

An Energy Efficient Framework for the Analysis of MIMO Slow Fading Channels

Vineeth Varma, Samson Lasaulce, Merouane Debbah, Salah Eddine Elayoubi

▶ **To cite this version:**

Vineeth Varma, Samson Lasaulce, Merouane Debbah, Salah Eddine Elayoubi. An Energy Efficient Framework for the Analysis of MIMO Slow Fading Channels. IEEE Transactions on Signal Processing, 2013, 61 (10), pp.2647-2659. 10.1109/TSP.2013.2252170 . hal-00822055

HAL Id: hal-00822055

https://centralesupelec.hal.science/hal-00822055

Submitted on 13 May 2013

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 Energy-Efficient Framework for the Analysis of MIMO Slow Fading Channels

Vineeth S. Varma, Samson Lasaulce, Merouane Debbah, and Salah Eddine Elayoubi,

***Abstract*—In this work, a new energy-efficiency performance metric is proposed for MIMO (multiple input multiple output) point-to-point systems. In contrast with related works on energy-efficiency, this metric translates the effects of using finite blocks for transmitting, using channel estimates at the transmitter and receiver, and considering the total power consumed by the transmitter instead of the radiated power only. The main objective pursued is to choose the best pre-coding matrix used at the transmitter in the following two scenarios : 1) the one where imperfect channel state information (CSI) is available at the transmitter and receiver ; 2) the one where no CSI is available at the transmitter. In both scenarios, the problem of optimally tuning the total used power is shown to be non-trivial. In scenario 2), the optimal fraction of training time can be characterized by a simple equation. These results and others provided in the paper, along with the provided numerical analysis, show that the present work can therefore be used as a good basis for studying power control and resource allocation in energy-efficient multiuser networks.**

***Index Terms*—Channel training, energy efficiency, finite block length, green communication, imperfect channel state information, MIMO.**

I. Introduction

Over the past two decades, designing energy-efficient communication terminals has become an important issue. This is not surprising for terminals which have to be autonomous as far as energy is concerned, such as cellular phones, unplugged laptops, wireless sensors, and mobile robots. More surprisingly, energy consumption has also become a critical issue for the fixed infrastructure of wireless networks. For instance, Vodafone's global energy consumption for 2007-2008 was about 3000 GWh [1], which corresponds to emitting 1.45 million tons of CO2 and represents a monetary cost of a few hundred million Euros. This context explains, in part, why concepts like "green communications" have emerged as seen from [2], [3] and [4]. Using large multiple antennas, virtual multiple input multiple output (MIMO) systems, and small cells is envisioned to be one way of contributing to reducing energy consumption drastically. The work reported in this paper concerns point-to-point MIMO systems in which communication links evolve in a quasi-static manner, these channels are referred to as MIMO slow fading channels. The performance metric considered for measuring energy-efficiency of a MIMO communication corresponds to a trade-off between the net transmission rate (transmission benefit) and the consumed power (transmission cost).

The ultimate goal pursued in this paper is a relatively important problem in signal processing for communications. It consists of tuning the covariance matrix of the transmitted signal (called the pre-coding matrix) optimally. But, in contrast with the vast literature initiated by [5] in which the transmission rate is of prime interest, the present paper aims at optimizing the pre-coding matrix in the sense of energy-efficiency as stated in [6]. Interestingly, in [6] the authors bridge a gap between the pioneering work by Verdú on the capacity per unit cost for static channels [7] and the more pragmatic definition of energy-efficiency proposed by [8] for quasi-static single input single output (SISO) channels. Indeed, in [6], energy-efficiency is defined as the ratio of the probability that the channel mutual information is greater than a given threshold to the used transmit power. Assuming perfect channel state information at the receiver (CSIR) and the knowledge of the channel distribution at the transmitter, the pre-coding matrix is then optimized for several special cases. While [6] provides interesting insights into how to allocate and control power at the transmitter, a critical issue is left unanswered; to what extent do the conclusions of [6] hold in more practical scenarios such as those involving imperfect CSI? Answering this question was one of the motivations for the work reported here. Below, the main differences between the approach used in this work and several existing relevant works are reviewed.

In the proposed approach, the goal pursued is to maximize the number of information bits transmitted successfully per Joule consumed at the transmitter. This is different from the most conventional approach which consists in minimizing the transmit power under a transmission rate constraint: [9] perfectly represents this body of literature. In the latter and related works, efficiency is not the main motivation. [10] provides a good motivation as to how energy-efficiency can be more relevant than minimizing power under a rate constraint. Indeed, in a communication system without delay constraints, rate constraints are generally irrelevant whereas the way energy is used to transmit the (sporadic) packets is of prime interest. Rather, our approach follows the original works on energy-efficiency which includes [11], [8], [12], [13], [14]. The current state of the art indicates that, since [6], there have been no works where the MIMO case is treated by exploiting the cumulative distribution of the channel mutual information (i.e., the outage probability) at the numerator of the performance metric. As explained below, our analysis goes much further than [6] by considering effects such as channel estimation error

V. S. Varma and S. E. Elayoubi are with the Orange Labs, 92130 Issy Les Moulineaux, France (e-mail: vineethsvarma@gmail.com; salaheddine.elayoubi@orange.com).

S. Lasaulce and M. Debbah are with SUPELEC, 91192 Gif-sur-Yvette, France (e-mail: Samson.LASAULCE@lss.supelec.fr; merouane.debbah@supelec.fr).

effects. In the latter respect, several works address the issue of power allocation for outage probability minimization [15], [16], [17] under imperfect channel state information. The latter will serve as a basis for the analysis conducted in the present paper. At this point, it is possible to state the contributions of the present work.

In comparison to [6], which is the closest related work, the main contributions of the paper can be summarized as follows:

- one of the scenarios under investigation concerns he case where CSI is also available at the transmitter (only the case with CSIR and CSI distribution at the transmitter is studied in [6]).
- The assumption of perfect CSI is relaxed. In Sec. III, it is assumed that only imperfect CSIT and imperfect CSIR is available. Sec. IV considers the case with no CSIT and imperfect CSIR. In particular, this leads us to the problem of tuning the fraction of training time optimally. Exploiting existing works for the transmission rate analysis [18] and [19], it is shown that this problem can also be treated for energy-efficiency.
- The realistic assumption of finite block length is made. This is particularly relevant, since block finiteness is also a source of outage and therefore impacts energy-efficiency. Note that recent works on transmission under the finite length regime such as [20] provide a powerful theoretical background for possible extensions of this paper.
- Instead of considering the radiated power only for the cost of transmitting, the total power consumed by the transmitter is accounted for. Based on works such as [21], an affine relation between the two is assumed. Although more advanced models can be assumed, this change is sufficient to show that the behavior of energy-efficiency is also modified.

The paper is therefore structured as follows. Sec. II describes the proposed framework to tackle the aforementioned issues. Sec. III and IV treat the case with and without CSIT respectively. They are followed by a section dedicated to numerical results (Sec. V) whereas Sec. VI concludes the paper with the main messages of this paper and some relevant extensions.

II. System model

A point-to-point multiple input and multiple output communication unit is studied in this work. In this paper, the dimensionality of the input and output is given by the numbers of antennas but the analysis holds for other scenarios such as virtual MIMO systems [22]. If the total transmit power is given as P, the average SNR is given by :

$$\rho = \frac{P}{\sigma^2} \tag{1}$$

where σ^2 is the reception noise variance.The signal at the receiver is modeled by :

$$\underline{y} = \sqrt{\frac{\rho}{M}}\mathbf{H}\underline{s} + \underline{z} \tag{2}$$

where $\mathbf{H}$ is the $N \times M$ channel transfer matrix and M (resp. N) the number of transmit (resp. receive) antennas. The entries of $\mathbf{H}$ are i.i.d. zero-mean unit-variance complex Gaussian random variables. The vector $\underline{s}$ is the M-dimensional column vector of transmitted symbols follows a complex normal distribution, and $\underline{z}$ is an N-dimensional complex white Gaussian noise distributed as $\mathcal{N}(\underline{0}, \mathbf{I})$. Denoting by $\mathbf{Q} = \mathbb{E}[\underline{s}\underline{s}^H]$ the input covariance matrix (called the pre-coding matrix), which satisfies

$$\frac{1}{M}\mathrm{Tr}(\mathbf{Q}) = 1 \tag{3}$$

where Tr stands for the trace operator. The power constraint is expressed as :

$$P \leq P_{\max} \tag{4}$$

where $P_{\max}$ is the maximum available power at the transmitter.

The channel matrix $\mathbf{H}$ is assumed to evolve in a quasi-static manner : the channel is constant for some time interval, after which it changes to an independent value that it holds for the next interval [18]. This model is appropriate for the slow-fading case where the time with which $\mathbf{H}$ changes is much larger than the symbol duration.

A. Defining the energy efficiency metric

In this section, we introduce and justify the proposed definition of energy-efficiency of a communication system with multiple input and output antennas, and experiences slow fading.

In [8], the authors study multiple access channels with SISO links and use the properties of the energy efficiency function defined as $\frac{f(\rho)}{P}$ to establish a relation between the channel state (channel complex gain) (h) and the optimal power (P^*). This can be written as:

$$P^* = \frac{\mathrm{SNR}^*\sigma^2}{|h|^2} \tag{5}$$

where SNR^* is the optimal SNR for any channel state and (when f is a sigmoidal/S-shaped function, i.e, it is initially convex and after some point becomes concave) is the unique strictly positive solution of

$$xf'(x) - f(x) = 0 \tag{6}$$

where $1 - f(.)$ is the outage probability. Formulating this problem in the case of MIMO channels is non-trivial as there is a problem of choosing the total transmit power as well as the power allocation.

When the same (imperfect) CSI is available at the transmitter and receiver, by estimating the channel for t time, and sending the information to the transmitter for t_f time, the energy-efficiency ν_T is defined as:

$$\nu_T(P, \mathbf{Q}, \hat{\mathbf{H}}) = \frac{R\left(1 - \frac{t+t_f}{T}\right) F_L\left[I_{\mathrm{ICSITR}}(P, \mathbf{Q}, \hat{\mathbf{H}}) - \frac{R}{R_0}\right]}{aP + b} \tag{7}$$

where R is the transmission rate in bit/s, T is the block duration in s, R_0 is a parameter which has unit Hz (e.g., the system bandwidth), and $a > 0$, $b \geq 0$ are parameters to relate the transmitter radiated power to its total consumed power ;

we define $\xi = \frac{R}{R_0}$ as the spectral efficiency. $I_{\mathrm{ICSITR}}(P, \mathbf{Q}, \hat{\mathbf{H}})$ denotes the mutual information with imperfect CSITR (the receiver also has the exact same CSI as the transmitter). This form of the energy-efficiency is inspired from early definitions provided in works like [8], and studies the gain in data rate with respect to the cost which is the power consumed. The numerator represents the benefit associated with transmitting namely, the net transmission rate (called the goodput in [23]) of the communication and is measured in bit/s. The goodput comprises a term $1-\frac{t+t_f}{T}$ which represents the loss in terms of information rate due to the presence of a training and feedback mechanism (for duration t seconds and t_f seconds resp. in a Ts long block) [1]. The denominator of (7) represents the cost of transmission in terms of power. The proposed form for the denominator of (7) is inspired from [21] where the authors propose to relate the average power consumption of a transmitter (base stations in their case),to the average radiated or radio-frequency power by an affine model.

The term $F_L(.)$ represents the transmission success probability. The quantity $F_L(.)$ gives the probability that the "information" denoted by $\hat{I}$ as defined in [24]) is greater than or equal to the coding rate (ξ), i.e., it is the complementary cumulative distribution function of the information $\hat{I}$, $\mathrm{Prob}(\hat{I} \geq \xi)$. Formally, $\hat{I}$ is defined as $\hat{I} = \log \frac{\mathrm{PDF}_{X,Y}(x,y)}{\mathrm{PDF}_X(x)\mathrm{PDF}_Y(y)}$, where $\mathrm{PDF}_{X,Y}$ and PDF_X represents the joint and marginal probability distribution functions, x and y are samples of the process X and Y, which in this case represent the transmitted and received signals. The average mutual information $I = E(\hat{I})$ is used to calculate this probability and $F_L(.)$ depends on the difference between I and ξ. $F_L(.)$ can be verified to be sigmoidal (this is the cumulative probability distribution function of a variable with a single peaked probability distribution function) and $F_L(0) = 0.5$ (If $\xi = I$, $F_L(.)$ is the probability that a random variable is equal to or larger than its mean). When CSIT is available, it is possible to ensure that the data transmission rate is just below the channel capacity. If this is done, then there is no possibility of outage when the block length is infinite [25]. However, in most practical cases, the block length is finite and this creates an outage effect which depends on the block length L [24].

The bounds on F_L can be expressed as $F_L(I_{\mathrm{ICSITR}}(0, 0, \mathbf{H}) - \xi) = 0$ (no reliable communication when transmit power is zero) and as $F_L \rightarrow 1$ when $P \rightarrow \infty$. This proposed form for this function, $F_L(I_{\mathrm{ICSITR}}(P, \mathbf{Q}, \hat{\mathbf{H}}) - \xi)$, is supported by works like [24] and [26]. An approximation for this function based on the automatic repeat request protocol [27] is $F_L(x) = Q_{func}(-Tx)$, where Q_{func} is the tail probability of the standard normal distribution.

Therefore, in the presence of CSI at the transmitter, outage occurs even when the mutual information is more than the targeted rate due to the noise and finite code-lengths. In this scenario, the energy-efficiency is maximized when the parameters $\mathbf{Q}$ and P are optimized.

[1]In this case, we assume that the feedback mechanism is sufficient to result in perfect knowledge of $\hat{\mathbf{H}}$ at the transmitter. This is done because, assuming a different imperfect CSI at the transmitter from the receiver creates too much complexity and this problem is beyond the scope of a single paper.

In the absence of CSI at the transmitter, the earlier definition of energy efficiency is not suitable since $\mathbf{H}$ is random, ν_T is also a random quantity. Additionally, in this case, it is impossible to know if the data transmission rate is lower than the instantaneous channel capacity as the channel varies from block to block. Therefore, in this case, the source of outage is primarily the variation of the channel [28], and using (7) directly is not suitable. As the channel information is unavailable at the transmitter, define $\mathbf{Q} = \frac{\mathbf{I}_M}{M}$, meaning that the transmit power is allocated uniformly over the transmit antennas; in Sec. IV-C, we will comment more on this assumption. Under this assumption, the average energy-efficiency can be calculated as the expectation of the instantaneous energy-efficiency over all possible channel realizations. This can be rewritten as:

$$\nu_R(P,t) = \frac{R\left(1-\frac{t}{T}\right)\mathbb{E}_{\mathbf{H}}\left(F_L\left[I_{\mathrm{ICSIR}}(P,\mathbf{Q},\hat{\mathbf{H}}) - \frac{R}{R_0}\right]\right)}{aP+b}. \tag{8}$$

For large L, it has been shown in [28] (and later used in other works like [6]) that the above equation can be well approximated to :

$$\nu_R(P,t) = \frac{R\left(1-\frac{t}{T}\right)\mathrm{Pr}_{\mathbf{H}}\left[I_{\mathrm{ICSIR}}(P,t,\hat{\mathbf{H}}) \geq \xi\right]}{aP+b} \tag{9}$$

where $\mathrm{Pr}_{\mathbf{H}}$ represents the probability evaluated over the realizations of the random variable $\mathbf{H}$. Here, I_{ICSIR} represents the mutual information of the channel with imperfect CSI at the receiver. Let us comment on this definition of energy efficiency. This definition is similar to the earlier definition in all most ways. Here the parameter t, represents the length of the training sequence used to learn the channel at the receiver[2]. The major difference here is that the expression for the success rate is the probability that the associated mutual information is above a certain threshold. This definition of the outage is shown to be appropriate and compatible with the earlier definition when only statistical knowledge of the channel is available [28].

Although very simple, these models allow one, in particular, to study two regimes of interest.

- The regime where $\frac{b}{a}$ is small allows one to study not only communication systems where the power consumed by the transmitter is determined by the radiated power but also those which have to been green in terms of electromagnetic pollution or due to wireless signal restrictions (see e.g., [29]).
- The regime where $\frac{b}{a}$ is large allows one to study not only communication systems where the consumed power is almost independent of the radiated power but also those where the performance criterion is the goodput.

Note that when $b = 0$, $t \rightarrow +\infty$, $T \rightarrow +\infty$, and $\frac{t}{T} \rightarrow 0$ equation (9) boils down to the performance metric investigated in [6].

[2]In this case, the optimization is done over P and t assuming imperfect CSI at the receiver. A parameter here not explicitly stated, but indicated nevertheless, is M due to the number of transmit antennas affecting the effectiveness of training

B. Modeling channel estimation noise

Each transmitted block of data is assumed to comprise a training sequence in order for the receiver to be able to estimate the channel; the training sequence length in symbols is denoted by t_s and the block length in symbols by T_s. Continuous counterparts of the latter quantities are defined by $t = t_s S_d$ and $T = T_s S_d$, where S_d is the symbol duration in seconds. In the training phase, all M transmitting antennas broadcast orthogonal sequences of known pilot/training symbols of equal power on all antennas. The receiver estimates the channel, based on the observation of the training sequence, as $\widehat{\mathbf{H}}$ and the error in estimation is given as $\Delta\mathbf{H} = \mathbf{H} - \widehat{\mathbf{H}}$. Concerning the number of observations needed to estimate the channel, note that typical channel estimators generally require at least as many measurements as unknowns [19], that is to say $Nt_s \geq NM$ or more simply

$$t_s \geq M. \tag{10}$$

The channel estimate normalized to unit variance is denoted by $\widetilde{\mathbf{H}}$. From [19] we know that the mutual information is the lowest when the estimation noise is Gaussian. Taking the worst case noise, it has been shown in [18] that the following observation equation

$$\underline{\widetilde{y}} = \sqrt{\frac{\rho_{\text{eff}}(\rho,t)}{M}}\widetilde{\mathbf{H}}\underline{s} + \underline{\widetilde{z}} \tag{11}$$

perfectly translates the loss in terms of mutual information[3] due to channel estimation provided that the effective SNR $\rho_{\text{eff}}(\rho,t)$ and equivalent observation noise $\underline{\widetilde{z}}$ are defined properly namely,

$$\begin{cases} \underline{\widetilde{z}} & = & \sqrt{\frac{\rho}{M}}\Delta\mathbf{H}\underline{s} + \underline{z} \\ \rho_{\text{eff}}(\rho,t) & = & \frac{\frac{t}{MS_d}\rho^2}{1+\rho+\rho\frac{t}{MS_d}} \end{cases}. \tag{12}$$

As the worst case scenario for the estimation noise is assumed, all formulas derived in the following sections give lower bounds on the mutual information and success rates. Note that the lower bound is tight (in fact, the lower bound is equal to the actual mutual information) when the estimation noise is Gaussian which is true in practical cases of channel estimation. The effectiveness of this model will not be discussed here but has been confirmed in many other works of practical interest (see e.g., [31]). Note that the above equation can be utilized for the cases of imperfect CSITR and CSIR as well as the case of imperfect CSIR with no CSITR. This is because in both cases, the outage is determined by calculating the mutual information I_{ICSITR} or I_{ICSIR} respectively.

III. Optimizing energy-efficiency with imperfect CSITR available

When perfect CSITR or CSIR is available, the mutual information of a MIMO system, with a pre-coding scheme $\mathbf{Q}$ and channel matrix $\mathbf{H}$ can be expressed as:

$$I_{\text{CSITR}}(P,\mathbf{Q},\mathbf{H}) = \log\left|\mathbf{I}_M + \frac{P}{M\sigma^2}\mathbf{H}\mathbf{Q}\mathbf{H}^H\right| \tag{13}$$

The notation $|\mathbf{A}|$ denotes the determinant of the (square) matrix $\mathbf{A}$. With imperfect CSIT, which is exactly the same as the CSIR (i.e., both the transmitter and the receiver have the same channel estimate $\hat{\mathbf{H}}$), a lower bound on the mutual information can be found from several works like [15], [17] etc. This lower bound for I_{ICSITR} is used, which is expressed as:

$$I_{\text{ICSITR}}(P,\mathbf{Q},\hat{\mathbf{H}}) = \log\left|\mathbf{I}_M + \hat{\mathbf{H}}\frac{P}{M\sigma^2(1+\rho\sigma_E^2)}\mathbf{Q}\hat{\mathbf{H}}^H\right| \tag{14}$$

where $\hat{\mathbf{H}}$ is the estimated channel and $1-\sigma_E^2$ is the variance of $\hat{\mathbf{H}}$. Considering the block fading channel model, from [15] and [19] we conclude that $\sigma_E^2 = \frac{1}{1+\rho\frac{t}{M}}$. Simplifying :

$$I_{\text{ICSITR}}(P,\mathbf{Q},\widehat{\mathbf{H}}) = \log\left|\mathbf{I}_M + \frac{\rho_{\text{eff}}}{M}\widehat{\mathbf{H}}\mathbf{Q}\widehat{\mathbf{H}}^H\right|. \tag{15}$$

Having defined the mutual information to be used for (7), we proceed with optimizing ν_T.

A. Optimizing the pre-coding matrix $\mathbf{Q}$

Studying (7) and (15), we see that varying the power allocation (or the corresponding pre-coding matrix) $\mathbf{Q}$, affects only the success rate $F_L(.)$ and the total power P is the only term that is present outside $F_L(.)$. As $F_L(.)$ is known to be an increasing function, if the total power is a constant, optimizing the energy efficiency ν_T amounts to simply maximizing the mutual information $I_{\text{ICSITR}}(P,\mathbf{Q},\hat{\mathbf{H}})$. This is a well documented problem and it gives a "water-filling" type of solution [32]. Rewriting (13) as

$$I_{\text{ICSITR}}(P,\mathbf{Q},\widehat{\mathbf{H}}) = \log\left|\mathbf{I}_M + \frac{\rho_{\text{eff}}}{M}\mathbf{D}\mathbf{S}\mathbf{D}^H\right| \tag{16}$$

where the optimal covariance matrix $\mathbf{Q} = \mathbf{V}\mathbf{S}\mathbf{V}^H$ is achieved through the singular value decomposition of the channel matrix $\widehat{\mathbf{H}} = \mathbf{U}\mathbf{D}\mathbf{V}^H$ and an optimal diagonal covariance matrix $\mathbf{S} = \text{diag}[s_1,\ldots,s_{\min(M,N)},0,\ldots,0]$. The water-filling algorithm can be performed by solving:

$$s_i = \left(\mu - \frac{1}{\rho\|d_i\|^2}\right)^+, \text{ for } i = 1,2,\cdots,\min(M,N) \tag{17}$$

where d_i are the diagonal elements of $\mathbf{D}$ and μ is selected such that $\Sigma_{i=1}^{\min(M,N)} s_i = M$. Here $(x)^+ = \max(0,x)$, this implies that s_i can never be negative. The actual number of non-vanishing entries in $\mathbf{S}$ depends on the values of d_i as well ρ (and thus P). Examining (17), we can see that when $\rho \to 0$, the water-filling algorithm will lead to choosing $s_j = M$ and $s_i = 0$ for all $i \neq j$, where j is chosen such that $d_j = \max(d_i)$ (beamforming). Similarly for $\rho \to \infty$, $s_i = \frac{M}{\min(M,N)}$ (uniform power allocation).

[3]It is implicitly assumed that the mutual information is taken between the system input and output; this quantity is known to be very relevant to characterize the transmission quality of a communication system (see e.g. [30] for a definition).

B. Determining the optimal total power

$\mathbf{Q}$ has been optimized in the previous section. From (7), we see that the parameters that can be optimized in order to maximize the energy efficiency are $\mathbf{Q}$ and P. Therefore, in this section, we try to optimize P, the total power. Note that for every different P, the optimal power allocation $\mathbf{Q}$ changes according to (17) as ρ is directly proportional to P. Therefore optimizing this parameter is not a trivial exercise. Practically, P represents the total radio power, that is, the total power transmitted by the antennas. This power determines the total consumed power $b+aP$, of base stations or mobile terminals and so, optimizing this power is of great importance.

In this section, a theorem on the properties of $\nu_T(P, \mathbf{Q}_{WF(P)}, \widehat{\mathbf{H}})$ is provided, where $\mathbf{Q}_{WF(P)}$ is the power allocation obtained by using the water-filling algorithm and iteratively solving (17) with power P. This procedure is said to be "iterative" because, after solving equation 17, if any $s_j < 0$, then we set $s_j = 0$ and the equation is resolved until the all solutions are positive. For optimization, desirable properties on $\nu_T(P, \mathbf{Q}_{WF(P)}, \widehat{\mathbf{H}})$ are differentiability, quasi-concavity and the existence of a maximum. The following theorem states that these properties are in fact satisfied by ν_T.

Theorem 3.1: The energy-efficiency function $\nu_T(P, \mathbf{Q}_{WF(P)}, \widehat{\mathbf{H}})$ is quasi-concave with respect to P and has a unique maximum $\nu_T(P^*, \mathbf{Q}_{WF(P^*)}, \widehat{\mathbf{H}})$, where P^* satisfies the following equation :

$$\begin{aligned} &\tfrac{\partial F_L[I_{\mathrm{ICSITR}}(P^*, \mathbf{Q}_{WF(P^*)}, \widehat{\mathbf{H}}) - \xi]}{\partial P} \left(P^* + \tfrac{b}{a}\right) \\ &\quad - F_L[I_{\mathrm{ICSITR}}(P^*, \mathbf{Q}_{WF(P^*)}, \widehat{\mathbf{H}}) - \xi] \quad = 0 \end{aligned} \tag{18}$$

where $\frac{\partial}{\partial P}$ is the partial derivative.

The proof of this theorem can be found in Appendix A. From the above theorem and equation, we can conclude that the optimal transmit power for imperfect CSITR depends on several factors like

- the channel estimate $\widehat{\mathbf{H}}$,
- the target spectral efficiency ξ,
- the ratio of the constant power consumption to the radio-frequency (RF) power efficiency $\frac{b}{a}$,
- the channel training time t and
- the noise level σ^2.

Note that in this model, we always assume the CSI at the transmitter to be exactly identical to CSI at the receiver. Because of this, we take the feedback mechanism to be perfect and take a constant time t_f. Although in practice, t_f plays a role in determining the efficiency and the optimal power, in our model t_f is a constant and does not appear in the equation for P^*. In our numerical results we focus on the impact of $\widehat{\mathbf{H}}$, ξ and $\frac{b}{a}$ on P^* and ν^*. The impact of t is not considered for this case, but is instead studied where we have no CSITR and imperfect CSIR, this choice helps in making the results presented easier to interpret and understand.

C. An illustrative special case : SISO channels

A study on energy-efficiency in SISO systems have been studied in many works like [8] and [7]. However, the approach used in this paper is quite novel even for the SISO case and presents some interesting insights that have not been presented before. For the case of SISO, the pre-coding matrix is a scalar and $\mathbf{Q} = 1$. The optimal power can be determined by solving (18). For a SISO system with perfect CSITR and CSIR, F_L can be expressed as

$$\begin{aligned} &F_L[I_{\mathrm{ICSITR}}(P^*, \mathbf{Q}_{WF(P^*)}, \widehat{\mathbf{H}}) - \xi] = \\ &Q_{\mathrm{function}}\left(L(1+\|h\|^2\rho)\tfrac{\xi - \log(1+\|h\|^2\rho)}{\|h\|^2\rho}\right) \end{aligned} \tag{19}$$

from [24]. Using this expression, we can find P^* maximizing ν_T.

In the case of high SNR (and high ξ), a solution to this problem can be found as

$$\lim_{\rho\to\infty} F_L(P, 1, h) = Q_{func}\left(L[\xi - \log(1+\|h\|^2\rho])\right). \tag{20}$$

Solving (18)

$$\begin{aligned} &\tfrac{-L\|h\|^2}{\sqrt{\pi}(1+\|h\|^2\rho^*)} \exp\left(-L^2\left[\xi - \log(1+\|h\|^2\rho^*)\right]^2\right)\left(\rho^* + \tfrac{b}{a\sigma^2}\right) \\ &\quad - Q_{func}\left(L[\xi - \log(1+\|h\|^2\rho^*])\right) = 0 \end{aligned} \tag{21}$$

From the above equation it can be deduced that if $b = 0$, for large ξ, $\log(1+\rho^*) \approx \xi$. While for low SNR, $\lim_{\rho\to 0} F_L(P, 1, h) = Q_{func}\left(L\frac{\xi - \log(1+\|h\|^2\rho)}{\|h\|^2\rho}\right)$ and so, if $b = 0$,

$$\begin{aligned} &\tfrac{1}{\sqrt{\pi}}\left[L + L\tfrac{\xi - |h|^2\rho^*}{\|h\|^2\rho^*}\right] \exp\left(\tfrac{-1}{2}\left[L\tfrac{\xi - \|h\|^2\rho^*}{\|h\|^2\rho^*}\right]^2\right) \\ - &\quad Q_{func}\left(L\tfrac{\xi - \|h\|^2\rho^*}{\|h\|^2\rho^*}\right) = 0 \end{aligned} \tag{22}$$

Substitute $x = L\frac{\xi - \|h\|^2\rho^*}{\|h\|^2\rho^*}$ and we have

$$\frac{1}{\sqrt{\pi}}\left[L + x\right] \exp\left(\frac{-1}{2}x^2\right) - Q_{func}(x) = 0. \tag{23}$$

As seen from the above equation, the value of x depends only on L the block length. For example if $L = 10$, we get $x \approx -1.3$. So, $\rho^* = 1.14\frac{\xi}{\|h\|^2}$. Whereas if $L = 100$ we get $\rho^* = 1.02\frac{\xi}{\|h\|^2}$. Note that these calculations are true only for $\xi \to 0$ so that $\rho \to 0$ is satisfied.

The above equations signify that for finite block lengths, the energy efficiency at $\xi \to 0$ is lower than the value calculated in [7] (of course, a direct comparison does not make sense as in [7], infinite block lengths are assumed). This suggests that a non-zero value of ξ might optimize the energy efficiency. This value is evaluated in our numerical section and we find that the energy efficiency is optimized at a non-zero power.

D. Special Case: Infinite code-length and perfect CSITR

When a very large block is used then the achievable rate approaches the mutual information [25], i.e $\lim_{L\to\infty, I_{\mathrm{CSITR}}-\xi\to 0^+} F_L(I_{\mathrm{CSITR}} - \xi) = 1$. Therefore in this limit, we can now simplify (7) to:

$$\nu_T(P, \mathbf{Q}, \widehat{\mathbf{H}}) = \frac{R_0\left(1 - \frac{t+t_f}{T}\right) I_{\mathrm{CSITR}}(P, \mathbf{Q}, \widehat{\mathbf{H}})}{aP + b}. \tag{24}$$

This is done because we replace ξ with I_{CSITR} to maximize efficiency as F_L is 0 when $I_{\mathrm{CSITR}} < \xi$, and choosing

$\xi \to I_{\mathrm{CSITR}}$ maximizes efficiency. Water-filling optimizes the efficiency in this situation as well, and so we use $\mathbf{Q} = \mathbf{Q}_{WF(P)}$. It can be easily verified that for $b \to 0$: ν_T is maximized for $P \to 0$. And in this case, water-filling also implies that only the antenna with the best channel is used to transmit. Interestingly, when in the domain of finite code-lengths, our simulations indicate that there is a non-zero rate and power that optimizes the energy-efficiency function.

For general b, ν_T is optimized for P^* satisfying:

$$\frac{\partial I_{\mathrm{CSITR}}(P, \mathbf{Q}_{WF(P)}, \widehat{\mathbf{H}})}{\partial P}(aP + b) - I_{\mathrm{CSITR}}(P, \mathbf{Q}, \widehat{\mathbf{H}}) = 0. \tag{25}$$

The above equation admits a unique maximum because $I_{\mathrm{CSITR}}(P, \mathbf{Q}_{WF(P)}, \widehat{\mathbf{H}})$ is a concave function of P (can be seen from Appendix A) and is mathematically appealing. For $\lim_{b\to 0} P^* = 0$ and as $\frac{b}{a}$ increases, P^* also increases. A special case of this, with $b = 0$, and perfect CSITR, for a SISO channel has been studied in [7].

IV. Optimizing energy-efficiency with no CSIT and imperfect CSIR

This problem has already been well analyzed in [6] when perfect CSI is available at the receiver and $b = 0$. So, in this paper we focus on the case when imperfect CSI is available and is obtained through channel training. For $I_{ICSIR}(P, t, \mathbf{H})$, we use a lower bound on the mutual information obtained from the equivalent observation equation (11), derived in [19]:

$$I_{ICSIR}(P, t, \hat{\mathbf{H}}) = \log \left| \mathbf{I}_M + \frac{1}{M}\rho_{\mathrm{eff}} \left(\frac{LP}{\sigma^2}, t \right) \hat{\mathbf{H}}\hat{\mathbf{H}}^H \right| \tag{26}$$

Note that here, $Q = \frac{\mathbf{I}_M}{M}$ is used and has been shown to be optimal in [6]. In this section our focus is to generalize [6] to a more realistic scenario where the total power consumed by the transmitter (instead of the radiated power only) and imperfect channel knowledge are accounted for.

A. Optimal transmit power

By inspecting (9) and (26) we see that using all the available transmit power can be suboptimal. For instance, if the available power is large and all of it is used, then $\nu_R(P, t)$ tends to zero. Since $\nu_R(P, t)$ also tends to zero when P goes to zero (see [6]), there must be at least one maximum at which energy-efficiency is maximized, showing the importance of using the optimal fraction of the available power in certain regimes. The objective of this section is to study those aspects namely, to show that ν_R has a unique maximum for a fixed training time length and provide the equation determining the optimum value of the transmit power.

From [33] we know that a sufficient condition for the function $\frac{f(x)}{x}$ to have a unique maximum is that the function $f(x)$ be sigmoidal. To apply this result in our context, one can define the function f by

$$f(\rho_{\mathrm{eff}}) = \Pr\left[\log \left| \mathbf{I}_M + \frac{1}{M}\rho_{\mathrm{eff}} \mathbf{H}\mathbf{H}^H \right| \geq \xi \right]. \tag{27}$$

For the SISO case, for a channel with h following a complex normal distribution, it can be derived that $f(\rho) = \exp\left(-\frac{2^\xi - 1}{\rho}\right)$ which is sigmoidal. It turns out that proving that f is sigmoidal in the general case of MIMO is a non-trivial problem, as advocated by the current state of relevant literature [6], [34], [35]. In [6], $\nu_R(P)$ under perfect CSIR, was conjectured to be quasi-concave for general MIMO, and proven to be quasi-concave for the follwing special cases:

(a) $M \geq 1$, $N = 1$;
(b) $M \to +\infty$, $N < +\infty$, $\lim_{M\to\infty} \frac{N}{M} = 0$;
(c) $M < +\infty$, $N \to +\infty$, $\lim_{N\to\infty} \frac{M}{N} = 0$;
(d) $M \to +\infty$, $N \to +\infty$, $\lim\limits_{M\to+\infty, N\to+\infty} \frac{M}{N} = \ell < +\infty$;
(e) $\sigma^2 \to 0$;
(f) $\sigma^2 \to +\infty$;

In the following proposition, we give a sufficient condition to ensure that $\nu_R(P, t)$ is quasi-concave w.r.t P.

Proposition 4.1 (Optimization of $\nu_R(P, t)$ w.r.t P): If $\nu_R(P)$ with perfect CSIR is quasi-concave w.r.t P, then $\nu_R(P, t)$ is a quasi-concave function with respect to P, and has a unique maximum.

This proposition is proved in Appendix B. The above proposition makes characterizing the unique solution of $\frac{\partial \nu_R}{\partial P}(P, t) = 0$ relevant. This solution can be obtained through the root ρ^*_{eff} (which is unique because of [33]) of:

$$\frac{L}{\sigma^2}\left(P + \frac{b}{a}\right) \frac{\tau\rho\left[(\tau+1)\rho + 2\right]}{\left[(\tau+1)^2 + 1\right]^2} f'(\rho_{\mathrm{eff}}) - f(\rho_{\mathrm{eff}}) = 0 \tag{28}$$

with $\tau = \frac{t_s}{M}$. Note that P is related to ρ through $P = \sigma^2\rho$ and ρ is related to ρ_{eff} through (12) and can be expressed as

$$\rho = \frac{1}{2\tau}\rho_{\mathrm{eff}}\sqrt{(1+\tau)^2 + \frac{4\tau}{\rho_{\mathrm{eff}}}}. \tag{29}$$

Therefore (28) can be expressed as a function of ρ_{eff} and solved numerically; once ρ^*_{eff} has been determined, ρ^* follows by (29), and eventually P^* follows by (1). As a special case we have the scenario where $b = 0$ and $\tau \to +\infty$; this case is solved by finding the unique root of $\rho^* f'(\rho^*) - f(\rho^*) = 0$ which corresponds to the optimal operating SNR in terms of energy-efficiency of a channel with perfect CSI (as training time is infinite). Note that this equation is identical to that in [8] and in this work, we provide additional insights into the form of the function $f(.)$.

Quasi-concavity is an attractive property for the energy-efficiency as quasi-concave functions can be easily optimized numerically. Additionally, this property can also be used in multi-user scenarios for optimization and for proving the existence of a Nash Equilibrium in energy-efficient power control games [8], [36], [37].

B. Optimal fraction of training time

The expression of $\nu_R(P, t)$ shows that only the numerator depends on the fraction of training time. Choosing $t = 0$ maximizes $1 - \frac{t}{T}$ but the block success rate vanishes. Choosing $t = T$ maximizes the latter but makes the former term go to zero. Again, there is an optimal trade-off to be found. Interestingly, it is possible to show that the function $\nu_R(P^*, t)$ is strictly concave w.r.t. t for any MIMO channels in terms of

(M, N), where P^* is a maximum of ν_R w.r.t P. This property can be useful when performing a joint optimization of ν_R with respect to both P and t simultaneously. This is what the following proposition states.

Proposition 4.2 (Maximization of $\nu(P^*(t), t)$ w.r.t t): The energy-efficiency function $\nu_R(P^*(t), t)$ is a strictly concave function with respect to t for any $P^*(t)$ satisfying $\frac{\partial \nu_R}{\partial P}(P^*, t) = 0$ and $\frac{\partial^2 \nu_R}{\partial P^2}(P^*, t) < 0$, i.e, at the maximum of ν_R w.r.t. P.

The proof of this proposition is provided in Appendix C. The parameter space of ν_R is two dimensional and continuous as both P and t are continuous and thus the set $\nu(P^*(t), t)$ is also continuous and the proposition is mathematically sound. The proposition assures that the energy-efficiency can been maximized w.r.t. the transmit power and the training time jointly, provided $\nu_R(P, t)$ is quasi-concave w.r.t P for all t. Based on this, the optimal fraction of training time is obtained by setting $\frac{\partial \nu_R}{\partial t}(P, t)$ to zero which can be written as:

$$\left(\frac{T_s}{M} - \tau\right) \frac{\rho^2(\rho+1)}{[\tau\rho + \rho + 1]^2} f'(\rho_{\text{eff}}) - f(\rho_{\text{eff}}) = 0 \qquad (30)$$

again with $\tau = \frac{t_s}{M}$. In this case, following the same reasoning as for optimizing the ν_R w.r.t. P, it is possible to solve numerically the equation w.r.t. ρ_{eff} and find the optimal t_s, which is denoted by t_s^*.

Note that the energy-efficiency function is shown to be concave only when it has already been optimized w.r.t P. The optimization problem studied here is basically, a joint-optimization problem, and we show that once $\nu(P, t)$ is maximized w.r.t P for all t, then, $\nu(P^*(t), t)$ is concave w.r.t t. A solution to (30) exists only if ν_R has been optimized w.r.t P. However, in many practical situations, this optimization problem might not be readily solved as the optimization w.r.t P for all t has to be implemented first.

The following proposition describes how the optimal training time behaves as the transmit power is very large:

Proposition 4.3 (Optimal t in the high SNR regime): We have that: $\lim_{P \to +\infty} t_s^* = M$ for all MIMO systems.

The proof for this can be found in Appendix D.

C. Optimal number of antennas

So far we have always been assuming that the pre-coding matrix was chosen to be the identity matrix i.e., $\mathbf{Q} = \mathbf{I}_M$. Clearly, if nothing is known about the channel, the choice $\mathbf{Q} = \mathbf{I}_M$ is relevant (and may be shown to be optimal by formulating the problem as an inference problem). On the other hand, if some information about the channel is available (the channel statistics as far as this paper is concerned), it is possible to find a better pre-coding matrix. As conjectured in [5] and proved in some special cases (see e.g., [34]), the outage probability is minimized by choosing a diagonal pre-coding matrix and a certain number of 1's on the diagonal. The position of the 1's on the diagonal does not matter since channel matrices with i.i.d. entries are assumed. However, the optimal number of 1's depends on the operating SNR. The knowledge of the channel statistics can be used to compare the operating SNR with some thresholds and lead to this optimal number. Although we consider (9) as a performance metric instead of the outage probability, we are in a similar situation to [6], meaning that the optimal pre-coding matrix in terms of energy-efficiency is conjectured to have the same form and that the number of used antennas have to be optimized. In the setting of this paper, as the channel is estimated, an additional constraint has to be taken into account that is, the number of transmit antennas used, M, cannot exceed the number of training symbols t_s. This leads us to the following conjecture.

Conjecture 4.4 (Optimal number of antennas): For a given coherence time T_s, ν_R is maximized for $M^* = 1$ in the limit of $P \to 0$. As P increases, M^* also increases monotonically until some P_+ after which, M^* and t_s^* decreases. Asymptotically, as $P \to \infty$, $M^* = t_s^* = 1$.

This conjecture can be understood intuitively by noting that the only influence of M on ν_R is through the success rate. Therefore, optimizing M for any given P and t amounts to minimizing outage. In [5], it is conjectured that the covariance matrices minimizing the outage probability for MIMO channels with Gaussian fading are diagonal with either zeros or constant values on the diagonal. This has been proven for the special case of MISO in [34], we can conclude that the optimal number of antennas is one in the very low SNR regime and that it increments as the SNR increases. However, the effective SNR decreases by increasing M (seen from expression of ρ_{eff} and τ) , this will result in the optimal M for each P with training time lower than or equal to the optimal M obtained with perfect CSI. Concerning special cases, it can be easily verified that the optimal number of antennas is 1 at very low and high SNR.

At last, we would like to mention a possible refinement of the definition in (9) regarding M. Indeed, by creating a dependency of the parameter b towards M one can better model the energy consumption of a wireless device. For instance, if the transmitter architecture is such that one radio-frequency transmitter is used per antenna, then, each antenna will contribute to a separate fixed cost. In such a situation the total power can written as $aP + Mb_0$ where b_0 is the fixed energy consumption per antenna. It can be trivially seen that this does not affect the goodput in any manner and only brings in a constant change to the total power as long as M is kept a constant. Therefore, the optimization w.r.t P and t will not change but it will cause a significant impact on the optimal number of antennas to use.

V. Numerical results and interpretations

We present several simulations that support our conjectures as well as expand on our analytical results. All simulations are performed using Monte-Carlo simulations as there is no expression available for the outage of a general MIMO system.

A. With imperfect CSITR available

The F_L we use here is based on the results in [24], $F_L = Q_{func}(\frac{\xi - I_{\text{ICSITR}}(P, \mathbf{Q}_{WF}, \mathbf{H})}{\sqrt{\frac{2\rho}{(1+\rho)L}}})$, L being the code-length. This is the Gaussian approximation that is very accurate for L large enough and from simulations we observe that for $L \geq 10$ the approximation is quite valid.

First of all, numerical results are presented that support and present our analytical results through figures. The first two figures shown assume imperfect CSITR obtained through training and use a 2×2 MIMO system. The quasi-concavity of the the energy-efficiency function w.r.t the transmit power is shown in Figure 1 for $\xi = 1$ and $\xi = 4$, and $t_s = 2$ and $t_s = 10$. This figure shows that for a higher target rate, a longer training time yields a better energy-efficiency. We also observe that using a higher ξ can results in a better energy-efficiency as in this figure. This motivates us to numerically investigate if there is also an optimal spectral efficiency to use, given a certain T_s, $\frac{b}{a}$ and L. Figures 3 and 2 present the results of this study.

Surprisingly, we observe that our plots are quasi-concave and so there is an optimal target rate to use for each channel condition and code-length. In Figure 2, ν_T is always optimized over P and $\mathbf{Q}$. Observe that $\nu_T^*(\xi)$ is also quasi-concave and has a unique maximum for each value of d_i and t_s (representing the channel Eigen-values as from equations (16), (17) and training time lengths). d_i is ordered in an ascending order, i.e. in this case, with $d_1^2 \leq d_2^2$. The parameters used are: $M = N = 2$, $R_0 = 1$bps, $T_s = 100$, $L = 100$ and $\frac{b}{a} = 1$ mW with $t_s = 2, 10$ and 20 for $d_1^2 = 1$, $d_2^2 = 3$, and $t_s = 2$ for $d_1^2 = d_2^2 = 1$.This figure also implies that the training time and target rate can be optimized to yield the maximum energy-efficiency for a given coherence-time and channel fading. For infinite code-length the plot is maximized at the solution of (25). While for Figure 3, perfect CSIT is assumed with $b = 0$, at infinite block length, the optimal transmit rate/power is zero as expected (also seen from (25)). However, remarkably, for finite code-lengths there is a non-zero optimal rate and corresponding optimal power as seen from the figure.

Finally in Figure 4, we compare our energy efficiency function that uses optimized power allocation to uniform power allocation, and present the gain from having CSIT. In both cases, the training time and the transmit power is optimized and we plot the optimized energy efficiency v.s $P_{\max}$. Note that the optimized PA always yields a better performance when compared to UPA and at low power, UPA has almost zero efficiency while the optimal PA yields a finite efficiency. The gain observed can be considered as the major justification in using non-uniform power allocation and sending the channel state information to the transmitter. However, when the block length is small, imperfect CSIT results in a smaller gain as seen from the relatively larger gap between $T_s = 100$ and $T_s = 10000$ when compared to the size of the gap in UPA.

B. With no CSIT

We start off by confirming our conjecture that for a general MIMO system, $\nu_R(P, t)$ has a unique maximum w.r.t P. We also confirm that optimal values of training lengths and transmit antennas represented by t_s^* and M^* are as conjectured.

Once the analytical results have been established, we explore further and find out the optimal number of antennas

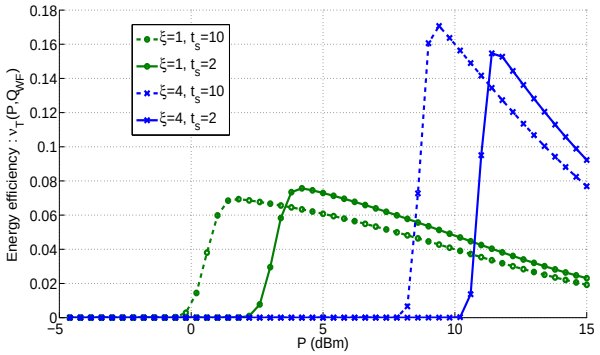


Fig. 1. Energy efficiency (ν_T) in bits/J v.s transmit power (P) in dBm for a MIMO system with imperfect CSITR, $M = N = 2$, $R_0 = 1$bps, $T_s = 100$, $\frac{b}{a} = 10$ mW for certain values of ξ and t_s.

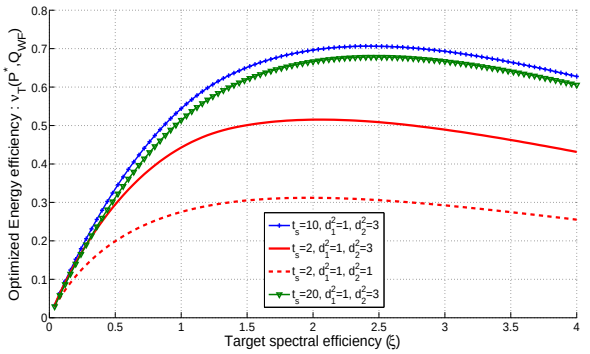


Fig. 2. Optimal energy-efficiency ($\nu_T(P^*, \mathbf{Q}_{WF})$) in bits/J v.s spectral efficiency (ξ) for a MIMO system with imperfect CSITR, $M = N = 2$, $R_0 = 1$bps, $T_s = 100$, $L = 100$ and $\frac{b}{a} = 1$ mW.

and training time when ν_R has been optimized w.r.t P. For this we use the optimized energy efficiency defined as $\nu^*(P, t) = \max\{\nu_R(p, t) \| p \in [0, P]\}$. As we know ν_R to be quasi-concave w.r.t P and having a unique maximum, this newly defined ν^* will indicate what is the best energy efficiency achievable given a certain amount of transmit power P. Hence, plotting ν^* against P for various values of M or t_s can be useful to determine the optimal number of antennas and training time while using the optimal power.

In the following plots we take $\frac{\sigma^2}{L} = 1mW$ so that P can be expressed in dBm easily. Also note that $\frac{b}{a}$ has the unit of power and is expressed in Watts (W). We also use $S_d = 15$ μs from LTE standards [38].

Figure 5 studies the energy efficiency as a function of the transmit power (P) for different values of $\frac{b}{a}$ and illustrates the quasi-concavity of the energy efficiency function w.r.t P. The parameters used are $R = 1600$, $\xi = \frac{R}{R_0} = 16$, $T_s = 55$ and $M = N = t = 4$.

Figure 6 studies the optimized energy efficiency ν^* as a function of the transmit power with various values of t_s. The figure illustrates that beyond a certain threshold on the available transmit power, there is an optimal training sequence length that has to be used to maximize the efficiency, when the optimization w.r.t P has been done, which has been

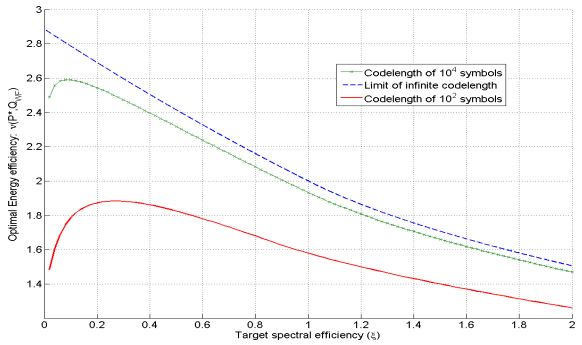


Fig. 3. Optimal energy-efficiency ($\nu_T(P^*, \mathbf{Q}_{WF})$) v.s spectral efficiency (ξ) for a MIMO system with perfect CSITR, $M = N = 2$, $R_0 = 1$bps and $\frac{b}{a} = 0$.

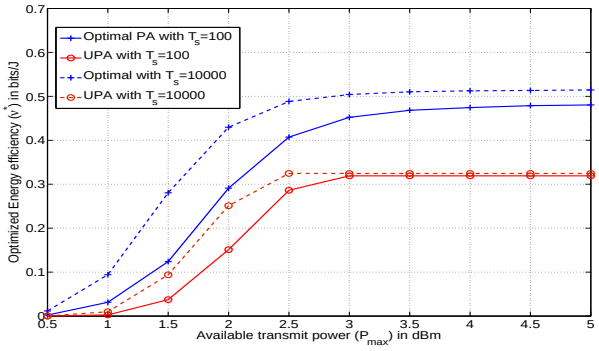


Fig. 4. Optimal energy-efficiency ($\nu_R(P^*)$) in bits/J v.s available transmit power ($\sup(P)$) for a MIMO system with imperfect CSITR, $M = N = 2$, $R_0 = 1$bps, $R = 1$bps and $\frac{b}{a} = 1$ mW.

proven analytically in proposition 4.2. The parameters are $R = 1$Mbps, $\xi = 16$, $\frac{b}{a} = 0$, $M = N = 4$, $\frac{b}{a} = 0$ and $T_s = 55$.

Figure 7 studies the optimal training sequence length t_s as a function of the transmit power P. Note that in this case, we are not optimizing the efficiency with respect to P and so this figure illustrates proposition 4.3. With P large enough $t_s = M$ becomes the optimal training time and for P small enough $t_s = T_s - 1$ as seen from the figure. The parameters are $R = 1600$, $\frac{b}{a} = 0$ W, $\xi = 16$ and $T_s = 10$. (We use $T_s = 10$, as if the coherence time is too large, the outage probabilities for low powers that maximize the training time, such that $t_s^* = T_s - 1$, become too small for any realistic computation.)

Figure 8 studies the optimal number of antennas M^* as a function of the transmit power P with the training time optimized jointly with M. With P large enough $M = t_s = 1$ becomes the optimal number of antennas and for P small enough $M = 1$ as seen from the figure. This figure illustrates conjecture 4.4. The parameters are $R_0 = 1$Mbps, $\frac{b}{a} = 10$ mW and $T_s = 100$.

From all of our theoretical and numerical results so far, we can conclude that given a target spectral efficiency ξ, a coherence block length T_s and number of receive antennas, there is an optimal transmit power P^*, transmit antennas M^* and training time t_s^* to use that optimizes the energy efficiency.

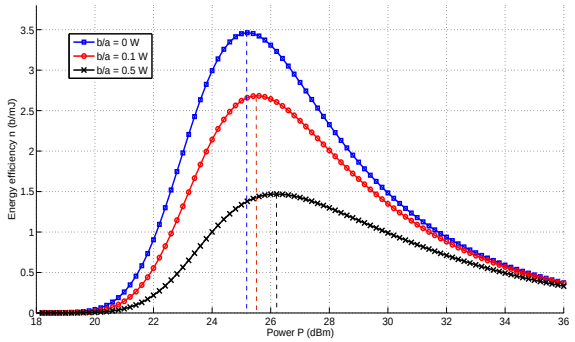


Fig. 5. Energy efficiency (ν_R) v.s transmit power (P) with $t_s = M = N = 4$,$R = 1600$bps, $\xi = \frac{R}{R_0}$=16 and $T_s = 55$ symbols.

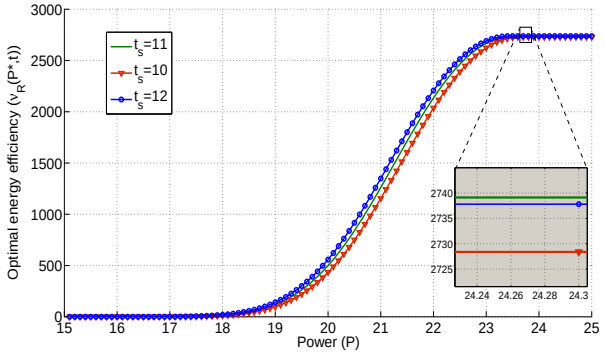


Fig. 6. Optimized efficiency (ν^*) vs. maximum transmit power (P) for a MIMO system with $M = N = 4$, $R = 1$ Mbps, $\xi = \frac{R}{R_0}$=16, $T_s = 55$ and $\frac{b}{a} = 0$W.

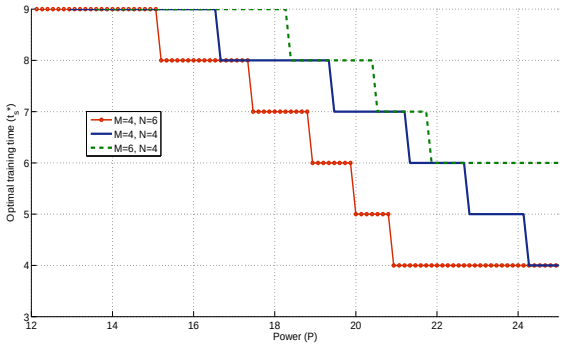


Fig. 7. Optimal training sequence length (t_s) vs. Transmit Power (P) MIMO system with $\xi = \frac{R}{R_0} = 16$,$R = 1$Mbps, $N = 4$, $T_s = 10$ symbols. The discontinuity is due to the discreteness of t_s.

VI. Conclusion

This paper proposes a framework for studying the problem of energy-efficient pre-coding (which includes the problem

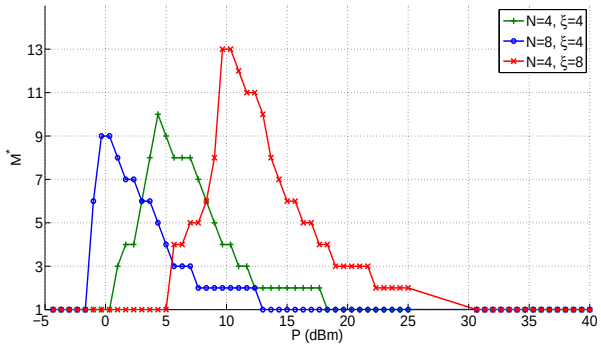


Fig. 8. Optimal number of antennas (M^*) vs. Transmit Power (P) in dBm for a MIMO system with $R_0 = 1$Mbps, $T_s = 100$, for certain values of ξ and N, and t_s optimized jointly with M. The discontinuity is due to the discreteness of M.

of power allocation and control) over MIMO channels under imperfect channel state information and the regime of finite block length. As in [8], energy-efficiency is defined as the ratio of the block success rate to the transmit power. But, in contrast with [8] and the vast majority of works originating from it, we do not assume an empirical choice for the success rate such as taking $f(x) = (1 - e^{-x})^L$, L is the block length. Instead, the numerator of the proposed performance metric is built from the notion of information, and more precisely from the average information (resp. mutual information) in the case where CSIT is available (resp. not available). This choice, in addition to giving a more fundamental interpretation to the metric introduced in [8], allows one to take into account in a relatively simple manner effects of practical interest such as channel estimation error and block length finiteness. Both in the case where (imperfect) CSIT is available and not available, it is shown that using all the available transmit power is not optimal. When CSIT is available, whereas determining the optimal power allocation scheme is a well known result (water-filling), finding the optimal total amount of power to be effectively used is a non-trivial choice. Interestingly, the corresponding optimization problem can be shown to be quasi-convex and have a unique solution, the latter being characterized by an equation which is easy to solved numerically. When CSIT is not available, solving the pre-coding problem in the general case amounts to solving the Telatar's conjecture. Therefore, a new conjecture is proposed and shown to become a theorem in several special cases. Interestingly, in this scenario, it is possible to provide a simple equation characterizing the optimal fraction of training time. Numerical results are provided to sustain the proposed analytical framework, from which interesting observations can be made which includes : block length finiteness gives birth to the existence of a non-trivial trade-off between spectral efficiency and energy efficiency ; using optimal power allocation brings a large gain in terms of energy-efficiency only when the channel has a large enough coherence time,demonstrating the value of CSIT and channel training.

The proposed framework is useful for engineers since it provides considerable insights into designing the physical layer of MIMO systems under several assumptions on CSI. The proposed framework also opens some interesting research problems related to MIMO transmission, which includes : finding the optimal pre-coding matrix for the general case of i.i.d. channel matrices under no CSIT. Even in the case of large MIMO systems, this problem is not solved ; extending the proposed approach to the case of Rician channels with spatial correlations ; tackling the important case of multiuser MIMO channels ; considering the problem of distributed energy-efficient pre-coding.

APPENDIX A
PROOF OF THEOREM 3.1

In order to prove that $\nu_T(P, \mathbf{Q}_{WF(P)}, \mathbf{H})$ is quasi-concave with respect to P and has a unique maximum $\nu_T(P^*, \mathbf{Q}_{WF(P^*)}, \mathbf{H})$, we exploit the result in [33] which states that if $f(x)$ is an "S"-shaped or sigmoidal function, then $\frac{f(x)}{x}$ is a quasi-concave function with a unique maximum. An "S"-shaped or sigmoidal function has been defined in [33] in the following manner. A function f is "S" shaped, if it satisfies the following properties:

1) Its domain is the interval $[0, \infty)$.
2) Its range is the interval $[0, 1)$.
3) It is increasing.
4) ("Initial convexity") It is strictly convex over the interval $[0, x_f]$, with x_f a positive number.
5) ("Eventual concavity") It is strictly concave over any interval of the form $[x_f, L]$, where $x_f < L$.
6) It has a continuous derivative.

Considering the non-constant terms in ν_T, we see that what we have to show is that $F_L(I_{\text{ICSITR}}(P, \mathbf{Q}_{WF(P)}, \mathbf{H}) - \xi)$ is "S"-shaped w.r.t P. We already have that $F_L(x)$ is sigmoidal, therefore all we have to show is that $F_L(g(P))$ is also sigmoidal where $g(P) = I_{\text{ICSITR}}(P, \mathbf{Q}_{WF(P)}, \mathbf{H}) - \xi$. Trivially, when $P = 0$, $F_L(I_{\text{ICSITR}}(P)) = 0$ and $\lim_{P\to\infty} F_L = 1$. The rest can be proved using the following arguments:

- $g(P)$ is continuous: As P varies, $\mathbf{Q}_{WF(P)}$ also is modified according to the iterative water-filling algorithm. This results in using one antenna for low ρ to all antennas for high values of ρ.
  There exists certain "threshold" points of the total power, $P_i^{\text{th}}, i = \{1, \ldots, M\}$, at which the number of antennas used changes. The convention being, for $P_{i-1}^{\text{th}} \leq P \leq P_i^{\text{th}}$, i number of antennas are used (s for the rest are set to zero). $P_0^{\text{th}} = 0$ and $P_M^{\text{th}} = \infty$. If $I_{\text{CSITR}}(P, \mathbf{Q}_{WF(P)}, \mathbf{H})$ is continuous at these points, then $g(P)$ is continuous. It can also be observed that in all other points, $I_{\text{CSITR}}(P, \mathbf{Q}_{WF(P)}, \mathbf{H})$ can be expressed as $\Sigma_{i=1}^{J} \log(1 + \alpha_i + \beta_i s_i), J \leq \min(M, N)$. ($\alpha$ and β is obtained from solving (17).) A "threshold" point occurs when $P = P_j^{\text{th}}$ $s_j = 0$ is obtained by solving (17). The left hand limit is that $j - 1$ antennas are used and so, $I_{\text{CSITR}}(P, \mathbf{Q}_{WF(P)}, \mathbf{H}) = \Sigma_{i=1}^{j-1} \log(1 + \alpha_i + \beta_i s_i)$. The right hand limit will be obtained by solving (17), with $s_1 \to 0$ (assuming without loss of generality that d_1^2 is the smallest). This will yield a solution which can

be easily seen to be the same as the left hand limit as $p_1 \to 0$.

- We have shown $g(P)$ to be a finite sum of logarithms of a monomial expansion of P in certain intervals (marked by P_i^{th}) . For each interval it is trivial to see that $F_L(g(P))$ is also "S"-shaped. As $g(P)$ is continuous, $F_L(g(P))$ is "S"-shaped for all P.
- From Lemma B proved in Appendix B we can show that $\frac{F_L(g(P))}{aP+b}$ is also "S"-shaped by a simple change of variable $x = aP + b$. Thus, we have $\nu_T(P, \mathbf{Q}_{WF(P)}, \mathbf{H})$ as a quasi-concave function with a unique maximum.
- With imperfect CSI, the only change is in $I_{\text{ICSITR}}(P^*, \mathbf{Q}_{WF(P^*)}, \widehat{\mathbf{H}})$ now given from (15). The water-filling algorithm now replaces $\mathbf{H}$ with $\widehat{\mathbf{H}}$ and so on. This maintains the continuity of $g(P)$. However we now have $I_{\text{ICSITR}}(P, \mathbf{Q}_{WF(P)}, \widehat{\mathbf{H}}) = \Sigma_{i=1}^{j} \log\left(1 + \frac{\alpha_i + \beta_i s_i}{1+\rho\sigma_E^2}\right)$. From [19] we have $\frac{1}{1+\rho\sigma_E^2}$ as a concave function and so even in this case, we have $F_L(g(P))$ as a sigmoidal function and $\nu_T(P, \mathbf{Q}_{WF(P)}, \widehat{\mathbf{H}})$ as quasi-concave with a unique maximum. As it is continuous and differentiable, the maximum can be found as the unique solution to the equation:

$$\begin{aligned} &\frac{\partial F_L[I_{\text{ICSITR}}(P^*, \mathbf{Q}_{WF(P^*)}, \widehat{\mathbf{H}}) - \xi]}{\partial P}\left(P^* + \frac{b}{a}\right) \\ &-F_L[I_{\text{ICSITR}}(P^*, \mathbf{Q}_{WF(P^*)}, \mathbf{H}) - \xi] \quad = 0 \end{aligned} \tag{31}$$

where $\frac{\partial}{\partial P}$ is the partial derivative.

QED

APPENDIX B
PROOF OF PROPOSITION 4.1

An "S"-shaped function has been defined in [33] in the following manner. A function f is "S" shaped, if it satisfies the properties as mentioned in Appendix A.

Lemma 1: If f is a "S" shaped function, the composite function $f \circ g(x)$ is also "S" shaped if g satisfies the following properties:

1) g also satisfies conditions 1, 3, 4 and 6 but with $g(0) = b, b > 0$.
2) $\lim_{x\to\infty} f'(x)g'(x) = 0$.
3) $g''(x)$ is a decreasing function such that $\lim_{x\to\infty} g''(x) = 0$.

The proof for the above Lemma is at the end of this section.

In [6], the authors prove that the energy efficiency function with perfect CSI defined as the goodput ration to transmitted RF signal power is a quasi concave function by showing that the success rate function, $f(\rho)$ is "S" shaped for the following cases:

(a) $M \geq 1$, $N = 1$;
(b) $M \to +\infty$, $N < +\infty$;
(c) $M < +\infty$, $N \to +\infty$;
(d) $M \to +\infty$, $N \to +\infty$, $\lim_{M\to+\infty, N\to+\infty} \frac{M}{N} = \ell < +\infty$;
(e) $\sigma^2 \to 0$;
(f) $\sigma^2 \to +\infty$;

So, if we can show that the success rate function in our situation is also "S" shaped, our proof is complete for all the cases mentioned above. From (11) we know that the worst case mutual information in the case of imperfect CSI with training is mathematically equivalent to that of perfect CSI but with ρ replaced by ρ_{eff}. Thus it is possible to replace $f(\rho)$, in the case of perfect CSI, by $f(\rho_{\text{eff}})$, when we study the case of imperfect CSI, and so we can study the energy efficiency function given by:

$$\nu_R(P, t) = R\zeta \frac{f(\rho_{\text{eff}}(p(x)))}{x} \tag{32}$$

where x is a new variable that represents the total consumed power and $p(x) = \frac{L(x-b)}{a\sigma^2}$. $p'(x) > 0$ and $p''(x) = 0$ and $\rho'_{\text{eff}}(\rho) > 0$ and $\lim_{\rho\to\infty} \rho''_{\text{eff}}(\rho) = 0$. Thus ρ_{eff} and p satisfy the conditions on g detailed in Lemma B. Hence we have proven that the numerator is "S" shaped with respect to x and then it immediately follows from the results in [33] that ν_R has a unique maximum and is quasi-concave for all the specified cases.

Proof of Lemma

Here we show that $f \circ g$ also satisfies all the properties of the "S" function as described in [33].

1) Its domain is the domain of g which is clearly the non-negative part of the real line; that is, the interval $[0, \infty)$.
2) Its range is the range of f, the interval $[0, 1)$.
3) It is increasing as both f and g are increasing.
4) ("Initial convexity"). Note that $f(g(x))'' = f''(y)g'(x) + g''(x)f'(y)$, with $y = g(x)$. As all terms in this expansion are positive in the interval $[0, x_f]$, $f \circ g$ is also convex in this interval. Also note that as g' and f' are strictly positive and g'' is decreasing, thus for $y > x_f$ once $f(g(x))'' < 0$ it stays negative till infinity. This implies that if there is an inflexion point, it is unique.
5) ("Eventual concavity") Consider $h(x) = f(g(x))' = f'(y)g'(x)$, due to the initial convexity and increasing nature of h, $h(x_f) = k, k > 0$. $\lim_{x\to\infty} f(g(x_f))' = 0$. As h is continuous the mean value theorem imposes $h'(x) < 0$ at some point. This implies that there exists some point $x_d > 0$ such that $f \circ g$ is concave in the interval $[x_d, \infty]$ and convex before it.
6) It has a continuous derivative. (all the functions used here are continuous)

Hence, $f \circ g$ is "S" shaped.

QED

APPENDIX C
PROOF OF PROPOSITION 4.2

Let us consider the second partial derivative of ν_R with respect to t. (Note that this is possible as t is a real number with the unit of time while t_s is a natural number.) From (32), $\nu_R(P, t) = K^{-1}(1 - \frac{t}{T})f(\rho_{\text{eff}})$, with $K^{-1} = \frac{R}{x}$ a constant if P is held a constant.

$$\begin{aligned} \frac{K\partial^2\nu}{\partial t^2} = \; & (1 - \tfrac{t}{T})f''(\rho_{\text{eff}})\rho'^2_{\text{eff}}(t) + (1 - \tfrac{t}{T})f'(\rho_{\text{eff}})\rho''_{\text{eff}}(t) \\ & -\tfrac{2}{T}f'(\rho_{\text{eff}})\rho'_{\text{eff}}(t) \end{aligned} \tag{33}$$

In the above sum, it can be easily verified that the terms $f'(\rho_{\text{eff}})\rho'_{\text{eff}}(t)$ and $f'(\rho_{\text{eff}})$ are positive and that $\rho''_{\text{eff}}(t) < 0$. Thus if we have $f''(\rho_{\text{eff}}) < 0$, then $\nu_R(t)$ is strictly concave.

The only way ν_R depends on P is through $\frac{f(\rho_{\text{eff}})}{aP+b}$ and $R\zeta$ stays a constant if only P changes. So if we use the fact that we are working at a maximum of ν_R with respect to ρ, i.e $\frac{\partial \nu}{\partial \rho} = 0$ and $\frac{\partial^2 \nu}{\partial \rho^2} < 0$, we have $\frac{\partial \nu}{\partial \rho}$ as:

$$0 = f'(\rho_{\text{eff}})\rho'_{\text{eff}}(\rho)\rho^{-1} - f(\rho_{\text{eff}})\rho^{-2} \tag{34}$$

And, substituting (34) in $\frac{\partial^2 \nu}{\partial \rho^2} < 0$, that is:

$$\begin{aligned} f''(\rho_{\text{eff}})(\rho'_{\text{eff}})^2\rho^{-1} - 2f'(\rho_{\text{eff}})\rho'_{\text{eff}}\rho^{-2} + 2f(\rho_{\text{eff}})\rho^{-3} &< 0 \\ f''(\rho_{\text{eff}})(\rho'_{\text{eff}})^2\rho^{-1} + \frac{2}{\rho}(0) &< 0 \\ f''(\rho_{\text{eff}}) < 0 \end{aligned} \tag{35}$$

Thus, using (35) in (33), we have the result that $\nu_R(P^*, t)$ is strictly concave w.r.t t.

QED

APPENDIX D
PROOF OF PROPOSITION 4.3

The equation that describes the optimal training time t_s^* can be written as:

$$(T - t_s^*)f'(\rho_{\text{eff}})\rho'_{\text{eff}}(t)_{t=S_d t_s^*} - f(\rho_{\text{eff}}) = 0 \tag{36}$$

Now, let us study the optimal training time t_s^* in the very high SNR regime, ie. when $\rho \to \infty$. (Note that $P \to \infty$ is equivalent to $\rho \to \infty$).

High SNR regime: Applying the limit of $\rho \to \infty$ in (36), we get that

$$T_s - t_s^* = \lim_{\rho\to\infty} \frac{f(t\rho(1+t)^{-1})}{\rho f'(t\rho(1+t)^{-1})} \tag{37}$$

We know from various works including [6] that $\lim_{\rho\to\infty} f(t\rho(1+t)^{-1}) = 1$. Now let us consider $f'(\frac{t}{1+t}\rho)$.

For a MISO system we know from [39] that:

$$f_{MISO}(\rho_{\text{eff}}) \quad = \quad \frac{\gamma\left(M, \frac{2^\xi - 1}{\sqrt{\rho_{\text{eff}}}}\right)}{\Gamma(M)} \tag{38}$$

where Γ is the Gamma function and γ is the lower incomplete Gamma function. Now we can use the special property of the incomplete gamma function, $\lim_{x\to 0} \frac{\gamma(M,x)}{x^M} = 1/M$ detailed in [40] to determine $\lim_{\rho\to\infty} \rho f(\rho_{\text{eff}})' \propto \rho_{\text{eff}}^{-M/2}$. Plugging this into (37) we have:

$$\lim_{\rho\to\infty} T_s - t_s^* = \frac{1+t}{t}\rho_{\text{eff}}^{M/2} \tag{39}$$

Thus ν_R is optimized when $t_s \to -\infty$, but as $M \le t_s < T_s$, we have $\lim_{\rho\to\infty} t_s^* = M$ for all MISO systems.

Now for any MIMO system in general, using the eigen value decomposition of $\mathbf{H}\mathbf{H}^H$ we have the eigenvalue decomposition $\log_2 |I_M + \frac{\rho}{M}\mathbf{H}\mathbf{H}^H| = \log_2(\Pi_{i=1}^L(1+\lambda_i))$, where $L = \min(M, N)$ and λ_i are the eigenvalues of $\mathbf{H}\mathbf{H}^H$. Applying the limit on ρ and ignoring lower order terms we have

$$\lim_{\rho\to\infty} f(\rho_{\text{eff}}) = \Pr\left[\Pi_{i=1}^L \lambda_i \ge \frac{2^\xi}{\rho_{\text{eff}}^L}\right]. \tag{40}$$

We can observe that the above expression is a cumulative distribution function of $\Pi_{i=1}^L\lambda_i$ and so it's derivative is simply the PDF of $\Pi_{i=1}^L\lambda_i$. For $\rho \to \infty$ we have $f'(\rho_{\text{eff}}) = \Pr[\Pi_{i=1}^L\lambda_i = \frac{2^\xi}{\rho_{\text{eff}}^L}]$ As we know that the in general, if the number of transmit antennas are the same, $\Pr[\lambda_{MISO} > x] \le \Pr[\Pi_{i=1}^L\lambda_i > x]$ for any $x > 0$ [35] . Thus $f'_{MIMO}(\rho_{\text{eff}}) < f'_{MISO}(\rho_{\text{eff}})$, implying that for all MIMO systems $\lim_{\rho\to\infty} t_s^* = M$ from (37) and (39).

REFERENCES

[1] D. Lister, "An Operator's View on Green Radio", Proc. IEEE Internat. Conf. on Commun, Workshops (ICC Workshops '09), 1st Internat. Workshop on Green Commun. (GreenComm '09) (Dresden, Ger, 2009).

[2] J. Palicot, and C. Roland, "On The Use Of Cognitive Radio For Decreasing The Electromagnetic Radiations", URSI 05, XXVIII General Assembly, New Delhi, India, October 23-29, 2005.

[3] GreenTouch, "Communications Turns Totally Green", Press Release, Jan. 11, 2010.

[4] Mobile VCE, Virtual Centre of Excellence in Mobile and Personal Communications, Core 5 Research Programme, "Green Radio - Sustainable Wireless Networks", Feb. 2009.

[5] E. Telatar, "Capacity of Multi-antenna Gaussian Channels", *European Trans. Telecommunications*, Vol. 10 (1999), pp. 585-595.

[6] E.V. Belmega, and S. Lasaulce, "Energy-efficient pre-coding for multiple-antenna terminals", IEEE. Trans. on Signal Processing, 59, 1, Jan. 2011.

[7] S. Verdú, "Spectral efficiency in the wideband regime", IEEE Trans. on Inf. Theory, vol. 48, no. 6, pp. 1319-1343, Jun. 2002.

[8] D. J. Goodman, and N. Mandayam, "Power Control for Wireless Data", IEEE Personal Communications, vol. 7, pp. 48-54, Apr. 2000.

[9] S. Cui, A. Goldsmith, and A. Bahai, "Energy-efficiency of MIMO and Cooperative MIMO Techniques in Sensor Networks" , IEEE Journal on Selected Areas in Communications, Vol. 22, No. 6, pp. 1089-1098, Aug. 2004.

[10] V. Chandar, A. Tchamkerten, and D. Tse, Asynchronous Capacity Per Unit Cost , IEEE Inter. Symposium on Information Theory Proceedings (ISIT), pp. 280-284, June 2010.

[11] 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. 427-1432, Sep. 1998.

[12] 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.

[13] S. Yatawatta, A.P. Petropulu, and C.J. Graff, "Energy-efficient channel estimation in MIMO systems", EURASIP Journal on Wireless Communications and Networking, v.2006 n.2, pp. 317-320, April 2006.

[14] S. Buzzi, H. V. Poor, and D. Saturnino, "Energy-efficient resource allocation in multiuser MIMO systems: a game-theoretic framework, Proc. of the 16th European Signal Processing Conference (EUSIPCO 2008), Losanna (Svizzera), August 2008.

[15] T. Yoo, N. Jindal, and A. Goldsmith, "Multi-Antenna Downlink Channels with Limited Feedback and User Selection", IEEE Journal on Selected Areas in Communications (JSAC), pp. 1478-1491, Sep. 2007.

[16] G. Caire, N. Jindal, M. Kobayashi, and N. Ravindran, "Multiuser MIMO achievable rates with downlink training and channel state", IEEE Transactions on Information Theory, 56(6): 2845-2866, 2010.

[17] C. C. Gaudes, and E Masgrau, "Bounds on Capacity over Gaussian MIMO Multiaccess Channels with Channel State Information Mismatch," IEEE ICASSP06, vol. 4, pp. 93-96, May 2006

[18] T. L. Marzetta, and B. M. Hochwald, "Capacity of a Mobile Multiple-Antenna Communication Link in Rayleigh Flat Fading", *IEEE Trans. on Information Theory*, Vol. 45, No. 1, Jan. 1999.

[19] B. Hassibi, and B. M. Hochwald, "How Much Training is Needed in Multiple-Antenna Wireless Links?", *IEEE Trans. on Information Theory*, Vol. 49, No. 4, April, 2003.

[20] Polyanskiy, H. Vincent Poor, and S. Verdú, "Channel Coding Rate in the Finite Blocklength Regime," IEEE Trans. Information Theory, vol. 56, no. 5, pp. 2307–2359, May 2010.
[21] F. Richter, A. J. Fehske, and G. Fettweis, "Energy Efficiency Aspects of Base Station Deployment Strategies for Cellular Networks", Proceedings of VTC Fall, pp. 1-5, Dresden, 2009.
[22] S. K. Jayaweera, "A virtual MIMO-based cooperative communications architecture for energy-constrained wireless sensor networks", IEEE Trans. Wireless Commun., vol. 5, pp. 984-989, May 2006.
[23] F. Meshkati, H. V. Poor, S. C. Schwartz, and N. B. Mandayam, "An energy-efficient approach to power control and receiver design in wireless data networks," IEEE Trans. Commun., vol. 52, pp. 1885-1894, Nov. 2005.
[24] D. Buckingham, and M.C. Valenti, "The information-outage probability of finite-length codes over AWGN channels", IEEE Information Sciences and Systems, 2008. CISS 2008.
[25] C.E. Shannon, "A Mathematical Theory of Communication", Bell System Technical Journal, vol. 27, pp. 379-423, 623-656, July, October, 1948.
[26] J. Hoydis, R. Couillet, P. Piantanida, and M. Debbah, "A Random Matrix Approach to the Finite Blocklength Regime of MIMO Fading Channels", IEEE International Symposium on Information Theory (ISIT), pp. 2181 - 2185, Cambridge, 2012.
[27] Q. Liu, S. Zhou, and G. B. Giannakis, "Cross-Layer Combining of Adaptive Modulation and Coding with Truncated ARQ Over Wireless Links, IEEE Trans. Wireless Commun., vol. 3, no. 5, pp. 1746-1755, Sept. 2004.
[28] L.H. Ozarow, S. Shamai, and A.D. Wyner, "Information theoretic considerations for cellular mobile radio", IEEE Vehicular Technology, vol. 43, No. 2, pp 359-378, May. 1994.
[29] FCC Sec.15.249, *Operation within the bands 902-928 MHz, 2400-2483.5 MHz, 5725-5875 MHZ, and 24.0-24.25 GHz*, Oct. 2011.
[30] T. M. Cover, and J. A. Thomas, "Elements of information theory", Wiley-Interscience, New York, NY, 1991.
[31] S. Lasaulce, and N. Sellami, "On the Impact of using Unreliable Data on the Bootstrap Channel Estimation Performance", *Proc. IEEE SPAWC*, Italy, June 2003, pp. 348-352.
[32] D. Love, R. Heath, V. Lau, D. Gesbert, B. Rao, and M. Andrews, "An overview of limited feedback in wireless communication systems", IEEE Journal on Selected Areas Communications, vol 26, pp. 1341-1365, 2008.
[33] V. Rodriguez, "An Analytical Foundation for Ressource Management in Wireless Communication", IEEE Proc. of Globecom, San Francisco, CA, USA, pp. 898-902, Dec. 2003.
[34] E. A. Jorswieck, and H. Boche, "Outage probability in multiple antenna systems", European Transactions on Telecommunications, vol. 18, pp. 217-233, 2006.
[35] A. Edelman, "Eigenvalues and Condition Numbers of Random Matrices", PhD thesis, Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA, 1989.
[36] S. Lasaulce, and H. Tembine, "Game Theory and Learning for Wireless Networks: Fundamentals and Applications", Academic Press, pp. 1-336, Aug. 2011, ISBN 978-0123846983.
[37] S. Lasaulce, M. Debbah, and E. Altman, Methodologies for analyzing equilibria in wireless games, IEEE Signal Processing Magazine, vol. 26, no: 5, pp. 41-52, Sep. 2009.
[38] 3GPP TR 36.814 v0.4.1, "Further Advancements for E-UTRA", *Physical Layer Aspects*, Feb. 2009.
[39] M. Alexander, F. A. Graybill, and D. C. Boes (1974). *"Introduction to the Theory of Statistics"* (Third Edition, p. 241-246). McGraw-Hill. ISBN 0-07-042864-6.
[40] Digital Library of Mathematical Functions, Incomplete Gamma functions, 8.11. [Online] Available: http://dlmf.nist.gov/8.11#i



Vineeth S Varma was born in Tripunithura, India. He is currently doing his PhD at LSS under the supervision of Dr. Samson Lasaulce. He obtained his Bachelors in Physics with Honors from Chennai Mathematical Institute, India in 2008. He then proceeded to join the Erasmus Mundus course in Optics and obtained Masters in Science and Technology from Friedrich-Schiller-University of Jena in 2009 and Warsaw University of Technology in 2010. His areas of interest are Physics, Applied mathematics and Telecommunication. Since December 2010 he is working on his PhD aimed at finding the fundamental limits of energy on telecommunication and optimizing energy efficiency in networks. He is the recipient of the 2012 ACM VALUETOOLS best student paper award.



Samson Lasaulce received his BSc and Agrégation degree in Physics from École Normale Supérieure (Cachan) and his MSc and PhD in Signal Processing from École Nationale Supérieure des Télécommunications (Paris). He has been working with Motorola Labs for three years (1999, 2000, 2001) and with France Télécom R&D for two years (2002, 2003). Since 2004, he has joined the CNRS and Supélec as a Senior Researcher. Since 2004, he is also Professor at École Polytechnique. His broad interests lie in the areas of communications, information theory, signal processing with a special emphasis on game theory for communications. Samson Lasaulce is the recipient of the 2007 ACM VALUETOOLS, 2009 IEEE CROWNCOM, 2012 ACM VALUETOOLS best student paper award, and the 2011 IEEE NETGCOOP best paper award. He is an author of the book Game Theory and Learning for Wireless Networks: Fundamentals and Applications. He is currently an Associate Editor of the IEEE Transactions on Signal Processing.



Mérouane Debbah entered the Ecole Normale Supérieure de Cachan (France) in 1996 where he received his M.Sc and Ph.D. degrees respectively. He worked for Motorola Labs (Saclay, France) from 1999-2002 and the Vienna Research Center for Telecommunications (Vienna, Austria) until 2003. He then joined the Mobile Communications department of the Institut Eurecom (Sophia Antipolis, France) as an Assistant Professor until 2007. He is now a Full Professor at Supelec (Gif-sur-Yvette, France), holder of the Alcatel-Lucent Chair on Flexible Radio and a recipient of the ERC starting grant MORE (Advanced Mathematical Tools for Complex Network Engineering). His research interests are in information theory, signal processing and wireless communications. He is a senior area editor for IEEE Transactions on Signal Processing. Mérouane Debbah is the recipient of the "Mario Boella" award in 2005, the 2007 General Symposium IEEE GLOBECOM best paper award, the Wi-Opt 2009 best paper award, the 2010 Newcom++ best paper award as well as the Valuetools 2007, Valuetools 2008, Valuetools 2012 and CrownCom2009 best student paper awards. He is a WWRF fellow. In 2011, he received the IEEE Glavieux Prize Award.



Salah Eddine Elayoubi received his MS in telecommunications and networking from the National Polytechnic Institute at Toulouse, France, in 2001, and his Ph.D. and Habilitation in Computer science from the University of Paris VI in 2004 and 2009, respectively. Since then, he has been working in Orange Labs, the research and development division of France Telecom group. He is now a senior research engineer, involved in several EU-funded research projects. His research interests include radio resource management, modeling and performance evaluation of mobile networks.