=1$: $R_{\mathrm{sum}}^{\mathrm{NE}}(0)=R_{\mathrm{sum}}^{\mathrm{NE}}(1)=R_{\mathrm{sum}}^{(\mathrm{C})}$. Indeed, let us consider that $p=1$. In this case, user 1 is always decoded in the second place ($\Pr[S=1]=1$). This means that there is no temporal power allocation game here and each user always allocates all of his available power for the case where $S=1$: $(\alpha_1^{\mathrm{NE}},\alpha_2^{\mathrm{NE}})=(1,0)$. Replacing in Eq. (4) the corresponding correlation matrices: $\mathbf{Q}_1^{(1)}=\mathbf{I}_{n_t}$, $\mathbf{Q}_2^{(1)}=\mathbf{I}_{n_t}$ and $\mathbf{Q}_1^{(2)}=\mathbf{O}_{n_t}$ (the square zero matrix), $\mathbf{Q}_2^{(2)}=\mathbf{O}_{n_t}$ we obtain that $R_{\mathrm{sum}}^{\mathrm{NE}}(1)=R_{\mathrm{sum}}^{(\mathrm{C})}$.

In the *high SNR regime*, where $\eta\to\infty$, we obtain from (8),(9) that $\eta\delta_1\to\frac{1}{\gamma_1}$, $\eta\delta_2\to\frac{1}{2\gamma_2}$ and thus γ_1 and γ_2 are the unique solutions of the following equations: $\frac{1}{n_t}\sum_{j=1}^{n_r}\frac{d^{(\mathrm{R})}(j)}{\gamma_1+d^{(\mathrm{R})}(j)}=1$, $\frac{1}{2n_t}\sum_{j=1}^{n_r}\frac{d^{(\mathrm{R})}(j)}{\gamma_2+d^{(\mathrm{R})}(j)}=1$. The

approximated utilities become:

$$\begin{aligned}\tilde{R}_1^{(1)}(\alpha_1,\alpha_2) &= \sum_{i=1}^{n_t}\log_2\left[1+\eta\alpha_1 P_1 d_1^{(\mathrm{T})}(i)\gamma_1\right]+\sum_{j=1}^{n_r}\log_2\left[1+\frac{d^{(\mathrm{R})}(j)}{\gamma_1}\right]-n_t\log_2 e\\ \tilde{R}_2^{(1)}(\alpha_1,\alpha_2) &= \sum_{i=1}^{n_t}\log_2\left[1+2\eta\alpha_1 P_1 d_1^{(\mathrm{T})}(i)\gamma_2\right]+\sum_{i=1}^{n_t}\log_2\left[1+2\eta\frac{1-\overline{p}\alpha_2}{p}P_2 d_2^{(\mathrm{T})}(i)\gamma_2\right]+\sum_{j=1}^{n_r}\log_2\left[1+\frac{d^{(\mathrm{R})}(j)}{\gamma_2}\right]\\ &\quad -2n_t\log_2 e-\tilde{R}_1^{(1)}(\alpha_1,\alpha_2).\end{aligned} \tag{10}$$

By setting the derivatives of $\tilde{u}_1(\cdot,\cdot)$ w.r.t. α_1 and $\tilde{u}_2(\cdot,\cdot)$ w.r.t. α_2 to zero, we obtain that, for each user, the PA at the NE is the uniform PA $(\alpha_1^{\mathrm{NE}},\alpha_2^{\mathrm{NE}})=(1,1)$, regardless of the distribution of the coordination signal $p\in[0,1]$. Therefore, at the equilibrium, we have that

$$\begin{aligned}R_{\mathrm{sum}}^{\mathrm{NE}}(p) &= pR_1^{(1)}(\alpha_1^{\mathrm{NE}},\alpha_2^{\mathrm{NE}})+\overline{p}R_1^{(2)}(\alpha_1^{\mathrm{NE}},\alpha_2^{\mathrm{NE}})+pR_2^{(1)}(\alpha_1^{\mathrm{NE}},\alpha_2^{\mathrm{NE}})+\overline{p}R_2^{(2)}(\alpha_1^{\mathrm{NE}},\alpha_2^{\mathrm{NE}})\\ &= p\mathbb{E}\log|\mathbf{I}+\rho_1\mathbf{H}_1\mathbf{H}_1^H|+\overline{p}\mathbb{E}\log|\mathbf{I}+\rho_1\mathbf{H}_1\mathbf{H}_1^H+\rho_2\mathbf{H}_2\mathbf{H}_2^H|-\overline{p}\mathbb{E}\log|\mathbf{I}+\rho_2\mathbf{H}_2\mathbf{H}_2^H| \qquad (11)\\ &\quad +p\mathbb{E}\log|\mathbf{I}+\rho_1\mathbf{H}_1\mathbf{H}_1^H+\rho_2\mathbf{H}_2\mathbf{H}_2^H|-p\mathbb{E}\log|\mathbf{I}+\rho_1\mathbf{H}_1\mathbf{H}_1^H|+\overline{p}\mathbb{E}\log|\mathbf{I}+\rho_2\mathbf{H}_2\mathbf{H}_2^H|\\ &= R_{\mathrm{sum}}^{(\mathrm{C})}.\end{aligned}$$

Knowing that the uniform spatial PA is optimal in the high SNR regime [17], [10], the centralized network sum-rate coincides with the sum-capacity of the centralized MAC channel, $R_{\mathrm{sum}}^{(\mathrm{C})}=C_{\mathrm{sum}}$.

In the *low SNR regime*, where $\eta\to 0$, we obtain from (8), (9) that $\eta\delta_1\to 0$, $\eta\delta_2\to 0$ and thus $\gamma_1=\frac{1}{n_t}\sum_{j=1}^{n_r}d^{(\mathrm{R})}(j)$, $\gamma_2=\frac{1}{2n_t}\sum_{j=1}^{n_r}d^{(\mathrm{R})}(j)$. Approximating $\ln(1+x)\approx x$ for $x<<1$, the achievable rates become:

$$\begin{aligned}\tilde{R}_1^{(1)}(\alpha_1) &= \frac{1}{n_t}\eta P_1\alpha_1\sum_{j=1}^{n_r}d^{(\mathrm{R})}(j)\sum_{i=1}^{n_t}d_1^{(\mathrm{T})}(i)\log_2 e\\ \tilde{R}_2^{(1)}(\alpha_1,\alpha_2) &= \frac{1}{n_t}\eta\frac{1-\overline{p}\alpha_2}{p}P_2\sum_{j=1}^{n_r}d^{(\mathrm{R})}(j)\sum_{i=1}^{n_t}d_1^{(\mathrm{T})}(i)\log_2 e\end{aligned}. \tag{12}$$

We see that the utilities $\tilde{u}_k(\alpha_1,\alpha_2)=\frac{1}{n_t}\eta P_k\sum_{j=1}^{n_r}d^{(\mathrm{R})}(j)\sum_{i=1}^{n_t}d_k^{(\mathrm{T})}(i)\log_2 e$ converge and the network sum-rate at the NE coincides here again with the centralized network sum-rate:
$R_{\mathrm{sum}}^{(\mathrm{C})}=\frac{1}{n_t}\sum_{j=1}^{n_r}d^{(\mathrm{R})}(j)\left(\eta P_1\sum_{i=1}^{n_t}d_1^{(\mathrm{T})}(i)+\eta P_2\sum_{i=1}^{n_t}d_2^{(\mathrm{T})}(i)\right)\log_2 e$. In this case also, the price of anarchy [28] is minimal for any distribution of the coordination signal.

To sum up we have seen that there is no loss of optimality in terms of sum-rate by decentralizing the PA procedure in at least four special cases: 1) $p=0$; 2) $p=1$; 3) when $\eta\to\infty$ for any $p\in[0,1]$; 4) when $\eta\to 0$ for any $p\in[0,1]$. Additionally, in case 3), since there is no loss by imposing the spatially uniform PA [17], [10], the centralized (and cooperative) MAC sum-capacity is achieved. If we further assume that there is no

correlation among the transmit antennas, $\mathbf{T}_k = \mathbf{I}$, the uniform spatial PA is optimal [17] for any η. Thus, the centralized sum-rate is always identical to the sum-capacity of the centralized MAC channel, $R_{\text{sum}}^{(\text{C})} = C_{\text{sum}}$. This means that if the BS chooses to use a completely unfair SIC-based decoding scheme, the selfish behavior of the users will always lead to achieving the centralized sum-capacity. This result is in agreement with [9], where the authors have proposed a water-filling game for the fast fading SISO MAC (assuming perfect CSIT and CSIR) and shown that the equilibrium sum-rate is equal to the maximum sum-rate point of the capacity region. However, as opposed to the SISO MAC with the proposed coordination mechanism [1], the decentralized MIMO MAC with coordination does not achieve the sum-rate of the equivalent virtual MIMO network for any value of p and for an arbitrary noise level at the BS. In particular, the fair choice $p = \frac{1}{2}$ is not optimal. We will quantify the corresponding performance gap through simulation results. Furthermore, in the low and high SNR regimes, the centralized sum-capacity is also achieved for any value of p. The consequence of these results is that any binary coordination signal can be used without loss of global optimality.

IV. Spatial power allocation game

In this section, we assume that the users are free to share their transmit power between their antennas but for each realization of the coordination signal the transmit power is constrained by Eq. (6). In other words we assume that the users cannot distribute their power over time: they cannot decide the amount of power they dedicate to a given realization of the coordination signal. As a consequence of this power constraint (Eq. (6)), the two precoding matrices that each user needs to choose can be optimized independently and each of them does not depend on p. Consider for example user 1. Its objective is to maximize its own payoff (Eq. (3)):

$$\begin{aligned} \max_{\mathbf{Q}_1^{(1)},\mathbf{Q}_1^{(2)}} u_1(\mathbf{Q}_1^{(s)},\mathbf{Q}_2^{(s)}) &= \max_{\mathbf{Q}_1^{(1)},\mathbf{Q}_1^{(2)}} \left\{ pR_1^{(1)}(\mathbf{Q}_1^{(1)}) + (1-p)R_1^{(2)}(\mathbf{Q}_1^{(2)},\mathbf{Q}_2^{(2)}) \right\} \\ &= p \max_{\mathbf{Q}_1^{(1)}} R_1^{(1)}(\mathbf{Q}_1^{(1)}) + (1-p) \max_{\mathbf{Q}_1^{(2)}} R_1^{(2)}(\mathbf{Q}_1^{(2)},\mathbf{Q}_2^{(2)}), \end{aligned} \tag{13}$$

where the last inequality follows directly form the power constraint (Eq. (6)). The strategy set of user k in the spatial PA (SPA) game is:

$$\begin{aligned} \mathcal{A}_k^{\text{SPA}} = \Big\{ \mathbf{Q}_k = (\mathbf{Q}_k^{(1)},\mathbf{Q}_k^{(2)}) \quad | \quad & \mathbf{Q}_k^{(1)} \succeq 0, \mathbf{Q}_k^{(1)} = \mathbf{Q}_k^{(1)H}, \text{Tr}(\mathbf{Q}_k^{(1)}) \leq n_t P_k \\ & \mathbf{Q}_k^{(2)} \succeq 0, \mathbf{Q}_k^{(2)} = \mathbf{Q}_k^{(2)H}, \text{Tr}(\mathbf{Q}_k^{(2)}) \leq n_t P_k \Big\}. \end{aligned} \tag{14}$$

Theorem 4 (Existence and uniqueness of an NE in Game 2): The SPA game defined by the set of players $\mathcal{K} = \{1,2\}$, the strategy sets $\mathcal{A}_k^{(\text{SPA})}$ and utilities $u_k(\alpha_k,\alpha_{-k})$ given by Eq. (3), has a unique NE.

Proof: The main feature of the game under the aforementioned power constraint is that there exists a unique NE in each sub-game defined by the realization of the coordination signal. The proof is much simpler than

that of the time PA problem since the use of Rosen's Theorem [23] is not required. Without loss of generality assume that $S = 1$. Whatever the strategy of user 2, user 1 sees no interference. Therefore he can choose $\mathbf{Q}_1^{(1)}$ independently of user 2. Because $R_1^{(1)}(\mathbf{Q}_1^{(1)}, \mathbf{Q}_2^{(1)})$ is a strictly concave function to be maximized over a convex set, there is a unique optimum strategy for user 1. As we assume a game with complete information and rational users, user 2 knows the utility of user 1 and thus the precoding matrix he will choose. The same concavity argument can be used for $R_2^{(1)}(\mathbf{Q}_1^{(1)}, \mathbf{Q}_2^{(1)})$ and therefore guarantees that user 2 employs a unique precoding matrix. ■

Determination of the NE. In order to find the optimum covariance matrices, we proceed in the same way as described in [10]. First we focus on the optimum eigenvectors and then we determine the optimum eigenvalues by approximating the utility functions under the large system assumption. In order to determine the optimum eigenvectors, the proof in [20] can be applied in our context to assert that there is no loss of optimality by restricting the search for the optimum covariance matrix when imposing the structure $\mathbf{Q}_k^{(s)} = \mathbf{U}_k \mathbf{P}_k^{(s)} \mathbf{U}_k^H$, where $\mathbf{U}_k$ is a unitary matrix coming from the spectral decomposition of transmit correlation matrix $\mathbf{T}_k = \mathbf{U}_k \mathbf{D}_k \mathbf{U}_k^H$ defined in Eq. (2) and the diagonal matrix $\mathbf{P}_k^{(s)} = \text{Diag}(P_k^{(s)}(1), ..., P_k^{(s)}(n_t))$ represents the powers user k allocates to the different eigenvectors. As a consequence, we can exploit once again the results of [24], [25], [26] assuming the asymptotic regime in terms of the number of antennas. The new approximated rates are:

$$\begin{aligned}
\tilde{R}_1^{(1)}(\mathbf{P}_1^{(1)}) &= \sum_{i=1}^{n_t} \log_2 \left[1 + \eta P_1^{(1)}(i) d_1^{(\mathrm{T})}(i) \gamma_1\right] + \sum_{j=1}^{n_r} \log_2 \left[1 + \eta d^{(\mathrm{R})}(j) \delta_1\right] - n_t \eta \gamma_1 \delta_1 \log_2 e \\
\tilde{R}_2^{(1)}(\mathbf{P}_1^{(1)}, \mathbf{P}_2^{(1)}) &= \sum_{\ell=1}^{2} \sum_{i=1}^{n_t} \log_2 \left[1 + 2\eta P_\ell^{(1)}(i) d_\ell^{(\mathrm{T})}(i) \gamma_2\right] + \sum_{j=1}^{n_r} \log_2 \left[1 + 2\eta d^{(\mathrm{R})}(j) \delta_2\right] - 4 n_t \eta \gamma_2 \delta_2 \log_2 e - \\
&\quad - \tilde{R}_1^{(1)}(\mathbf{P}_1^{(1)})
\end{aligned} \tag{15}$$

where $\forall k \in \{1, 2\}, d_k^{(\mathrm{T})}(i)$, $i \in \{1, ..., n_t\}$ are always the eigenvalues of the transmit correlation matrices $\mathbf{T}_k$, $d^{(\mathrm{R})}(j)$, $j \in \{1, ..., n_r\}$, are the eigenvalues of the receive correlation matrix $\mathbf{R}$ and the parameters γ_i, δ_j are the unique solutions of the following systems of equations:

$$\begin{cases}
\gamma_1 &= \dfrac{1}{n_t} \displaystyle\sum_{j=1}^{n_r} \dfrac{d^{(\mathrm{R})}(j)}{1 + \eta d^{(\mathrm{R})}(j) \delta_1} \\
\delta_1 &= \dfrac{1}{n_t} \displaystyle\sum_{i=1}^{n_t} \dfrac{P_1^{(1)}(i) d_1^{(\mathrm{T})}(i)}{1 + \eta P_1^{(1)}(i) d_1^{(\mathrm{T})}(i) \gamma_1}
\end{cases} \tag{16}$$

$$\left\{\begin{array}{rcl}\gamma_2 & = & \dfrac{1}{2n_t}\displaystyle\sum_{j=1}^{n_r}\frac{d^{(\mathrm{R})}(j)}{1+2\eta d^{(\mathrm{R})}(j)\delta_2}\\ \delta_2 & = & \dfrac{1}{2n_t}\displaystyle\sum_{l=1}^{2}\sum_{i=1}^{n_t}\frac{P_l^{(1)}(i)d_l^{(\mathrm{T})}(i)}{1+2\eta P_l^{(1)}(i)d_l^{(\mathrm{T})}(i)1\gamma_2}.\end{array}\right. \tag{17}$$

Then, optimizing the approximated rates $\tilde{R}_k^{(1)}(\cdot)$ w.r.t. $P_k^{(1)}(i)$ leads to the following water-filling equations:

$$\forall k\in\{1,2\},\ P_k^{(1),\mathrm{NE}}(i)=\left[\frac{1}{\ln 2\lambda_k^{(1)}}-\frac{1}{\eta d_k^{(\mathrm{T})}(i)\gamma_k}\right]^+ \tag{18}$$

where $\lambda_k^{(1)}\geq 0$, $k\in\{1,2\}$, are the Lagrangian multipliers tuned in order to meet the power constraints given in (6): $\sum_{i=1}^{n_t}P_k^{(1),\mathrm{NE}}(i)=n_tP_k$. We use the same iterative PA algorithm as the one described in [10]. Under the large systems assumption, in this game also, the approximated utilities have the same properties as the exact utilities.

Sum-rate efficiency of the NE. Unlike the temporal PA game, we have not assumed a particular structure for the precoding matrices and thus the centralized solution coincides with the sum-capacity of the virtual MIMO network, $R_{\mathrm{sum}}^{(\mathrm{C})}=C_{\mathrm{sum}}$. Another important point to notice here is that the equilibrium precoding matrices do not depend on p. This considerably simplifies the BS's choice for the sum-rate optimal value for p. Indeed, as we have already mentioned, the precoding matrices do no depend on p and therefore the sum-rate $R_{\mathrm{sum}}(p)$ is merely a linear function of p: $R_{\mathrm{sum}}^{\mathrm{NE}}(p)=ap+b$ where

$$\begin{array}{rcl}a & = & \mathbb{E}\log|\mathbf{I}+\eta\mathbf{H}_1\mathbf{Q}_1^{(1),\mathrm{NE}}\mathbf{H}_1^H+\eta\mathbf{H}_2\mathbf{Q}_2^{(1),\mathrm{NE}}\mathbf{H}_2^H|-\mathbb{E}\log|\mathbf{I}+\eta\mathbf{H}_1\mathbf{Q}_1^{(2),\mathrm{NE}}\mathbf{H}_1^H+\eta\mathbf{H}_2\mathbf{Q}_2^{(2),\mathrm{NE}}\mathbf{H}_2^H|\\ b & = & \mathbb{E}\log|\mathbf{I}+\eta\mathbf{H}_1\mathbf{Q}_1^{(2),\mathrm{NE}}\mathbf{H}_1^H+\eta\mathbf{H}_2\mathbf{Q}_2^{(2),\mathrm{NE}}\mathbf{H}_2^H|.\end{array} \tag{19}$$

Depending on the sign of a, if the BS wants to maximize the sum-rate, it will choose either $p=0$ or $p=1$. If it wants a fair game it will choose $p=\frac{1}{2}$ and accept a certain loss of global optimality. Note that even for $p\in\{0,1\}$ the sum-capacity is not reached in general: this is because the matrix $\mathbf{Q}_1^{(1),\mathrm{NE}}$ (resp. $\mathbf{Q}_2^{(2),\mathrm{NE}}$) does not coincide with the first (resp. second) component of the pair of precoding matrices that maximizes the (strictly concave) network sum-rate. However, as we did for the temporal PA game, in the low and high SNR regimes one can show that the decentralized MIMO MAC has the same performance (w.r.t. the sum-rate) as its equivalent $Kn_t\times n_r$ virtual MIMO network.

V. SIMULATION EXAMPLES

All the results will be provided by assuming the asymptotic regime in the numbers of antennas. We know, from many contributions (see e.g., [10], [27], [29], [30]) that large-system approximates of ergodic rates are accurate even for relatively small systems. We also assume that $\mathbf{R}=\mathbf{I}$.

For the TPA problem, we look at the case where there is no transmit correlation, $\mathbf{T}_k = \mathbf{I}$. We have seen that the performance of decentralized MAC depends on the rule of the game i.e., the value of p. This is exactly what Fig. 1 depicts for the following scenario: $P_1 = 1$, $P_2 = 10$, $\eta = 5$ dB, $n_t = n_r = 4$. First, we see that the MAC sum-rate is a convex function of p and the maximum of $R_{\text{sum}}^{\text{NE}}(p)$ is reached for $p \in \{0, 1\}$. In these points, which correspond to the most unfair decoding schemes (either user 1 or 2 is always decoded first) the centralized sum-capacity of the MAC is achieved. One important observation to be made is that the minimum and maximum only differ by about 1%. Many other simulations have confirmed this observation. This shows that whatever the value of p, the gap between the sum-rate of a decentralized MIMO MAC with selfish users and the sum-capacity of the equivalent cooperative MAC (virtual MIMO network) is in fact very small. Now, we want to evaluate the benefits brought by using a SIC instead of single-user decoding [10]. For the scenario where $P_1 = P$, $P_2 = 10P$ with $P \in [0, 20]$, $n_r = n_t = 4$ and $\eta = 5$ dB, Fig. 2 shows the achievable network sum-rate at the NE versus the available power at the first transmitter P. For the SUD scheme, the users are decoded simultaneously at the receiver. In this case both users see all the interference coming from the others. We see that the SIC scheme performs much better than the proposed SUD scheme, regardless of the distribution of the coordination signal: this comparison makes sense especially for the point $p = \frac{1}{2}$ since both decoding schemes are fair.

From now on, we consider the SPA problem. In this case we assume an exponential correlation profile for $\mathbf{T}_k$ such that $\mathbf{T}_k(i, j) = t_k^{|i-j|}$ (note that $\text{Tr}(\mathbf{T}_k) = n_t$), where $0 \leq t_k \leq 1$ is the corresponding correlation coefficient [31], [32]. We already know that the sum-rate is a linear function of p and therefore is maximized when either $p = 0$ or $p = 1$. It turns out that this slope has a small value. Furthermore, it has been observed to be even 0 for a symmetric MAC, i.e., $P_1 = P_2$ and $t_1 = t_2$. These observations have been confirmed by many simulations. In Fig. 3 we have plotted the sum-rate achieved by varying p for the scenario: $P_1 = 5$, $P_2 = 50$, $\eta = 3$ dB, $n_t = n_r = 4$, $t_1 = 0.4$, $t_2 = 0.3$. Even in this scenario, which was thought to be a bad case in terms of sub-optimality, the sum-rate is not far from the sum-capacity of the centralized MAC. For the same scenario, we have plotted in Fig. 4 the achievable rate region and compared it to that obtained with SUD. We observe that in large MIMO MAC channels, the capacity region comprises a full cooperation segment (approximately) just like SISO MAC channels. The coordination signal allows one to move along an almost straight line, corresponding to a relatively large range of rates.

VI. CONCLUSION

We have provided complete proofs for the existence and uniqueness of an NE in fast fading MIMO MACs with CSIR and CDIT where the transmission rate is chosen as user utility. By exploiting random matrix theory, we have also provided the corresponding optimum selfish PA policies. We have seen that the BS can, through a single parameter (i.e., $p \in [0,1]$, which represents the distribution of the coordination signal), force the system to operate at many different points that correspond to a relatively large range of achievable transmission rate pairs. We know, from [1], [9] that for Gaussian MACs with single antenna terminals, this set of rate pairs corresponds to the full cooperation segment of the centralized MAC. Said otherwise a *decentralized* Gaussian SISO MAC with coordination achieves the same rate pairs as a MAC with full cooperation or *virtual MIMO system*. The goal here was to know to what extent this key result is valid for fading MAC with multi-antenna terminals. It turns out this is almost true in the MIMO setting. In the cases of interest, where the power is optimally allocated either over space or time, the performance gap is relatively small even though the proposed coordination mechanism was a priori sub-optimal since it does take into account the channel realizations (known to the receiver). Interestingly in large MIMO MACs, the capacity region comprises a full cooperation segment just like SISO MACs. The coordination signal precisely allows one to move along the corresponding (almost) straight line. This shows the relevance of large systems in decentralized networks since they allow to determine the capacity region of certain systems whereas it is unknown in the finite setting. Furthermore, they induce an averaging effect, which makes the users' behavior predictable. Indeed, in large MIMO MACs the knowledge of the CSIT does not improve the performance w.r.t. the case with CDIT. To conclude we review some extensions of this work which we have suggested throughout it. It would be interesting to study the case of the decentralized space-time PA, which, in particular, would require the generalization of Lemma 1 to arbitrary positive matrices and exploitation of some results in [21]. A second useful extension would be to evaluate the impact of a non-perfect SIC on the PA problem. At last, we will mention that it would be useful to evaluate analytically or bounding the price of anarchy of the NE, which would require to find a bounding technique different from that used for non-atomic games [34], [35], [36].

APPENDIX A

We want to prove that the diagonally strict concavity condition is met for the time PA problem i.e., for all $(\alpha_1',\alpha_1'') \in \left(\mathcal{A}_1^{\mathrm{TPA}}\right)^2$ and $(\alpha_2',\alpha_2'') \in \left(\mathcal{A}_2^{\mathrm{TPA}}\right)^2$ such that either $\alpha_1' \neq \alpha_1''$ or $\alpha_2' \neq \alpha_2''$ we want to prove that:

$$\mathcal{C} = (\alpha_1'' - \alpha_1') \left[\frac{\partial R_1}{\partial \alpha_1}(\alpha_1', \alpha_2') - \frac{\partial R_1}{\partial \alpha_1}(\alpha_1'', \alpha_2'') \right] + (\alpha_2'' - \alpha_2') \left[\frac{\partial R_2}{\partial \alpha_2}(\alpha_1', \alpha_2') - \frac{\partial R_2}{\partial \alpha_2}(\alpha_1'', \alpha_2'') \right] > 0. \qquad (20)$$

We can write $\mathcal{C} = p\mathcal{T}^{(1)} + \overline{p}\mathcal{T}^{(2)}$ where for all $s \in \{1,2\}$:

$$\mathcal{T}^{(s)} = (\alpha_1'' - \alpha_1') \left[\frac{\partial R_1^{(s)}}{\partial \alpha_1}(\alpha_1', \alpha_2') - \frac{\partial R_1^{(s)}}{\partial \alpha_1}(\alpha_1'', \alpha_2'') \right] + (\alpha_2'' - \alpha_2') \left[\frac{\partial R_2^{(s)}}{\partial \alpha_2}(\alpha_1', \alpha_2') - \frac{\partial R_2^{(s)}}{\partial \alpha_2}(\alpha_1'', \alpha_2'') \right] \quad (21)$$

By expanding $\mathcal{T}^{(1)}$ we have

$$\begin{aligned} \mathcal{T}^{(1)} &= (\alpha_1'' - \alpha_1')\mathbb{E}\mathrm{Tr}\left\{[(\mathbf{I} + \rho_1\alpha_1'\mathbf{H}_1\mathbf{H}_1^H)^{-1} - (\mathbf{I} + \rho_1\alpha_1'\mathbf{H}_1\mathbf{H}_1^H)^{-1}]\rho_1\mathbf{H}_1\mathbf{H}_1^H\right\} \\ &\quad + (\alpha_2'' - \alpha_2')\mathbb{E}\mathrm{Tr}\left\{(\mathbf{I} + \rho_1\alpha_1'\mathbf{H}_1\mathbf{H}_1^H + \tfrac{1-\overline{p}\alpha_2'}{p}\rho_2\mathbf{H}_2\mathbf{H}_2^H)^{-1}\rho_2\tfrac{\overline{p}}{p}\mathbf{H}_2\mathbf{H}_2^H \right. \\ &\quad \left. - (\mathbf{I} + \rho_1\alpha_1''\mathbf{H}_1\mathbf{H}_1^H + \tfrac{1-\overline{p}\alpha_2''}{p}\rho_2\mathbf{H}_2\mathbf{H}_2^H)^{-1}\rho_2\tfrac{\overline{p}}{p}\mathbf{H}_2\mathbf{H}_2^H\right\} \\ &= \mathbb{E}\mathrm{Tr}\left\{(\mathbf{A}'' - \mathbf{A}')[(\mathbf{I} + \mathbf{A}')^{-1} - (\mathbf{I} + \mathbf{A}'')^{-1}] + (\mathbf{B}'' - \mathbf{B}')[(\mathbf{I} + \mathbf{B}' + \mathbf{A}')^{-1} - (\mathbf{I} + \mathbf{B}'' + \mathbf{A}'')^{-1}]\right\}, \end{aligned} \quad (22)$$

where $\mathbf{A}' = \rho_1\alpha_1'\mathbf{H}_1\mathbf{H}_1^H$, $\mathbf{A}'' = \rho_1\alpha_1''\mathbf{H}_1\mathbf{H}_1^H$, $\mathbf{B}' = \rho_2\frac{1-\overline{p}\alpha_2'}{p}\mathbf{H}_2\mathbf{H}_2^H$, $\mathbf{B}'' = \rho_2\frac{1-\overline{p}\alpha_2''}{p}\mathbf{H}_2\mathbf{H}_2^H$. We observe that the matrices $\mathbf{A}'$, $\mathbf{A}''$, $\mathbf{B}'$ and $\mathbf{B}''$ verify the assumptions of Lemma 1. First, they are Hermitian and non-negative. Second, as they write as $\mathbf{A}' = a'\mathbf{H}_1\mathbf{H}_1^H$, $\mathbf{A}'' = a''\mathbf{H}_1\mathbf{H}_1^H$, $\mathbf{B}' = b'\mathbf{H}_2\mathbf{H}_2^H$ and $\mathbf{B}'' = b''\mathbf{H}_2\mathbf{H}_2^H$, we also see that the matrix order $\succeq$ is total for each of the pairs of matrices $(\mathbf{A}', \mathbf{A}'')$ and $(\mathbf{B}', \mathbf{B}'')$. This directly follows from the fact that the scalar order $\geq$ is total, which implies that either $a'' \geq a'$ or $a'' \leq a'$ and either $b'' \geq b'$ or $b'' \leq b'$. By considering the particular structure of the four matrices and applying Lemma 1, it is straightforward to see that the term $\mathcal{T}^{(1)}$ is strictly positive, $\mathcal{T}^{(1)} > 0$. In a similar way we can prove that $\mathcal{T}^{(2)} > 0$ and thus the diagonally strict concavity condition is met: $\mathcal{C} > 0$.

APPENDIX B

Proving Lemma 1 amounts to showing that

$$\mathcal{T} = \mathrm{Tr}\left\{(\mathbf{A} - \mathbf{B})(\mathbf{B}^{-1} - \mathbf{A}^{-1}) + (\mathbf{C} - \mathbf{D})[(\mathbf{B} + \mathbf{D})^{-1} - (\mathbf{A} + \mathbf{C})^{-1}]\right\} > 0 \quad (23)$$

where the matrices $\mathbf{A} = \mathbf{I} + \mathbf{A}''$, $\mathbf{B} = \mathbf{I} + \mathbf{A}'$, $\mathbf{C} = \mathbf{B}''$ and $\mathbf{D} = \mathbf{B}'$ have been introduced for more clarity. Since the matrix order $\succeq$ is total for $\mathbf{A}$ and $\mathbf{B}$, and $\mathbf{C}$ and $\mathbf{D}$ it suffices to prove that $\mathcal{T} > 0$ for the four following cases: (1) $\mathbf{A} \succeq \mathbf{B}$ and $\mathbf{C} \succeq \mathbf{D}$; (2) $\mathbf{A} \preceq \mathbf{B}$ and $\mathbf{C} \preceq \mathbf{D}$; (3) $\mathbf{A} \succeq \mathbf{B}$ and $\mathbf{C} \preceq \mathbf{D}$; (4) $\mathbf{A} \preceq \mathbf{B}$ and $\mathbf{C} \succeq \mathbf{D}$.

Case (1): $\mathbf{A} \succeq \mathbf{B}$ and $\mathbf{C} \succeq \mathbf{D}$. To prove the desired result in this case we use the following lemma.

Lemma 2: If $\mathbf{M}$ is a Hermitian and non-negative ($\mathbf{M} = \mathbf{M}^H \succeq \mathbf{0}$) and $\mathbf{N}$ is non-negative ($\mathbf{N} \succeq 0$) but not necessarily Hermitian, then $\mathrm{Tr}(\mathbf{MN}) \geq 0$.

Proof: We write $\mathrm{Tr}(\mathbf{MN}) = \mathrm{Tr}(\mathbf{M}^{1/2}\mathbf{N}\mathbf{M}^{1/2}) \geq 0$. We have used the fact that $\mathbf{M}$ is a Hermitian non-negative matrix to write $\mathbf{M} = \mathbf{M}^{1/2}\mathbf{M}^{1/2}$. Knowing that $\mathbf{N}$ is a non-negative matrix one can easily check

that $\mathbf{M}^{1/2}\mathbf{N}\mathbf{M}^{1/2}$ is also a non-negative matrix and thus the trace (sum of the non-negative eigenvalues) is non-negative. ■

The quantity $\mathcal{T}$ writes as $\mathcal{T} = \mathrm{Tr}(\mathbf{M}_1\mathbf{N}_1) + \mathrm{Tr}(\mathbf{M}_2\mathbf{N}_2)$ where $\mathbf{M}_1 = \mathbf{A}-\mathbf{B}$, $\mathbf{N}_1 = \mathbf{B}^{-1}-\mathbf{A}^{-1}$, $\mathbf{M}_2 = \mathbf{C}-\mathbf{D}$ and $\mathbf{N}_2 = (\mathbf{B}+\mathbf{D})^{-1} - (\mathbf{A}+\mathbf{C})^{-1}$. Clearly these four matrices are Hermitian. Since by assumption $\mathbf{M}_1 \succeq 0$ and $\mathbf{M}_2 \succeq 0$ we only need to verify that $\mathbf{N}_1 \succeq 0$ and $\mathbf{N}_2 \succeq 0$ to be able to apply Lemma 2 to $\mathcal{T}$. The matrix $\mathbf{N}_1$ is non-negative because for any pair of invertible matrices $(\mathbf{X}, \mathbf{Y})$: $\mathbf{X} \succeq \mathbf{Y} \Leftrightarrow \mathbf{Y}^{-1} \succeq \mathbf{X}^{-1}$ (see e.g., [33]). The same result applies to $\mathbf{N}_2$ since by assumption $\mathbf{A}+\mathbf{C} \succeq \mathbf{B}+\mathbf{D}$. Using lemma 2 concludes the proof.

Case (3): $\mathbf{A} \succeq \mathbf{B}$ and $\mathbf{C} \preceq \mathbf{D}$. To treat this case we first prove the following auxiliary Lemma.

Lemma 3: Let $\mathbf{X}$ and $\mathbf{Y}$ be two distinct, Hermitian and positive matrices of size n: $\mathbf{X} = \mathbf{X}^H \succ 0$, $\mathbf{Y} = \mathbf{Y}^H \succ 0$ and $\mathbf{X} \neq \mathbf{Y}$. Then $\mathrm{Tr}[(\mathbf{X}-\mathbf{Y})(\mathbf{Y}^{-1}-\mathbf{X}^{-1})] \geq 0$.

Proof: It is easy to see that $\mathrm{Tr}[(\mathbf{X}-\mathbf{Y})(\mathbf{Y}^{-1}-\mathbf{X}^{-1})] = \mathrm{Tr}[\mathbf{Z}+\mathbf{Z}^{-1}-2\mathbf{I}]$, with the Hermitian and positive matrix $\mathbf{Z} \triangleq \mathbf{X}^{\frac{1}{2}}\mathbf{Y}^{-1}\mathbf{X}^{\frac{1}{2}}$ and thus we further have $\mathrm{Tr}[(\mathbf{X}-\mathbf{Y})(\mathbf{Y}^{-1}-\mathbf{X}^{-1})] = \sum_{i=1}^{n} \frac{(\lambda_Z(i)-1)^2}{\lambda_Z(i)} \geq 0$ where the matrix $\mathbf{\Lambda}_Z = \mathrm{Diag}(\lambda_Z(1), ..., \lambda_Z(n))$ corresponds to the spectral decomposition of . ■

By applying this lemma to $\mathcal{T}$ we have that:

$$\begin{aligned} \mathcal{T} &= \mathrm{Tr}\left\{(\mathbf{A}-\mathbf{B})(\mathbf{B}^{-1}-\mathbf{A}^{-1}) + [(\mathbf{C}+\mathbf{A})-(\mathbf{B}+\mathbf{D})][(\mathbf{B}+\mathbf{D})^{-1}-(\mathbf{A}+\mathbf{C})^{-1}]\right. \\ &\quad \left. -(\mathbf{A}-\mathbf{B})[(\mathbf{B}+\mathbf{D})^{-1}-(\mathbf{A}+\mathbf{C})^{-1}]\right\} \\ &\geq \mathrm{Tr}\left\{(\mathbf{A}-\mathbf{B})(\mathbf{B}^{-1}-\mathbf{A}^{-1}) - (\mathbf{A}-\mathbf{B})[(\mathbf{B}+\mathbf{D})^{-1}-(\mathbf{A}+\mathbf{C})^{-1}]\right\}. \end{aligned} \tag{24}$$

We know that $\mathbf{C} \preceq \mathbf{D}$ then $\mathbf{C}+\mathbf{A} \preceq \mathbf{D}+\mathbf{A}$ and also that $(\mathbf{C}+\mathbf{A})^{-1} \succeq (\mathbf{D}+\mathbf{A})^{-1}$. Using the fact that $\mathbf{A} \succeq \mathbf{B}$ and also Lemma 2 we have that $\mathrm{Tr}[(\mathbf{A}-\mathbf{B})(\mathbf{C}+\mathbf{A})^{-1}] \geq \mathrm{Tr}[(\mathbf{A}-\mathbf{B})(\mathbf{D}+\mathbf{A})^{-1}]$ and the trace becomes lower bounded as $\mathcal{T} \geq \mathrm{Tr}\left\{(\mathbf{A}-\mathbf{B})(\mathbf{B}^{-1}-\mathbf{A}^{-1}) - (\mathbf{A}-\mathbf{B})[(\mathbf{B}+\mathbf{D})^{-1}-(\mathbf{A}+\mathbf{D})^{-1}]\right\}$. Now, we are going to prove that this lower bound, say $\mathcal{T}_{LB}$, is positive:

$$\begin{aligned} \mathcal{T}_{LB} &= \mathrm{Tr}\left\{(\mathbf{A}-\mathbf{B})(\mathbf{B}^{-1}-\mathbf{A}^{-1}) - [(\mathbf{A}+\mathbf{D})-(\mathbf{B}+\mathbf{D})]\left[(\mathbf{B}+\mathbf{D})^{-1}-(\mathbf{A}+\mathbf{D})^{-1}\right]\right\} \\ &= \mathrm{Tr}\left\{(\widetilde{\mathbf{A}}-\widetilde{\mathbf{B}})(\widetilde{\mathbf{B}}^{-1}-\widetilde{\mathbf{A}}^{-1}) - \left[(\widetilde{\mathbf{A}}+\mathbf{I})-(\widetilde{\mathbf{B}}+\mathbf{I})\right]\left[(\widetilde{\mathbf{B}}+\mathbf{I})^{-1}-(\widetilde{\mathbf{A}}+\mathbf{I})^{-1}\right]\right\} \end{aligned} \tag{25}$$

where we have made the following change of variables: $\mathbf{A} = \mathbf{D}^{1/2}\widetilde{\mathbf{A}}\mathbf{D}^{1/2}$, $\mathbf{B} = \mathbf{D}^{1/2}\widetilde{\mathbf{B}}\mathbf{D}^{1/2}$ such that $\widetilde{\mathbf{A}} = \mathbf{D}^{-1/2}\mathbf{A}\mathbf{D}^{-1/2} = \widetilde{\mathbf{A}}^H \succ 0$ and $\widetilde{\mathbf{B}} = \mathbf{D}^{-1/2}\mathbf{B}\mathbf{D}^{-1/2} = \widetilde{\mathbf{B}}^H \succ 0$. By applying the Woodbury formula $(\widetilde{\mathbf{A}}+\mathbf{I})^{-1} = \widetilde{\mathbf{A}}^{-1} - \widetilde{\mathbf{A}}^{-1}(\widetilde{\mathbf{A}}+\mathbf{I})^{-1}$ and $(\widetilde{\mathbf{B}}+\mathbf{I})^{-1} = \widetilde{\mathbf{B}}^{-1} - \widetilde{\mathbf{B}}^{-1}(\widetilde{\mathbf{B}}+\mathbf{I})^{-1}$, the lower bound $\mathcal{T}_{LB}^{(1)}$ rewrites as:

$$\begin{aligned} \mathcal{T}_{LB} &= \mathrm{Tr}\left\{(\widetilde{\mathbf{A}}-\widetilde{\mathbf{B}})\left[\widetilde{\mathbf{B}}^{-1}-\widetilde{\mathbf{A}}^{-1}-\widetilde{\mathbf{B}}^{-1}+\widetilde{\mathbf{B}}^{-1}(\widetilde{\mathbf{B}}+\mathbf{I})^{-1}+\widetilde{\mathbf{A}}^{-1}-\widetilde{\mathbf{A}}^{-1}(\widetilde{\mathbf{A}}+\mathbf{I})^{-1}\right]\right\} \\ &= \mathrm{Tr}\left\{\widetilde{\mathbf{A}}\widetilde{\mathbf{B}}^{-1}(\widetilde{\mathbf{B}}+\mathbf{I})^{-1} + \widetilde{\mathbf{B}}\widetilde{\mathbf{A}}^{-1}(\widetilde{\mathbf{A}}+\mathbf{I})^{-1} - (\widetilde{\mathbf{A}}+\mathbf{I})^{-1} - (\widetilde{\mathbf{B}}+\mathbf{I})^{-1}\right\}. \end{aligned} \tag{26}$$

Let us denote the ordered eigenvalues of the two matrices $\widetilde{\mathbf{A}}$ and $\widetilde{\mathbf{B}}$ as $\lambda_{\widetilde{A}}(1) \leq \lambda_{\widetilde{A}}(2) \leq \ldots \leq \lambda_{\widetilde{A}}(n)$ and $\lambda_{\widetilde{B}}(1) \leq \lambda_{\widetilde{B}}(2) \leq \ldots \leq \lambda_{\widetilde{B}}(n)$. From [37] we know that for two matrices $\mathbf{X}$ and $\mathbf{Y}$ of size n, $\mathrm{Tr}(\mathbf{X}\mathbf{Y}) \geq$

$\sum_{i=1}^{n} \lambda_X(i)\lambda_Y(n-i+1)$, which implies directly that $\mathrm{Tr}(\mathbf{X}\mathbf{Y}^{-1}) \geq \sum_{i=1}^{n} \frac{\lambda_X(i)}{\lambda_Y(i)}$, where $\lambda_X(i)$ and $\lambda_Y(i)$ are the ordered eigenvalues (in the previously specified order) of the corresponding matrices. Applying this result we find that

$$\mathrm{Tr}\left[\widetilde{\mathbf{A}}\widetilde{\mathbf{B}}^{-1}(\mathbf{I}+\widetilde{\mathbf{B}})^{-1}\right] \geq \sum_{i=1}^{n} \frac{\lambda_{\widetilde{A}}(i)}{\lambda_{\widetilde{B}}(i)(1+\lambda_{\widetilde{B}}(i))}, \quad \mathrm{Tr}\left[\widetilde{\mathbf{B}}\widetilde{\mathbf{A}}^{-1}(\mathbf{I}+\widetilde{\mathbf{A}})^{-1}\right] \geq \sum_{i=1}^{n} \frac{\lambda_{\widetilde{B}}(i)}{\lambda_{\widetilde{A}}(i)(1+\lambda_{\widetilde{A}}(i))}, \tag{27}$$

and finally obtain that:

$$\mathcal{T}_{LB} \geq \sum_{i=1}^{n} \frac{[\lambda_{\widetilde{A}}(i)-\lambda_{\widetilde{B}}(i)]^2[1+\lambda_{\widetilde{A}}(i)+\lambda_{\widetilde{B}}(i)]}{\lambda_{\widetilde{A}}(i)\lambda_{\widetilde{B}}(i)[1+\lambda_{\widetilde{A}}(i)][1+\lambda_{\widetilde{B}}(i)]} \geq 0. \tag{28}$$

To conclude the global proof one can easily check that Case (2) (resp. Case (4)) can be readily proved from the proof of Case (1) (resp. Case (3)) by interchanging the role of $\mathbf{A}$ and $\mathbf{B}$ and $\mathbf{C}$ and $\mathbf{D}$.

REFERENCES

[1] E. V. Belmega, S. Lasaulce and M. Debbah, "Power control in distributed multiple access channels with coordination", *Proc. of the Intl. Symp. on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks and Workshops (WiOpt), Berlin, Germany*, pp. 1–8, Apr. 2008.

[2] S. A. Grandhi, R. Vijayan and D. J. Goodman, "Distributed algorithm for power control in cellular radio systems", *Proc. of Annual Allerton Conf. on Comm. Control and Computing, Monticello, Illinois, USA*, Sep./Oct. 1992.

[3] S. A. Grandhi, R. Vijayan and D. J. Goodman, "Distributed power control in cellular radio systems", *IEEE Trans. on Comm.*, Vol. 42, No. 234, pp. 226–228, Feb./Mar./Apr. 1994.

[4] H. Ji and C.-Y. Huang, "Non-cooperative uplink power control in cellular radio systems", *Wireless Networks*, Vol. 4, Issue. 3, pp. 233–240, Mar. 1998.

[5] S.-J. Oh, T. L. Olsen and K. M. Wasserman, "Distributed power control and spreading gain allocation in CDMA data networks", *IEEE Proc. of Conf. of Computer and Comm. Societies (INFOCOM), Tel Aviv, Israel*, Vol. 2, pp. 379–385, Mar. 2000.

[6] D. J. Goodman and N. B. Mandayam, "Power control for wireless data", *IEEE Person. Comm.*, Vol. 7, No. 2, pp. 48–54, 2000.

[7] F. Meshkati, M. Chiang, H. V. Poor and S. C. Schwartz, "A game-theoretic approach to energy-efficient power control in multi-carrier CDMA systems", *IEEE J. on Selected Areas in Comm.*, Vol. 24, No. 6, pp. 1115–1129, Jun. 2006.

[8] W. Yu, G. Ginis and J. M. Cioffi, "Distributed multiuser power control for digital subscriber lines", *IEEE J. of Selected Areas in Comm.*, Vol. 20, No. 5, pp. 1105–1115, Jun. 2002.

[9] L. Lai and H. El Gamal, "The water-filling game in fading multiple-access channels", *IEEE Trans. on Inform. Theory*, Vol. 54, No. 5, pp. 2110–2122, May 2008.

[10] S. Lasaulce, A. Suarez, M. Debbah and L. Cottatellucci, "Power allocation game for fading MIMO multiple access channels with antenna correlation", *ACM Proc. of the Intl Conf. on Game Theory in Comm. Networks (Gamecomm), Nantes, France*, pp. 1–9, Oct. 2007.

[11] G. Arslan, M. F. Demirkol and Y. Song, "Equilibrium efficiency improvement in MIMO interference systems: a Decentralized Stream Control Approach", *IEEE Trans. on Wireless Comm.*, Vol. 6, No. 8, pp. 2984–2993, Aug. 2007.

[12] G. Scutari, D. P. Palomar and S. Barbarossa, "Competitive design of multiuser MIMO systems based on game theory: a unified view", *IEEE J. of Selected Areas in Comm.*, Vol. 26, No. 7, pp. 1089–1103, Sep. 2008.

[13] A. Carleial, "Interference channels", *IEEE Trans. on Inform. Theory*, Vol. 24, No.1, pp. 60–70, 1978.

[14] D. P. Palomar, J. M. Cioffi, M. A. Lagunas, "Uniform power allocation in MIMO channels: a game theoretic approach", *IEEE Trans. on Inform. Theory*, Vol. 49, No. 7, pp. 1707–1727, Jul. 2003.

[15] A. D. Wyner, "Recent results in Shannon theory", *IEEE Trans. on Inform. Theory*, Vol. 20, Issue 1, pp. 2–10, Jan. 1974.

[16] T. Cover, "Some advances in broadcast channels", *in Advances in Comm. Systems*, Vol. 4, Academic Press, 1975.

[17] E. Telatar, "Capacity of multi-antenna Gaussian channels", *European Trans. on Telecomm.*, Vol. 10, No. 6, pp. 585–596, 1999.

[18] 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, pp. 502–513, Mar. 2000.

[19] E. Riegler and G. Taricco, "On the ergodic capacity region of the separately correlated Rician fading multiple access MIMO channel", *IEEE Global Telecomm. Conf. (Globecom), Nantes, France*, pp. 1535–1539, Nov. 2007.

[20] A. Soysal and S. Ulukus, "Optimality of beamforming in fading MIMO multiple access channels", *IEEE Trans. on Comm.*, in press.

[21] V. L. Girko, "Theory of stochastic canonical equations", *Kluwer Academic Publishers*, Vol. I, The Netherlands, 2001.

[22] D. Tse and P. Viswanath, "Fundamentals of wireless communication", *Cambridge University Press*, 2005.

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

[24] A. Tulino and S. Verdu, "Random matrices and wireless communications", *Foundations and trends in comm. and inform. theory, NOW, The Essence of Knowledge*, 2004.

[25] J. W. Sylverstein and Z. D. Bai, "On the empirical distribution of eigenvalues of a class of large dimensional random matrices", *J. of Multivariate Analysis*, Vol. 54, No. 2, pp. 175–192, 1995.

[26] A. Tulino and S. Verdu, "Impact of antenna correlation on the capacity of multi-antenna channels", *IEEE Trans. on Inform. Theory*, Vol. 51, No. 7, pp. 2491–2509, Jul. 2005.

[27] J. Dumont, W. Hachem, S. Lasaulce, P. Loubaton and J. Najim, "On the capacity achieving covariance matrix of Rician MIMO channels: an asymptotic approach", *IEEE Trans. on Inform. Theory*, revised, http://arxiv.org/abs/0710.4051, Oct. 2007.

[28] C. H. Papadimitriou, " Algorithms, games, and the Internet", *ACM Proc. on Annual Symposium on Theory of Computing, Crete, Grece*, Jul. 2001.

[29] E. Biglieri, G. Taricco and A. Tulino, "How far is infinity? Using asymptotic analyses in multiple-antennas systems", *Proc. of the Intl. Symposium on Software Testing and Analysis (ISSTA), Rome, Italy*, Vol. 1, pp. 1–6, Jul. 2002.

[30] J. Dumont, P. Loubaton and S. Lasaulce, "On the capacity achieving transmit covariance matrices of MIMO correlated Rician channels: a large system approach", *IEEE Proc. of Globecom Technical Conf., San Francisco, California, USA*, Nov./Dec. 2006.

[31] M. Chiani, M. Z. Win, A. Zanella, "On the capacity of spatially correlated MIMO Rayleiggh-Fading Channels", *IEEE Trans. on Inform. Theory*, Vol. 49, No. 10, pp. 2363–2371, Jul. 2003.

[32] A. Skupch, D. Seethaler, F. Hlawatsch, "Free probability based capacity calculation for MIMO channels with transmit or receive correlation", *Intl Conf. on Wireless Networks, Comm. and Mobile Computing, New York, USA*, Jun. 2005.

[33] R. A. Horn and C. R. Johnson, "Topics in matrix anlysis", *Cambridge University Press*, 8th printing 2007.

[34] T. Roughgarden, "The price of anarchy is independent of the network topology", *J. of Computer and System Sciences*, Vol. 67, No. 2, pp. 341–364, 2003.

[35] J. R. Correa, A. S. Schulz and N. E. Stier-Moses, "On the inefficiency of equilibria in congestion games", *Proc. of the 11th Conf. on Integer Programming and Combinatorial Optimization (IPCO'05), Berlin, Germany*, Jun. 2005.

[36] P. Maillé and N. E. Stier-Moses, "Eliciting coordination with rebates", *Columbia W. P.* ♯ DRO-2007-04, May 2007.

[37] J. B. Lasserre, "A trace inequality for matrix product", *IEEE Trans. on Automatic Control*, Vol. 40, No. 8, pp. 1500–1501, 1995.

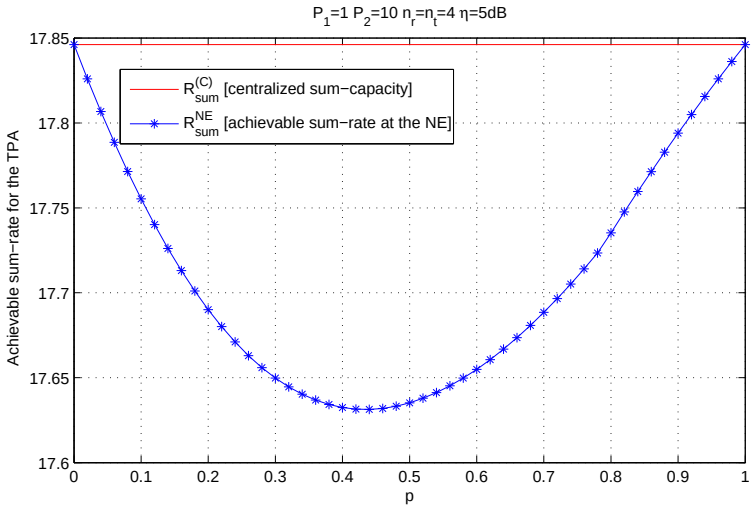

Fig. 1. Temporal PA game. Achievable network sum-rate versus p for $P_1 = 1$, $P_2 = 10$, $n_r = n_t = 4$, $\eta = 5$ dB. The sum-capacity of fading MIMO MACs is reached for both unfair SIC decoding schemes ($p_1^* = 0$ and $p_2^* = 1$) and is very close to this upper bound for any distribution of the coordination signal, $\forall p \in (0, 1)$.

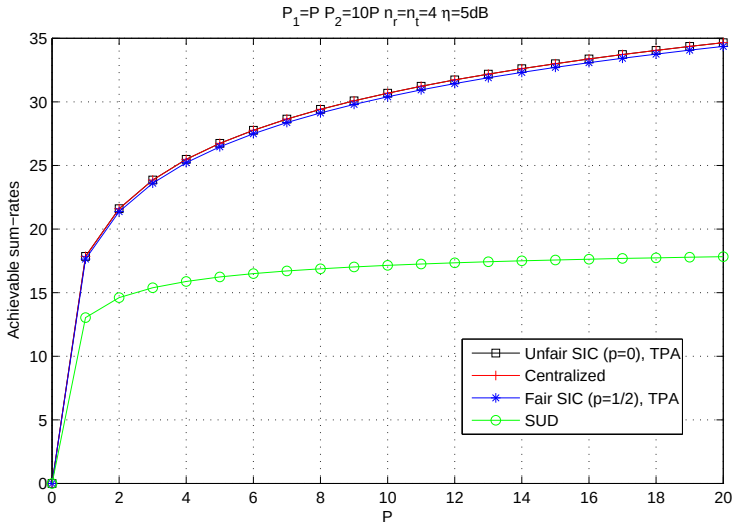

Fig. 2. Temporal PA game. MAC sum-rate versus the transmit power P for $P_1 = P$, $P_2 = 10P$, $n_r = n_t = 4$, $\eta = 5$ dB. Comparison between the fair SIC decoding scheme ($p = \frac{1}{2}$), the unfair SIC scheme ($p = 0$), and SUD decoding scheme.

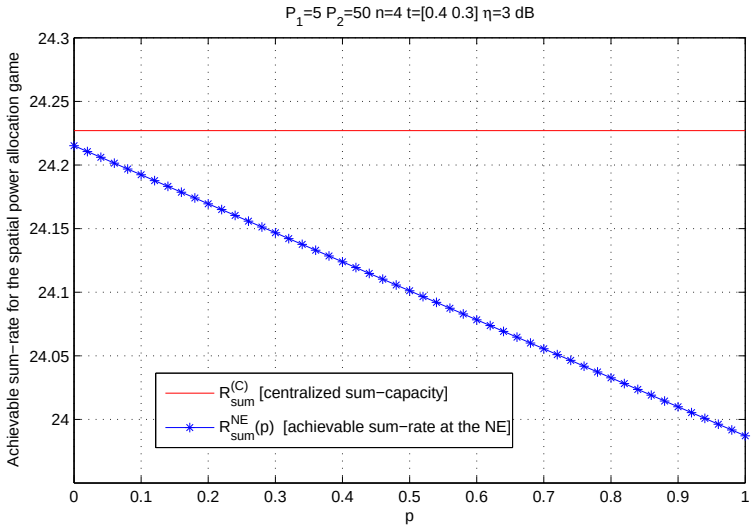

Fig. 3. Spatial PA game. MAC sum-rate versus p for $P_1 = 5$, $P_2 = 50$, $n_r = n_t = 4$, $\eta = 3$ dB, $t_1 = 0.4$, $t_2 = 0.3$. The achievable network sum-rate of fading MIMO MACs is linear w.r.t. $p \in [0, 1]$ and is very close to the centralized upper bound. The optimal distribution obtained with the Stackelberg game is $p^* = 0$.

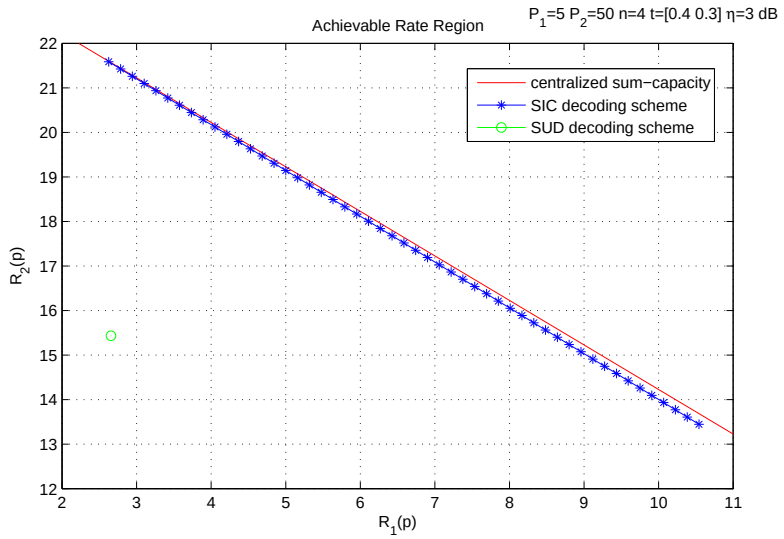

Fig. 4. Spatial PA game. Achievable rate region for $P_1 = 5$, $P_2 = 50$, $n_r = n_t = 4$, $\eta = 3$ dB, $t_1 = 0.4$, $t_2 = 0.3$. By varying p allows to move along a segment close to the centralized sum-capacity, similar to the SISO MAC channels.