



An information-theoretic look at MIMO energy-efficient communications

Elena Veronica Belmega, Samson Lasaulce

► To cite this version:

Elena Veronica Belmega, Samson Lasaulce. An information-theoretic look at MIMO energy-efficient communications. ACM International Conference on Performance Evaluation Methodologies and Tools (VALUETOOLS), Oct 2009, Pisa, Italy. pp.1-10. hal-00446961

HAL Id: hal-00446961

<https://hal.science/hal-00446961>

Submitted on 14 Jan 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

An information-theoretic look at MIMO energy-efficient communications

Elena Veronica Belmaga
Laboratoire des Signaux et Systèmes
CNRS - Supélec - Paris 11
91191, Gif-sur-Yvette, France
belmaga@lss.supelec.fr

Samson Lasaulce
Laboratoire des Signaux et Systèmes
CNRS - Supélec - Paris 11
91191, Gif-sur-Yvette, France
lasaulce@lss.supelec.fr

ABSTRACT

One of the main objectives of this paper is to provide an information-theoretic answer on how to maximize energy-efficiency in MIMO (multiple input multiple output) systems. In static and fast fading channels, for which arbitrarily reliable communications are possible, it is shown that the best precoding scheme (which includes power allocation) is to transmit at very low power ($\mathbf{Q} \rightarrow \mathbf{0}$). Whereas energy-efficiency is maximized in this regime, the latter also corresponds to communicating at very small transmission rates ($R \rightarrow 0$). In slow fading or quasi-static MIMO systems (where reliability cannot be ensured), based on the proposed information-theoretic performance measure, it is proven that energy-efficiency is maximized for a non-trivial precoding scheme; in particular, transmitting at zero power or saturating the transmit power constraint is suboptimal. The determination of the best precoding scheme is shown to be a new open problem. Based on this statement, the best precoding scheme is determined in several special but useful cases. As a second step, we show how to use the proposed energy-efficiency measure to analyze the important case of distributed power allocation in MIMO multiple access channels. Simulations show the benefits brought by multiple antennas for saving energy while guaranteeing the system to reach a given transmission rate target.

Categories and Subject Descriptors

H.4 [Information Systems Applications]: Miscellaneous;
D.2.8 [Software Engineering]: Metrics—*complexity measures, performance measures*

General Terms

Energy, Information Theory, Game Theory.

Keywords

Energy efficiency, MIMO channels, Outage probability, Power allocation games, Power control games.

The authors would like to thank the Network of Excellence NEWCOM++, Work Package 6.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

VALUETOOLS 2009, October 20-22, 2009 - Pisa, Italy.

Copyright 2009 ICST 978-963-9799-70-7/00/0004 \$5.00.

1. INTRODUCTION

During the past decade energy consumption has become a more and more important issue in wireless networks. In cellular networks for instance, mobile terminals are now equipped with relatively large screens, required to offer more and more functionalities (evolving to PCs), required to operate at higher transmission rates and better qualities and used for a longer duration within a day. On the fixed infrastructure side of cellular networks, energy consumption is also important since the number of base stations has increased dramatically and energy cost plays an important role in OPEX (operating expense); according to [1] energy costs due to (cellular) fixed networks is expected to be multiplied by about six within the decade 2002–2012. On the other hand, very significant progresses have been made in the art of designing wireless transmitters and receivers. This includes e.g., antennas and electronic circuits technology, signal processing algorithms, channel coding techniques and network protocols. The natural question which arises is: Will technological progresses be fast enough to control/decrease energy consumption at the (autonomous) terminal side and energy costs generated by the network infrastructure? Answering such a question is not easy and only elements of response can be provided. For this purpose, exploiting Moore's law and Information Theory can be of great help to get some insights into this issue. As far as this paper is concerned, we want to propose and exploit a suited information-theoretic performance metric to better understand this type of problems. Among the technological breakthroughs in communications we find MIMO (multiple input multiple output) systems [2][3][4]. Indeed, it is well known by now that for a point-to-point communication and given transmission quality target (say in terms of bit error rate), using multiple antennas at the transmitter or/and receiver in full diversity allows one to decrease the transmit power dramatically. Would therefore MIMO be the choicest solution to energy consumption issues? This paper provides some elements of response to this question.

The closest related works to the problem under consideration are [5][6][7]. In [5][6], the authors introduced a nice definition of energy-efficiency for the distributed power control problem over flat fading (single input single output, SISO) channels: it is defined as the net number of information bits that are transmitted without error per time unit (goodput) to the transmit power level. However, the nice framework of [6] needs to be broadened in order to attain the aforementioned objectives. First, we need to consider multiple antennas at the terminals (MIMO systems). In [7], the authors

generalized [6] to multi-carrier CDMA (code division multiple access) systems. Here we go a step further since multi-carrier systems are a special case of MIMO systems. More importantly, [6][7] do not provide an information-theoretic answer to the problem and therefore no indication on the ultimate performance achievable in a system where power control/allocation is decentralized. Another important technical difference is that optimizing the proposed performance metric only requires the knowledge of the statistics of the channel. Indeed, the proposed formulation aims at knowing how does a transmitter adapt its power to some channel parameters like path loss and share its power between its antennas to maximize energy-efficiency. In fact, the current state of the literature seems to indicate [9] that information-theoretic answers to the problem of energy-efficient communications, and therefore related problems like the distributed power control problem, are not satisfying. Indeed, a natural idea would be to take the Shannon transmission rate for the numerator of the ratio goodput to transmit power. For a single-user channel the energy-efficiency performance metric

would be $\frac{\log_2 \left(1 + \frac{p|h|^2}{\sigma^2} \right)}{p}$ where p is the transmit power, σ^2 the receive noise power and h the channel gain. We readily see that energy-efficiency of a point-to-point communication is maximized when $p \rightarrow 0$ that is, by transmitting with a very low power. This can be a very unsatisfying answer for engineers when a certain transmission rate target has to be reached since the corresponding (Shannon) transmission rate also tends to zero. Is this answer generalizable to MIMO systems? If yes, is this really all what information theory has to say about energy-efficient transmissions? Is there not an other way of defining energy-efficiency which leads to a more satisfying answer? One of the purposes of this paper is to answer these questions. An additional new feature with respect to [5][6][7] and related works (e.g., [8][9][10][11]) is that we conduct a discussion on the importance of the channel model that is, we not only look at slow fading channels (also called quasi-static or block fading channels) but also at static and fast fading channels.

The paper is structured as follows. We begin in Sec. 2 by describing the system model and the general assumptions. We investigate the power allocation policy that maximizes the energy-efficiency for a single user MIMO channel for different channel models: static channel (see Sec. 2.1), fast fading channel (see Sec. 2.2), slow fading channel (see Sec. 2.3). For the latter case, we define an energy efficiency function based on the outage probability and see that an analytical expression of the optimal power allocation policy is difficult to be found in general and will be further investigated assuming a uniform power allocation over the transmit antennas (power control problem). In Sec. 3, we conjecture that and provide analytical results for several particular but interesting cases. This conjecture will be used in Sec. 4 where the multi-user setting (the MIMO multiple access channel) is considered to prove the existence of a Nash equilibrium in the defined non-cooperative power control game. Some numerical simulation will be discussed in Sec. 5 and we end with concluding remarks and possible extensions.

2. CHANNEL MODEL AND UTILITY FUNCTION

Apart from one section, in all the paper, we focus on the

case of single-user MIMO channels. Indeed, the case of multiple access channels, which follows quite easily, will be addressed in Sec. 4. We thus assume one transmitter and one receiver and a standard signal model for the received signal. The equivalent baseband signal at the receiver can be written as:

$$\underline{y}(\tau) = \mathbf{H}(\tau)\underline{x}(\tau) + \underline{z}(\tau), \quad (1)$$

where \mathbf{H} is the $n_r \times n_t$ channel transfer matrix and n_t (resp. n_r) the number of transmit (resp. receive) antennas. The entries of \mathbf{H} are i.i.d. zero-mean unitary variance complex Gaussian random variables. The vector \underline{x} is the n_t -dimensional column vector of symbols transmitted and \underline{z} is an n_r -dimensional complex white Gaussian noise distributed as $\mathcal{N}(\underline{0}, \sigma^2 \mathbf{I})$. In this paper we are interested in the optimal way of allocating the transmit power between the available transmit antennas. We will denote by $\mathbf{Q} = \mathbb{E}[\underline{x}\underline{x}^H]$ the input covariance matrix (also called precoding matrix), which translates the chosen power allocation (PA) policy. The corresponding total power constraint is

$$\text{Tr}(\mathbf{Q}) \leq \bar{P}. \quad (2)$$

At last, the time index τ will be removed for the sake of clarity. In fact, depending on the rate at which τ varies three dominant classes of channels can be distinguished:

1. the class of static channels;
2. the class of fast fading channels;
3. the class of slow fading channels.

In this paper, we will dedicate most of our attention to the latter type of channels where an *important* point has to be mentioned. In this case, we introduce the general power allocation problem where the transmitter has to optimize the precoding matrix \mathbf{Q} based on the knowledge of the statistics of \mathbf{H} . However, we will mostly focus on the (already non-trivial) case where $\mathbf{Q} = p\mathbf{I}$. Under this assumption, the power allocation problem becomes a power control problem in which the transmitter, based on the statistics of \mathbf{H} , has to tune p optimally to maximize its energy-efficiency. The general problem is therefore left as (a very challenging) extension of this paper. We will now explain our motivation by analyzing cases 1 and 2.

2.1 Case of static channels

Here the frequency at which the channel matrix varies is strictly zero that is, \mathbf{H} is a constant matrix. We assume that both the transmitter and receiver know this matrix. We are exactly in the same framework as [4] and know that, for a given precoding strategy \mathbf{Q} , the transmitter can send $\log_2 \left| \mathbf{I} + \frac{\rho}{n_t} \mathbf{H} \mathbf{Q} \mathbf{H}^H \right|$ bits per channel use to the receiver reliably, with $\rho = \frac{1}{\sigma^2}$. Then, let us define the energy-efficiency of this communication by:

$$F_{\text{static}}(\mathbf{Q}) = \frac{\log_2 \left| \mathbf{I} + \frac{\rho}{n_t} \mathbf{H} \mathbf{Q} \mathbf{H}^H \right|}{\text{Tr}(\mathbf{Q})}. \quad (3)$$

The question is then to know whether the answer “do not transmit” (or more practically at a very small rate) to maximize energy-efficiency we have given in Sec. 1 for SISO channels also applies to MIMO channels. The answer is given by the following Theorem.

THEOREM 2.1 (STATIC MIMO CHANNELS). *The energy-efficiency of a MIMO communication over a static channel, measured by F_{static} , is maximized when $\mathbf{Q} = \mathbf{0}$ and this maximum is*

$$F_{\text{static}}^* = \frac{1}{\ln 2} \frac{\text{Tr}(\mathbf{H}\mathbf{H}^H)}{n_t \sigma^2}. \quad (4)$$

PROOF. As \mathbf{Q} is a non-negative Hermitian matrix, it can always be spectrally decomposed as $\mathbf{Q} = \mathbf{U}\mathbf{P}\mathbf{U}^H$ where \mathbf{P} is a diagonal matrix representing a given PA policy and \mathbf{U} a unitary matrix. Our goal is to prove that, for every \mathbf{U} , F_{static} is maximized when $\mathbf{P} = \text{diag}(0, 0, \dots, 0)$. To this end we rewrite F_{static} as

$$F_{\text{static}}(\mathbf{P}, \mathbf{U}) = \frac{\log_2 \left| \mathbf{I} + \sum_{i=1}^{n_t} p_i \underline{g}_i \underline{g}_i^H \right|}{\sum_{i=1}^{n_t} p_i}, \quad (5)$$

where \underline{g}_i represents the i^{th} column of the $n_r \times n_t$ matrix $\mathbf{G} = \sqrt{\frac{\rho}{n_t}} \mathbf{H}\mathbf{U}$ and proceed by induction on $n_t \geq 1$.

First, we introduce an auxiliary quantity (whose role will be made clear a little further)

$$E^{(n_t)}(p_1, \dots, p_{n_t}) \triangleq \text{Tr} \left(\left(\mathbf{I} + \sum_{i=1}^{n_t} p_i \underline{g}_i \underline{g}_i^H \right)^{-1} \left(\sum_{i=1}^{n_t} p_i \underline{g}_i \underline{g}_i^H \right) \right) - \log_2 \left| \mathbf{I} + \sum_{i=1}^{n_r} p_i \underline{g}_i \underline{g}_i^H \right|. \quad (6)$$

and prove by induction that it is negative that is, $\forall (p_1, \dots, p_{n_t}) \in \mathbb{R}_+^{n_t}$, $E^{(n_t)}(p_1, \dots, p_{n_t}) \leq 0$.

For $n_t = 1$, we have $E^{(1)}(p_1) = \text{Tr} \left[(\mathbf{I} + p_1 \underline{g}_1 \underline{g}_1^H)^{-1} \underline{g}_1 \underline{g}_1^H \right] - \log_2 \left| \mathbf{I} + p_1 \underline{g}_1 \underline{g}_1^H \right|$. The first order derivative of $E^{(1)}(p_1)$ w.r.t. p_1 is:

$$\frac{\partial E^{(1)}}{\partial p_1} = -p_1 [\underline{g}_1^H (\mathbf{I} + p_1 \underline{g}_1 \underline{g}_1^H)^{-1} \underline{g}_1]^2 \leq 0 \quad (7)$$

and thus $E^{(1)}(p_1) \leq E^{(1)}(0) = 0$.

Now, we assume that $E^{(n_t-1)}(\underline{p}) \leq 0$ and want to prove that $E^{(n_t)}(\underline{p}, p_{n_t}) \leq 0$, where $\underline{p} = (p_1, \dots, p_{n_t-1})$. It turns out that:

$$\frac{\partial E^{(n_t)}}{\partial p_{n_t}} = - \sum_{j=1}^{n_t} p_j \left| \underline{g}_j^H (\mathbf{I} + \sum_{i=1}^{n_t} p_i \underline{g}_i \underline{g}_i^H)^{-1} \underline{g}_{n_t} \right|^2 \leq 0, \quad (8)$$

and therefore $E^{(n_t)}(p_1, \dots, p_{n_t-1}, p_{n_t}) \leq E^{(n_t)}(p_1, \dots, p_{n_t-1}, 0) = E^{(n_t-1)}(p_1, \dots, p_{n_t-1}) \leq 0$.

As a second step of the proof, we want to prove by induction on $n_t \geq 1$ that

$$\arg \max_{\underline{p}, p_{n_t}} F_{\text{static}}^{(n_t)}(\underline{p}, p_{n_t}) = \underline{0}. \quad (9)$$

For $n_t = 1$ we have $F_{\text{static}}^{(1)}(p_1) = \frac{\log_2 |\mathbf{I} + p_1 \underline{g}_1 \underline{g}_1^H|}{p_1} = \frac{\log_2(1 + p_1 \underline{g}_1^H \underline{g}_1)}{p_1}$ which reaches its maximum in $p_1 = 0$.

Now, we assume that $\arg \max_{\underline{p}} F_{\text{static}}^{(n_t-1)}(\underline{p}) = \underline{0}$ and want to prove that $\arg \max_{(\underline{p}, p_{n_t})} F_{\text{static}}^{(n_t)}(\underline{p}, p_{n_t}) = \underline{0}$.

Let $k = \arg \min_{i \in \{1, \dots, n_t\}} \text{Tr} \left[\left(\mathbf{I} + \sum_{j=1}^{n_t} p_j \underline{g}_j \underline{g}_j^H \right)^{-1} \underline{g}_i \underline{g}_i^H \right]$. By

calculating the first order derivative of $F_{\text{static}}^{(n_t)}$ w.r.t. p_k one obtains that:

$$\frac{\partial F_{\text{static}}^{(n_t)}}{\partial p_k} = \frac{\mathcal{N}}{\left(\sum_{i=1}^{n_t} p_i \right)^2}, \quad (10)$$

with

$$\mathcal{N} = \left(\sum_{i=1}^{n_t} p_i \right) \text{Tr} \left[\left(\mathbf{I} + \sum_{j=1}^{n_t} p_j \underline{g}_j \underline{g}_j^H \right)^{-1} \underline{g}_k \underline{g}_k^H \right] - \log_2 \left| \mathbf{I} + \sum_{i=1}^{n_t} p_i \underline{g}_i \underline{g}_i^H \right| \quad (11)$$

and thus $\frac{\partial F_{\text{static}}^{(n_t)}}{\partial p_k} \leq \frac{E^{(n_t)}(p_1, \dots, p_{n_t})}{\left(\sum_{i=1}^{n_t} p_i \right)^2} \leq 0$ and $p_k^* = 0$ for

all $p_1, \dots, p_{k-1}, p_{k+1}, \dots, p_{n_t}$. We obtain that

$F_{\text{static}}^{(n_t)}(p_1, \dots, p_{k-1}, 0, p_{k+1}, \dots, p_{n_t}) = F_{\text{static}}^{(n_t-1)}(p_1, \dots, p_{k-1}, p_{k+1}, \dots, p_{n_t})$, which is maximized when $(p_1, \dots, p_{k-1}, p_{k+1}, \dots, p_{n_t}) = \underline{0}$ by assumption. We therefore have that $\mathbf{Q}^* = \mathbf{U}\mathbf{0}\mathbf{U}^H = \mathbf{0}$ is the solution that maximizes the function $F_{\text{static}}(\mathbf{Q})$. Lastly, to find the maximum reached by F_{static} one just needs to consider the equivalent of the $\log_2 \left| \mathbf{I} + \frac{\rho}{n_t} \mathbf{H}\mathbf{Q}\mathbf{H}^H \right|$ around $\mathbf{Q} = \mathbf{0}$

$$\log_2 \left| \mathbf{I} + \frac{\rho}{n_t} \mathbf{H}\mathbf{Q}\mathbf{H}^H \right| \sim \frac{\rho}{n_t} \text{Tr}(\mathbf{H}\mathbf{Q}\mathbf{H}^H) \quad (12)$$

and takes $\mathbf{Q} = q\mathbf{I}$ with $q \rightarrow 0$. \square

Therefore we see that for static MIMO channels, if energy-efficiency is defined by Eq. (3), it is maximized by transmitting at a very low power. As already mentioned, from a more practical perspective, this means that if the considered application supports very low data rates, one has to transmit with a very low power to make the communication energy-efficient. This kind of scenario occurs for example, when deploying sensors in the ocean to measure a temperature field (which varies very slowly). But this answer is not acceptable from the moment a minimum transmission rate is required, which is the most common situation (cellular networks, satellite communications, wireless local area networks, etc.)

2.2 Case of fast fading channels

In this section, the frequency at which the channel matrix varies is the reciprocal of the symbol duration ($\underline{x}(\tau)$ being a symbol) that is, it can be different for each channel use. Therefore the channel varies over a transmitted codeword (or packet) and more precisely each codeword sees as many channel realizations as the number of symbols per codeword. Because of the corresponding self-averaging effect, the following Shannon transmission rate (also called EMI for ergodic mutual information) can be achieved on each transmitted codeword by using the precoding strategy \mathbf{Q} :

$$R_{\text{fast}}(\mathbf{Q}) = \mathbb{E}_{\mathbf{H}} \left[\log_2 \left| \mathbf{I} + \frac{\rho}{n_t} \mathbf{H}\mathbf{Q}\mathbf{H}^H \right| \right]. \quad (13)$$

Interestingly, $R_{\text{fast}}(\mathbf{Q})$ can be maximized w.r.t. \mathbf{Q} by knowing the statistics of \mathbf{H} only that is, $\mathbb{E}[\mathbf{H}\mathbf{H}^H]$ if we make

the standard assumption that the entries of \mathbf{H} are complex Gaussian random variables. In practice, this means that only the knowledge of the path loss, power-delay profile, antenna correlation profile, etc is required at the transmitter to maximize the transmission rate. At the receiver however, the instantaneous knowledge of \mathbf{H} is required. In this framework, let us define energy-efficiency by:

$$F_{\text{fast}}(\mathbf{Q}) = \frac{\mathbb{E}_{\mathbf{H}} \left[\log_2 \left| \mathbf{I} + \frac{\rho}{n_t} \mathbf{H} \mathbf{Q} \mathbf{H}^H \right| \right]}{\text{Tr}(\mathbf{Q})}. \quad (14)$$

By rewriting $F_{\text{fast}}(\mathbf{Q})$ as

$$F_{\text{fast}}(\mathbf{Q}) = \mathbb{E}_{\mathbf{H}} \left[\frac{\log_2 \left| \mathbf{I} + \sum_{i=1}^{n_t} p_i \underline{g}_i \underline{g}_i^H \right|}{\sum_{i=1}^{n_t} p_i} \right] \quad (15)$$

and applying Theorem 2.1 for each realization of the channel matrix, the following theorem can be proven.

THEOREM 2.2 (FAST FADING MIMO CHANNELS). *The energy-efficiency of a MIMO communication over a fast fading channel, measured by F_{fast} , is maximized when $\mathbf{Q} = \mathbf{0}$ and this maximum is*

$$F_{\text{fast}}^* = \frac{1}{\ln 2} \frac{\text{Tr}(\mathbb{E}[\mathbf{H} \mathbf{H}^H])}{n_t \sigma^2}. \quad (16)$$

Here again, we see that for fast fading MIMO channels, it is not possible to have both an energy-efficient communication in the sense of F_{fast} and transmit reliably (in the sense of Shannon i.e., with an arbitrary small frame error rate) at a minimum required transmission rate. Interestingly, in slow fading MIMO channels, where outage events are unavoidable, we have found that in general there is a non-trivial information-theoretic and very interesting answer to this question. This is the purpose of the remaining of this paper.

2.3 Case of slow fading channels

In this section and in the remaining of this paper, the frequency at which the channel matrix varies is the reciprocal of the codeword/packet/frame/block duration that is, the channel remains constant over a codeword and varies from codeword to codeword. As a consequence, when the channel matrix remains constant over a certain block duration much smaller than the channel coherence time, the averaging effect we have mentioned for fast fading MIMO channels does not occur here and one has to communicate at rates smaller than the ergodic capacity (maximum of the EMI). The maximum EMI is therefore a rate upper bound for slow fading MIMO channels and only a fraction of it can be achieved (see [12] for more information about the famous diversity-multiplexing tradeoff). In fact, since the mutual information becomes a random variable, varying from block to block, it is not possible (in general) to guarantee at 100 % that it is above a certain threshold. A suited performance metric to study slow-fading channels [14] is the probability of an outage for a given transmission rate target R . It allows one to quantify the probability that the rate target R is not reached by using a good channel coding scheme and is defined as follows:

$$P_{\text{out}}(\mathbf{Q}, R) = \Pr \left[\log_2 \left| \mathbf{I} + \frac{\rho}{n_t} \mathbf{H} \mathbf{Q} \mathbf{H}^H \right| < R \right]. \quad (17)$$

In terms of information assumption, here again, it can be checked that only the second-order statistics of \mathbf{H} are required to optimize the precoding strategy \mathbf{Q} (and therefore the power allocation policy over its eigenvalues). In this framework we propose to define the energy-efficiency as follows:

$$F_{\text{slow}}(\mathbf{Q}, R) = R \frac{1 - P_{\text{out}}(\mathbf{Q}, R)}{\text{Tr}(\mathbf{Q})}. \quad (18)$$

It is important to keep in mind that even for $P_{\text{out}}(\mathbf{Q}, R)$ the optimization problem over \mathbf{Q} is still an open problem [4][13]. More precisely the optimal solution has been conjectured [4] but not proven. We will come back to this issue later on.

PROPOSITION 2.3 (SLOW FADING MIMO CHANNELS). *The energy-efficiency function $F_{\text{slow}}(\mathbf{Q}, R)$ is maximized, in general, for $\mathbf{Q} \neq \mathbf{0}$.*

PROOF. The proof has two parts. First, we start by proving that if the optimal solution is different than the uniform spatial power allocation $\mathbf{P}^* \neq p\mathbf{I}$ with $p \in (0, \frac{\bar{P}}{n_t}]$ then the solution is not trivial $\mathbf{P}^* \neq \mathbf{0}$. We proceed by reductio ad absurdum. We assume that the optimal solution is trivial $\mathbf{P}^* = \mathbf{0}$. This means that when fixing $(p_2, \dots, p_{n_t}) = (0, \dots, 0)$ the optimal $p_1 \in [0, \bar{P}]$ that maximizes the energy efficiency function is $p_1^* = 0$. The energy efficiency function becomes:

$$F_{\text{slow}}(p_1, R) = R \frac{1 - \Pr \left[\log_2(1 + \frac{\rho}{n_t} p_1 \|\underline{h}_1\|^2) < R \right]}{p_1} \quad (19)$$

where \underline{h}_1 represents the first column of the channel matrix \mathbf{H} . Knowing that the elements in \underline{h}_1 are i.i.d. $h_{1j} \sim \mathcal{CN}(0, 1)$ for all $j \in \{1, \dots, n_r\}$ we have that $|h_{1j}|^2 \sim \text{expon}(1)$. The random variable $\|\underline{h}_1\|^2 = \sum_{j=1}^{n_r} |h_{1j}|^2$ is the sum of n_r i.i.d. exponential random variables of parameter $\lambda = 1$ and thus follows an $2n_r$ chi-square distribution (or an n_r Erlang distribution) whose c.d.f. is known and given by $\zeta(x) = 1 - \exp(-x) \sum_{k=0}^{n_r-1} \frac{x^k}{k!}$. We can explicitly calculate the outage probability and obtain the energy efficiency function:

$$F_{\text{slow}}(p_1, R) = R \exp \left(-\frac{c}{p_1} \right) \sum_{k=0}^{n_r-1} \frac{c^k}{k!} \frac{1}{p_1^{k+1}} \quad (20)$$

where $c = \frac{n_t(2^R-1)}{\rho} > 0$. It is easy to check that $\lim_{p_1 \rightarrow 0} F_{\text{slow}}(p_1, R) = 0$, $\lim_{p_1 \rightarrow \infty} F_{\text{slow}}(p_1, R) = 0$. By evaluating the first derivative w.r.t. p_1 , it is easy to check that the maximum is achieved for $p_1^* = \frac{c}{x^*} \geq 0$ where x^* is the unique positive solution of the following equation:

$$\frac{1}{(n_r-1)!} x^{n_r} - \sum_{k=0}^{n_r-1} \frac{x^k}{k!} = 0. \quad (21)$$

Considering the power constraint the optimal transmission power is $p_1^* = \min\{\frac{n_t(2^R-1)}{x^* \rho}, \bar{P}\}$, which contradicts the hypothesis and thus if the optimal solution is different than the uniform spatial power allocation then the solution is not trivial $\mathbf{P}^* \neq \mathbf{0}$. \square

We make the following conjecture on F_{slow} .

CONJECTURE 2.4 (SLOW FADING MIMO CHANNELS). *The energy-efficiency function $F_{\text{slow}}(\mathbf{Q}, R)$ is a quasi-concave function of p where $\mathbf{Q} = p\mathbf{I}$.*

One of our objectives in this paper is to prove that, for special but useful cases, part (ii) can be proven and therefore Conjecture 2.4 becomes a theorem in these cases.

3. THE FUNCTION F_{SLOW} IS QUASI-CONCAVE

In this section, we focus on slow fading or quasi-static MIMO channels and study the proposed energy-efficiency function (Eq. (18)). It is useful to note that we do not need to analyze that F_{slow} with respect to \mathbf{Q} but to \mathbf{P} only. Indeed, for any spectral decomposition of $\mathbf{Q} = \mathbf{U}\mathbf{P}\mathbf{U}^H$ (always with \mathbf{U} unitary) we have that $\text{Tr}(\mathbf{Q}) = \text{Tr}(\mathbf{P})$, which means that the denominator of F_{slow} is independent of the eigenvectors of \mathbf{Q} . Additionally the distribution of the matrix $\mathbf{H}\mathbf{U}$ is the same as the matrix \mathbf{H} since the latter is assumed to be a Gaussian random matrix with i.i.d. entries. Therefore the numerator is independent of \mathbf{U} provided it is unitary. Finally, this shows that there is no loss of optimality by choosing $\mathbf{U} = \mathbf{I}$ that is, the search for the optimal precoding matrix can be restricted to diagonal non-negative matrices. This explains why we can restrict our analysis only w.r.t. \mathbf{P} (which is diagonal). In this paper, in part because of the lack of space and simplicity reasons, we restrict our attention to the uniform power allocation policies i.e., $\mathbf{P} = p\mathbf{I}$ and therefore want to prove that F_{slow} is quasi-concave in p , which is already a non-trivial very difficult problem. A third motivation for making this assumption is that, even for the outage probability alone, the optimization problem w.r.t. \mathbf{Q} is known to be an open problem but all observed results have confirmed the conjecture by Telatar [4] namely, allocating the transmit power uniformly over a subset of antennas minimizes P_{out} . Under the assumption made ($\mathbf{Q} = p\mathbf{I}$), it is known from [15] that when the benefit function, which is

$$1 - P_{\text{out}}(p, R) \triangleq f(p) \quad (22)$$

in our framework, is sigmoidal in p the energy-efficiency function F_{slow} is quasi-concave in p . Since the release of [15] there seems to be only one example of energy-efficiency function used in the literature. Indeed, all papers dealing with energy-efficiency ([6][7][11] etc) the authors always consider the following (empirical) formula for the packet error rate $f(x) = (1 - e^{-x})^M$ (x being the SINR), where M is the packet length. Very interestingly, the proposed (information-theoretic) benefit function, has also the nice properties stated by [15], which gives a more fundamental view to the problem initially tackled by [5]. In particular we will see that for all the cases treated $f(p)$ is sigmoidal. We will treat the case of SISO channels and MIMO channels with arbitrary numbers of antennas in the low and high SNR regimes. Finally, we will treat the case of large MIMO systems that is, when the numbers of antennas are large. Here are some motivations explaining why the latter case is in fact more common than what could be a priori thought:

- First of all, approximates provided under the large system assumption can be very accurate for reasonably small numbers of antennas [16][17]. In fact, to provide a rigorous answer to the question whether it is very accurate in our context a complete convergence analysis should be conducted, which is not done here because it is not our main objective and also because of the lack of space (see [18] to know more about this type of analysis).

- The matrix \mathbf{H} can be effectively large in some contexts, like virtual MIMO networks, where the network sum-rate is needed. In this case, the number of rows (resp. columns) can be the total number antennas of the group of base (resp. mobile) stations forming the virtual receiver (resp. transmitter).
- In this paper, the signal model under consideration is used to study energy-efficiency of communications with multi-antenna transmitters and receivers but the proposed analysis is directly exploitable for other systems where the dimensions of \mathbf{H} correspond to other quantities like the number of sub-carriers in OFDM systems, the spreading factor in CDMA systems, etc. This means that the large system assumption has also to be considered regarding this possibility.
- The case of large systems constitutes an additional special case for which Conjecture 2.4 becomes a Theorem. In fact, it will allow us to prove easily certain results whose proof requires much more efforts in the finite case.
- An important result will be drawn from the proposed asymptotic analysis.

3.1 SISO systems

In this section, we assume that the transmitter and receiver have a single antenna that is, $\mathbf{Q} = p$. Our goal is to prove that $F_{\text{slow}}(p, R)$ is quasi-concave in p for a given rate target R and determine in which point energy-efficiency is maximized. The following proposition summarizes the results we have found for the SISO case.

PROPOSITION 3.1 (SISO CHANNELS). *The energy-efficiency function $F_{\text{slow}}(p, R)$ is quasi-concave in p and reaches its maximum in*

$$p^* = \min \left\{ \frac{2^R - 1}{\rho}, \bar{P} \right\}. \quad (23)$$

PROOF. By assumption we have that $h \sim \mathcal{CN}(0, 1)$ and thus $|h|^2$ is an exponentially distributed random variable. The benefit function $f(p)$ can therefore be explicated:

$$f(p) = \exp \left(-\frac{c}{p} \right) \quad (24)$$

with $c = \frac{2^R - 1}{\rho}$. From the second derivative of f

$$f''(p) = \frac{c(c - 2p) \exp(-\frac{c}{p})}{p^4} \quad (25)$$

we see that f is convex on the open interval $(0, \frac{c}{2}]$ and concave on $(\frac{c}{2}, +\infty)$. This means that f is sigmoidal and therefore $F_{\text{slow}}(p, R)$ is quasi-concave. Now, by looking at the first derivative of $F_{\text{slow}}(p, R)$ which is easily shown to be

$$\frac{\partial F_{\text{slow}}}{\partial p} = R \frac{(c - p) \exp(-\frac{c}{p})}{p^3} \quad (26)$$

we see that $F_{\text{slow}}(p, R)$ has a unique maximum in $p^* = c$ if $c < \bar{P}$ and in $p^* = \bar{P}$ otherwise. \square

We see that the proposed efficiency-function has all the desirable properties stated in [15]. In order to fully understand how the statistics of the channel influences the optimal power control policy we also provide the result obtained

when the channel variance is no longer assumed to be unitary that is, $\mathbb{E}[|h|^2] \neq 1$:

$$p^* = \min \left\{ \frac{\sigma^2}{\mathbb{E}[|h|^2]}(e^R - 1), \bar{P} \right\}. \quad (27)$$

3.2 Low SNR regime

In the regime of low SNR, defined as $\rho \rightarrow 0$, there is a very simple approximation for the determinant $|\mathbf{I} + \frac{\rho p}{n_t} \mathbf{H} \mathbf{H}^H|$, whatever the numbers of transmit and receive antennas. The corresponding approximate for F_{slow} can be shown to be quasi-concave and is easy to be maximized. This is the purpose of the following Proposition.

PROPOSITION 3.2 (LOW SNR REGIME). *When $\rho \rightarrow 0$, any first-order equivalent of $F_{\text{slow}}(p, R)$ can be shown to be quasi-concave and maximized in $p^* = \frac{c}{x^*} = \frac{n_t(2^R - 1)}{x^* \rho}$ where x^* is the unique positive solution of the following equation:*

$$\frac{1}{(n_r n_t - 1)!} x^{n_t n_r} - \sum_{k=0}^{n_r n_t - 1} \frac{x^k}{k!} = 0.$$

PROOF. When $\rho \rightarrow 0$ we have

$$\left| \mathbf{I} + \frac{\rho p}{n_t} \mathbf{H} \mathbf{H}^H \right| = 1 + \frac{\rho p}{n_t} \text{Tr}(\mathbf{H} \mathbf{H}^H) + \mathcal{O}(\rho^2). \quad (28)$$

Since $|h(i, j)|^2$ is exponentially distributed for any $(i, j) \in \{1, \dots, n_t\} \times \{1, \dots, n_r\}$, the quantity $\text{Tr}(\mathbf{H} \mathbf{H}^H) = \sum_{i=1}^{n_t} \sum_{j=1}^{n_r} |h_{ij}|^2$ is an $2n_r n_t$ chi-square distributed random variable. Its cumulative distribution function is

$$\varsigma(x) = 1 - \exp(-x) \sum_{k=0}^{n_r n_t - 1} \frac{x^k}{k!}. \quad (29)$$

It is therefore possible to explicit the energy-efficiency function:

$$F_{\text{slow}}(p, R) = R \exp\left(-\frac{c}{p}\right) \sum_{k=0}^{n_r n_t - 1} \frac{c^k}{k!} \frac{1}{p^{k+1}}. \quad (30)$$

By evaluating the first-order and second-order derivatives of the benefit function $f(p) = \exp\left(-\frac{c}{p}\right) \sum_{k=0}^{n_r n_t - 1} \frac{c^k}{k!} \frac{1}{p^k}$ which are

$$f'(p) = \frac{1}{(n_r n_t - 1)!} \frac{c^{n_r n_t}}{p^{n_r n_t + 1}} \exp\left(-\frac{c}{p}\right) \quad (31)$$

$$f''(p) = \frac{1}{(n_r n_t - 1)!} \frac{c^{n_r n_t}}{p^{n_r n_t + 3}} [c - p(n_r n_t + 1)] \exp\left(-\frac{c}{p}\right) \quad (32)$$

we observe that the benefit function has a unique inflexion point $\tilde{p} = \frac{c}{n_r n_t + 1}$ and thus $f(p)$ is a sigmoid shaped function w.r.t. p . From [15] we have thus that $F_{\text{slow}}(p, R)$ is quasi-concave w.r.t. p . By evaluating the first derivative of F_{slow} w.r.t. p we obtain:

$$\frac{\partial F_{\text{slow}}}{\partial p} = \frac{R}{p^2} \left[\frac{1}{(n_r n_t - 1)!} \left(\frac{c}{p}\right)^{n_r n_t} - \sum_{k=0}^{n_r n_t - 1} \left(\frac{c}{p}\right)^k \frac{1}{k!} \right] \exp\left(-\frac{c}{p}\right). \quad (33)$$

Applying the same reasoning as in the proof of Proposition 3.1 the wanted result follow. If we no longer assume unitary variances for the channel matrix entries but $\mathbb{E}[|h_{ij}|^2] =$

$\mathbb{E}[|h|^2] \neq 1$ for all i, j then the optimal power control policy becomes:

$$p^* = \frac{n_t(2^R - 1)}{x^* \rho \mathbb{E}[|h|^2]}. \quad (34)$$

□

3.3 High SNR regime

In the high SNR regime, the counterpart of Proposition 3.3 is as follows.

PROPOSITION 3.3 (HIGH SNR REGIME). *When $\rho \rightarrow +\infty$, any first-order equivalent of $F_{\text{slow}}(p, R)$ can be shown to be quasi-concave.*

PROOF. In the high SNR regime we can approximate

$$\begin{aligned} \left| \mathbf{I} + \frac{\rho p}{n_t} \mathbf{H} \mathbf{H}^H \right| &= \left| \frac{\rho p}{n_t} \mathbf{H} \mathbf{H}^H \right| \left| \mathbf{I} + \frac{n_t}{\rho p} (\mathbf{H} \mathbf{H}^H)^{-1} \right| \\ &= \left| \frac{\rho p}{n_t} \mathbf{H} \mathbf{H}^H \right| \left\{ 1 + \frac{n_t}{\rho p} \text{Tr}[(\mathbf{H} \mathbf{H}^H)^{-1}] + \mathcal{O}\left(\frac{1}{\rho^2}\right) \right\}, \end{aligned}$$

where the pdf of $|\mathbf{H} \mathbf{H}^H|$ is given in [24]. □

3.4 Large MIMO systems

In this section the noise level is assumed to be fixed that is, $\rho = \text{const}$ and can be arbitrary. The results we have obtained can be summarized in the following proposition.

PROPOSITION 3.4 (LARGE MIMO SYSTEMS). *If one operates in one of the following asymptotic regimes:*

- (a) $n_t < +\infty$ and $n_r \rightarrow \infty$;
- (b) $n_t \rightarrow +\infty$ and $n_r < +\infty$;
- (c) $n_t \rightarrow +\infty$, $n_r \rightarrow \infty$ with $\lim_{n_i \rightarrow +\infty, i \in \{t, r\}} \frac{n_t}{n_r} = \beta < +\infty$,

then $F_{\text{slow}}(p, R)$ is quasi-concave.

PROOF. *Regime (a).* In this asymptotic regime we know from [25] that the mutual information converges in distribution to a Gaussian random variable:

$$\log_2 \left| \mathbf{I} + \frac{\rho p}{n_t} \mathbf{H} \mathbf{H}^H \right| \rightarrow \mathcal{N} \left(n_t \log_2 \left(1 + \frac{n_r}{n_t} \rho p \right), \frac{n_t}{n_r} \log_2(e) \right). \quad (35)$$

As a consequence the benefit function can be easily expressed:

$$f(p) = Q \left(\frac{R - n_t \log_2 \left(1 + \frac{n_r}{n_t} \rho p \right)}{\sqrt{\frac{n_t}{n_r} \log_2(e)}} \right) \quad (36)$$

where $Q(x) \triangleq \frac{1}{\sqrt{2\pi}} \int_x^{+\infty} \exp\left(-\frac{t^2}{2}\right) dt$. From Eq. (36) the second derivative w.r.t. p can be calculated. Let us define the auxiliary function g by $g_a(p) = \frac{R - n_t \log_2 \left(1 + \frac{n_r}{n_t} \rho p \right)}{\sqrt{\frac{n_t}{n_r} \log_2(e)}}$. We thus have $f(p) = Q(g_a(p))$. The first order derivative of f is then given by:

$$f'(p) = -\frac{1}{\sqrt{2\pi}} g'_a(p) \exp\left(-\frac{g_a(p)^2}{2}\right) \quad (37)$$

and the second derivative is

$$f''(p) = \frac{1}{\sqrt{2\pi}} [g_a(p)(g'_a(p))^2 - g''_a(p)] \exp\left(-\frac{g_a(p)^2}{2}\right). \quad (38)$$

Finally we have that

$$f''(p) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{g_a(p)^2}{2}\right) \frac{n_r^3 \rho^2}{n_t^2 \log_2(e)} \frac{1}{\left(1 + \frac{n_r}{n_t} \rho p\right)^2} \left(-\sqrt{\frac{n_r}{n_t} \log_2(e)} + n_t g_a(p)\right) \quad (39)$$

We see that f has a unique inflection point:

$$\tilde{p} = \frac{n_t}{n_r \rho} \left\{ \exp\left[\frac{1}{n_t} \left(R - \frac{\log_2(e)}{n_r}\right)\right] - 1 \right\}, \quad (40)$$

which is positive when n_r is large. Therefore $f(\theta)$ is a sigmoidal function. Furthermore, when $n_r \rightarrow \infty$ the inflection point $\tilde{p} \rightarrow 0$ such that $f(p)$ is a concave function w.r.t. p which implies that the optimal transmission power $p^* \rightarrow 0$.

Regime (b). Here again we know from [25] that the mutual information converges in distribution to a Gaussian random variable:

$$\log_2 \left| \mathbf{I} + \frac{\rho p}{n_t} \mathbf{H} \mathbf{H}^H \right| \rightarrow \mathcal{N} \left(n_r \log_2(1 + \rho p), \left(\sqrt{\frac{n_r}{n_t}} \frac{\rho p \log_2(e)}{1 + \rho p} \right)^2 \right). \quad (41)$$

Here we have that $f(p) = Q(g_b(p))$ with

$$g_b(p) = \sqrt{\frac{n_t}{n_r}} \frac{1 + \rho p}{\rho p \log_2(e)} [R - n_r \log_2(1 + \rho p)]. \quad (42)$$

We observe that $g_b(p)$ is proportional to $\sqrt{n_t}$ and thus $g'_b(p)$, $g''_b(p)$ will also be proportional to $\sqrt{n_t}$. In the equation $f''(p) = 0$ the term in $g''(p)$ can be neglected in comparison to $(g'_b(p))^2 g_b(p)$ (see Eq. (38)) when $n_t \rightarrow \infty$ to obtain $(g'_b(p))^2 g_b(p) = 0$. Since

$$g'_b(p) = \sqrt{\frac{n_t}{n_r}} \frac{-n_r \rho p - R + n_r \log_2(1 + \rho p)}{\rho p^2 \log_2(e)} \leq 0 \quad (43)$$

we see that f has a unique inflection point $\tilde{p} = \frac{2^{R/n_r} - 1}{\rho}$. Here we observe that the inflexion point does not depend on $n_t \rightarrow \infty$ and thus the optimal transmit power will not be trivial $p^* > 0$.

Regime (c). The reasoning is always the same. For this regime we exploit some results from random matrix theory derived in [26]. The mutual information converges in distribution to a Gaussian random variable:

$$\log_2 \left| \mathbf{I} + \frac{\rho p}{n_t} \mathbf{H} \mathbf{H}^H \right| \rightarrow \mathcal{N}(n_t \mu_I, \sigma_I^2) \quad (44)$$

where $\mu_I = \beta \log_2(1 + \rho p(1 - \alpha)) - \alpha + \log_2(1 + \rho p(\beta - \alpha))$, $\sigma_I^2 = -\log_2\left(1 - \frac{\alpha^2}{\beta}\right)$, $\alpha = \frac{1}{2} \left(1 + \beta + \frac{1}{\rho p} - \sqrt{\left(1 + \beta + \frac{1}{\rho p}\right)^2 - 4\beta}\right)$.

It can be checked that $(g'_c(p))^2 g_c(p) - g''_c(p) = 0$ has a unique solution where $g_c(p) = \frac{R - n_t \mu_I(p)}{\sigma_I(p)}$. We obtain $g'_c(p) = \frac{n_t \mu_I \sigma'_I - n_t \mu'_I \sigma_I - R \sigma'_I}{\sigma_I^2}$ and

$$g''_c(p) = \frac{(n_t \mu_I \sigma''_I - n_t \mu'_I \sigma'_I - R \sigma''_I) \sigma_I^2 - 2 \sigma_I \sigma'_I (n_t \mu_I \sigma'_I - n_t \mu'_I \sigma_I - R \sigma'_I)}{\sigma_I^4}.$$

We observe that, in the equation $(g'_c(p))^2 g_c(p) - g''_c(p) = 0$, there are terms in n_t^3 , n_t^2 , n_t and constant terms w.r.t. n_t . When n_t becomes sufficiently large we can neglect the first order terms such that the solution is given by $\mu_I(p) = 0$. It can be shown that $\mu_I(0) = 0$ and that μ_I is an increasing function w.r.t. p which implies that the unique solution is $\tilde{p} = 0$. Similarly to regime (a) we obtain the trivial solution $p^* = 0$. \square

4. NON-COOPERATIVE POWER ALLOCATION GAMES

So far we have been considering point-to-point MIMO communications or single-user MIMO channels. In this section we show how the quasi-concavity property of the energy-efficiency function as defined by Eq. (18) can be used in multi-user channels. More specifically we study the distributed power allocation problem in MIMO multiple access channels (MAC). We assume one base station (BS) and K mobile stations (MS). Each mobile station is a transmitter sending some messages to the base station which decodes all the messages by applying single-user decoding. No interference cancellation technique is assumed here (see e.g., [28] to know more about power allocation games in MIMO MACs where the users' performance criterion is the transmission rate and receiver applies successive interference cancellation). The network is said to be distributed or decentralized in the sense that each user can choose freely its power allocation 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) transmitter is free to ignore the policy intended for him if the latter does not maximize its performance criterion. In the described framework the signal received at the BS writes as:

$$\underline{y} = \sum_{k=1}^K \mathbf{H}_k \underline{x}_k + \underline{z}, \quad (45)$$

where \mathbf{H}_k is the $n_r \times n_t$ channel matrix, \underline{x}_k the n_t -dimensional column vector of symbols transmitted by user $k \in \mathcal{K} \triangleq \{1, \dots, K\}$ and \underline{z} is an n_r -dimensional complex white Gaussian noise distributed as $\mathcal{N}(0, \sigma^2 \mathbf{I})$. We will denote by $\mathbf{Q}_k = \mathbb{E}[\underline{x}_k \underline{x}_k^H]$ the precoding matrix transmitter k implements. In practice, the latter is subject to a constraint:

$$\text{Tr}(\mathbf{Q}_k) \leq \bar{P}_k \quad (46)$$

but as we have already shown it is not always optimal to use all the available transmit power to maximize energy-efficiency.

Now we can define properly the power allocation game modeling the type of wireless networks under investigation. The appropriate game model is here a static strategic-form non-cooperative game with complete information and rational players.

DEFINITION 1 (POWER ALLOCATION GAME). *The power allocation game is completely described by the triplet $\mathcal{G} = (\mathcal{K}, \{\mathcal{S}_k\}_{k \in \mathcal{K}}, \{u_k\}_{k \in \mathcal{K}})$ where:*

- $\mathcal{K} = \{1, \dots, K\}$ is the set of transmitters (players);
- $\mathcal{S}_k = \{\mathbf{Q}_k \in \mathcal{M}_{n_t \times n_t}(\mathbb{C}) \mid \mathbf{Q}_k \succeq 0, \text{Tr}(\mathbf{Q}_k) \leq \bar{P}_k\}$ is the strategy set for player k , where $\mathcal{M}_{n_t \times n_t}(\mathbb{C})$ denotes the set of square complex matrices of dimension n_t ;
- $u_k(\mathbf{Q}_k, \mathbf{Q}_{-k}) = \frac{R_k L}{M} \frac{1 - \Pr[i_k(\mathbf{Q}_k, \mathbf{Q}_{-k}) \leq R_k]}{\text{Tr}(\mathbf{Q}_k)}$ is the utility function for transmitter k , where R_k is the transmission rate target for user k , the standard notation $-k$

stands for the set of players $\mathcal{K} \setminus \{k\}$ and

$$i_k(\mathbf{Q}_k, \mathbf{Q}_{-k}) = \log_2 \left| \mathbf{I} + \frac{\rho}{n_t} \mathbf{H}_k \mathbf{Q}_k \mathbf{H}_k^H + \frac{\rho}{n_t} \sum_{\ell \neq k} \mathbf{H}_\ell \mathbf{Q}_\ell \mathbf{H}_\ell^H \right| - \log_2 \left| \mathbf{I} + \frac{\rho}{n_t} \sum_{\ell \neq k} \mathbf{H}_\ell \mathbf{Q}_\ell \mathbf{H}_\ell^H \right|. \quad (47)$$

In this definition the function i_k represents the mutual information between the vector of symbols transmitted by user k i.e., \underline{x}_k and the received vector \underline{y} . For the set of strategies we will focus our attention on the case $\mathbf{Q}_k = p_k \mathbf{I}$ for the reasons we have already mentioned and consider the general case as an extension of this work. The problem under consideration then becomes a power control.

The game \mathcal{G} describes a conflicting situation of interaction between selfish players. An important question is to know whether this situation has a predictable issue. A well-known solution concept for this type of situations is the Nash equilibrium. In our context it corresponds to a state of the network which is stable to one deviation; every user has no incentives to unilaterally deviate from the equilibrium because otherwise it would lose in terms of utility. For making this paper sufficiently self-contained we review the definition of a Nash equilibrium [27].

DEFINITION 2 (NASH EQUILIBRIUM). *The strategy profile*

$$(\mathbf{Q}_1^{\text{NE}}, \dots, \mathbf{Q}_K^{\text{NE}}) \text{ is a pure Nash equilibrium if } \forall k \in \mathcal{K}, \forall \mathbf{Q}'_k \in \mathcal{A}_k, u_k(\mathbf{Q}_k^{\text{NE}}, \mathbf{Q}_{-k}^{\text{NE}}) \geq u_k(\mathbf{Q}'_k, \mathbf{Q}_{-k}^{\text{NE}}). \quad (48)$$

It turns out that there is at least one pure Nash equilibrium in \mathcal{G} where transmitters spread their power uniformly over their antennas, which is stated under a theorem form as follows.

THEOREM 4.1 (EXISTENCE OF AN NE). *Assume $\mathbf{Q}_k = p_k \mathbf{I}$ and $u_k(p_1, \dots, p_K)$ is a quasi-concave in p_k . Then, there exists at least one NE in the power control game \mathcal{G} .*

PROOF. The proof relies on an existing fixed-point theorem. It is the the Debreu-Fan-Glicksberg existence theorem stated properly in [20] and resulting from the contributions of [21], [22], [23]. This theorem is as follows.

THEOREM 4.2 (DEBREU-FAN-GLICKSBERG). *Consider a strategic-form non-cooperative game. If the following conditions are satisfied: (i) for every player, his utility function is continuous in the profile of strategies; (ii) for every player, the utility function is quasi-concave in its individual strategy; (iii) for every player, its strategy set is compact and convex; then the game has at least one pure Nash equilibrium.*

By assumption we know that (ii) is verified. Note that in this paper we have proved that this assumption is in fact a property for u_k in several useful special case and the authors conjecture that this property is general (Conjecture 2.4). Also the third condition (iii) is straightforward since $\mathcal{S}_k = [0, \bar{P}_k]$. We therefore just need to prove (i). For this, we start by proving that $f_k(p_1, \dots, p_K) = 1 - \Pr[i_k(p_1, \dots, p_K) \leq R_k]$ is continuous w.r.t. (p_1, \dots, p_K) . We know that $i_k(p_1, \dots, p_K)$ is a continuous function w.r.t.

p_ℓ and also w.r.t. $\mathbf{H}_\ell, \forall \ell \in \mathcal{K}$. We also know that i_k is strictly increasing function w.r.t. p_k and thus invertible w.r.t. p_k we have that $\Pr[i_k(p_k) \leq R_k] = 1 - \Pr[i_k^{-1}(R_k) \leq p_k]$ is a continuous function of p_k (the c.d.f. of a continuous random variable is continuous). We can use the same reasoning to prove the continuity w.r.t. p_ℓ for all $\ell \neq k$ knowing that i_k is strictly decreasing function of p_ℓ for all $\ell \neq k$. \square

5. NUMERICAL RESULTS

Uniform power allocation. First, we assume $K = 1$ and a uniform power allocation policy at the transmitter. In Fig. 1 we plot the energy efficiency function versus the transmit power in the scenario: $n_r = n_t = n \in \{1, 2, 4, 8\}$, $R = 1$ bpcu, $\rho = 10$ dB, $\bar{P} = 1$ W. We see that $F_{\text{slow}}(p, R)$ is quasi-concave w.r.t. $p \in [0, \frac{\bar{P}}{n}]$ and that a non-trivial maximum power $p^* > 0$ exists. We also can observe two interesting things. Saturating the transmit power constraint can be very suboptimal even with single-antenna terminals. For example, if the available transmit power is 1000 mW, saturating the power constraint leads to transmit 10 times the optimal value p^* and energy-efficiency is divided by more than 3. We also see the benefits brought by multiple antennas for saving energy. For example, using 8 antennas at each terminal allows to multiply by 8 energy efficiency while guaranteeing a minimum transmission rate of 1 bit per channel use (bpcu).

Optimal power allocation vs. Beam-forming power allocation vs. Uniform power allocation. Until now we have assumed uniform power allocation over the transmit antennas. Now we want to assess the performance gap between the uniform power allocation policy and the optimal policy, which can be found exhaustively for 2×1 MIMO systems. We also look at the case where $n_r = 1$ and $n_t = 2$ because analytical results on the optimal power allocation for minimizing the outage probability are available [13]: the result for the outage probability minimization is that in the low power regime the beam-forming power allocation is optimal and in the high power regime the uniform power allocation is optimal. The authors give the closed-form expression of the boundary that separates the two regimes: $\frac{2(2^R - 1)}{\rho \eta_0}$, where $\eta_0 = 1.2564$. In Fig. 3, we plot the success probability $(1 - P_{\text{out}})$ as a function of $\bar{P} \in [0, 10]$ W for the scenario: $R = 3$ bpcu, $\rho = 10$ dB. We observe that the analysis in [13] is also valid in our (more complex) case. Fig. 4 represents for the same scenario the energy efficiency function w.r.t. \bar{P} . The crucial difference in the two optimization problems is that when optimizing the outage probability all of the available power is used either with the uniform, beam-forming or optimal scheme which is no longer the case when optimizing the energy efficiency function in all three power allocation policies. What is mostly remarkable is the fact that the same analysis can be conducted regarding the energy-efficiency problem. There is a boundary (exactly the same as before) such that if \bar{P} is below then the beam-forming policy is optimal and otherwise the uniform policy is optimal.

In conclusion, we can conjecture the important result that the uniform power allocation optimizes the energy efficiency function provided that sufficient transmit power is available at the transmitter.

6. CONCLUSION

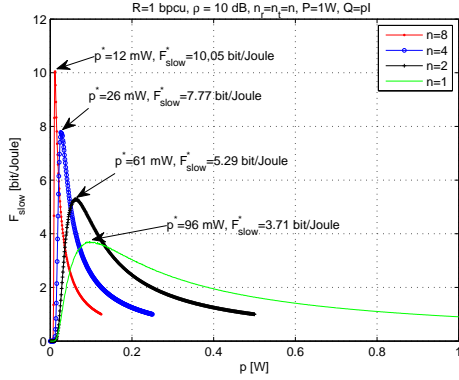


Figure 1: Uniform PA, $n_r = n_t = n$, $R = 1$ bpcu, $\rho = 10$ dB, $\bar{P} = 1$ W. For $n \in \{1, 2, 4, 8\}$ the energy efficiency function is quasi-concave w.r.t. $p \in [0, \frac{\bar{P}}{n}]$ and has a non trivial maximum point.

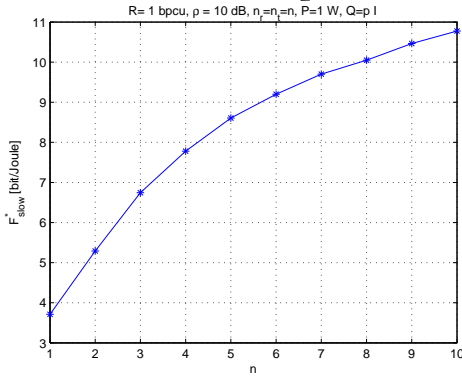


Figure 2: Uniform PA $n_r = n_t = n$, $R = 1$ bpcu, $\rho = 10$ dB, $\bar{P} = 1$ W. The energy efficiency at the maximum point is an increasing function of the number of antennas $n \in [1, 10]$.

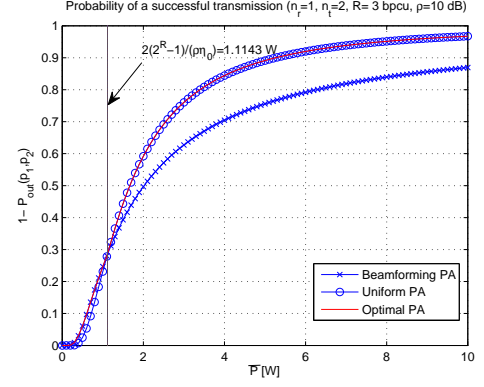


Figure 3: Optimal PA vs. Uniform PA vs. Beamforming PA, $n_r = 1$, $n_t = 2$, $R = 3$ bpcu, $\rho = 10$ dB, $\bar{P} \in [0, 10]$ W. The optimal PA that minimizes the outage probability is the Uniform PA in the high available power regime ($\bar{P} \geq \frac{2(2^R-1)}{\rho\eta_0} = 1.1143$ W) and the Beam-forming PA in the low available power regime.

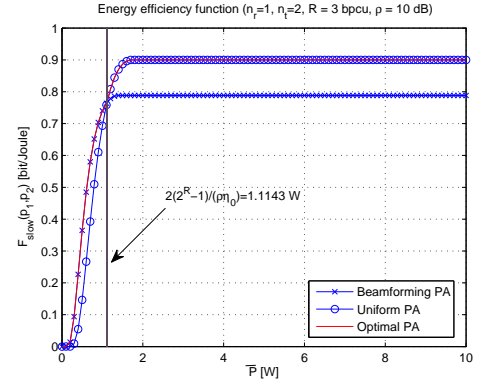


Figure 4: Optimal PA vs. Uniform PA vs. Beamforming PA, $n_r = 1$, $n_t = 2$, $R = 3$ bpcu, $\rho = 10$ dB, $\bar{P} \in [0, 10]$ W. The optimal PA that minimizes the energy efficiency function is the Uniform PA in the high available power regime ($\bar{P} \geq \frac{2(2^R-1)}{\rho\eta_0} = 1.1143$ W) and the Beam-forming PA in the low available power regime.

We have provided for different channel models an information-theoretic answer to the problem of energy-efficient communications. The most interesting scenario is that of quasi-static channels. Remarkably, the energy-efficiency function we introduce has all the nice properties stated by [15]. This shed a new light on [15] since, so far, only empirical efficiency-functions were used to analyze the problem of energy-efficient communications or power control. This also ensures the existence of a Nash equilibrium in MIMO multiple access channels implementing good coding schemes. In the case of MIMO single-user channels we have introduced a new open problem which is to determine the precoding matrix (and therefore the power allocation policy) maximizing the proposed energy-efficiency measure. We have solved this problem in several special and useful cases and conjectured the best precoding scheme. Simulations validates our conjectures. Many extensions of this work are possible. Here are some of them: 1. The most challenging one mathematically: solving the proposed open problem; 2. The most interesting one in terms of channel modeling: consider the case of arbitrary channel transfer matrices (Rice component, antenna correlations, etc); 3. Analyze the high SNR scenario by using the famous concept of diversity-multiplexing tradeoff; 4. From the energy standpoint: refine our energy-consumption analysis by including the consumption of the radio-frequency part of the transmitter e.g., the circuits associated with the transmit antennas.

7. REFERENCES

- [1] H. Kremling, "Making mobile broadband networks a success - Operator requirements", *Next Generation Mobile Networks Conference (NGMN)*, March 2008.
- [2] E. Telatar, "Capacity of multiantenna Gaussian channels", *AT&T Bell Laboratories Tech. Memo.*, June 1995.
- [3] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas", *Wireless Personal Commun.*, Kluwer Academic Publishers, vol. 6, pp. 311–335, March 1998.
- [4] E. Telatar, "Capacity of multi-antenna gaussian channels", *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–596, Nov./Dec. 1999.
- [5] V. Shah, N. B. Mandayam and D. J. Goodman, "Power control for wireless data based on utility and pricing", *IEEE Proc. of the 9th Intl. Symp. Personal, Indoor, Mobile Radio Communications (PIMRC)*, Boston, MA, pp. 1427–1432, Sep. 1998.
- [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] C. U. Saraydar, N. B. Mandayam and D. J. Goodman, "Efficient power control via pricing in wireless data networks", *IEEE Trans. on Communications*, vol. 50, no. 2, pp. 291–303, Feb. 2002.
- [9] F. Meshkati, H. Poor and S. Schwartz, "Energy-Efficient Resource Allocation in Wireless Networks", *IEEE Signal Processing Magazine*, vol. 58, pp. 58–68, May 2007.
- [10] N. Bonneau, M. Debbah, E. Altman and A. Hjørungnes, "Non-Atomic Games for Multi-User System", *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 7, pp. 1047–1058, Sep. 2008.
- [11] S. Lasaulce, Y. Hayel, R. El Azouzi and M. Debbah, "Introducing hierarchy in energy games", *IEEE Trans. on Wireless Communications*, vol. 8, no. 7, pp. 3833–3843, Jul. 2009.
- [12] L. Zheng, D.N. Tse, "Optimal Diversity-Multiplexing Tradeoff in Multiple Antenna Channels", *Proc. Allerton Conf. Comm., Control, Computing, Monticello*, pp. 835–844, Oct. 2001.
- [13] M. Katz, and S. Shamai, "On the outage probability of a multiple-input single-output communication link", *IEEE Trans. on Wireless Comm.*, vol. 6, no. 11, pp. 4120–4128, Nov. 2007.
- [14] L. H. Ozarow, S. Shamai, and A. D. Wyner, "Information theoretic considerations for cellular mobile radio", *IEEE Trans. on Vehicular Technology*, vol. 43, no. 2, pp. 359–378, May 1994.
- [15] V. Rodriguez, "An Analytical Foundation for Resource Management in Wireless Communication", *IEEE Proc. of Globecom*, 2003.
- [16] 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.
- [17] 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.
- [18] 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.
- [19] J. F. Nash, "Equilibrium points in n-points games", *Proc. of the National Academy of Science*, vol. 36, no. 1, pp. 48–49, Jan. 1950.
- [20] D. Fudenberg and J. Tirole, "Game Theory", *MIT Press*, 1991.
- [21] G. Debreu, "A social equilibrium existence theorem", *Proc. of the National Academy of Science*, vol. 38, pp. 886–893, 1952.
- [22] K. Fan, "Fixed point and minima theorems in locally convex topological linear spaces", *Proc. of the National Academy of Sciences*, vol. 38, pp. 121–126, 1952.
- [23] I. L. Glicksberg, "A further generalization of the Kakutani fixed point theorem with application to Nash equilibrium points", *Proc. of the American Mathematical Society*, vol. 3, pp. 170–174, 1952.
- [24] Y. Zhu, P. Y. Kam, and Y. Xin, "A new approach to the capacity distribution of MIMO Rayleigh Fading Channels", *IEEE Global Telecommunications Conference (GLOBECOM)*, San Francisco, CA, USA, Dec. 2006.
- [25] B. M. Hochwald, T. L. Marzetta, and V. Tarokh, "Multiple-antenna channel hardening and its implications for rate feedback and scheduling", *IEEE Trans. on Inform. Theory*, vol. 50, no. 9, pp. 1893–1909, Sep. 2004.
- [26] M. Debbah, and R. R. Muller, "MIMO channel modeling and the principle of maximum entropy", *IEEE Trans. on Inform. Theory*, vol. 51, no. 5, pp. 1667–1690, May 2005.
- [27] J. F. Nash, "Equilibrium points in n-person games", *Proc. of the Nat. Academy of Science*, vol. 36, no. 1, pp. 48–49, 1950.
- [28] E. V. Belmega, S. Lasaulce and M. Debbah, "Power allocation games for MIMO multiple access channels with coordination", *IEEE Trans. on Wireless Communications*, vol. 8, no. 6, pp. 3182–3192, Jun. 2009.