Conditions for Multi-Antenna Selection to be Optimal Given Channel Amplitude Information

Pengfei Xia, *Member, IEEE,* and Georgios B. Giannakis, *Fellow, IEEE*

***Abstract*— Transmitter designs based on partial channel state information (CSI) have become increasingly attractive in multi-antenna wireless communication systems. To capture partial CSI statistics at the transmitter, we rely on a channel amplitude information (CAI) model, based on which we aim for maximizing the random channel's average mutual information. Due to the high computational complexity required for such optimal transmissions, we resort to reduced complexity practical schemes and derive necessary and sufficient conditions for these alternatives to achieve maximal average mutual information.**

***Index Terms*— Antenna selection, channel amplitude information, multi-input single-output, Nakagami fading.**

I. INTRODUCTION AND MOTIVATION

MULTI-ANTENNA wireless communication systems are receiving increasing attention due to their potential for high data rates and resilient error performance [12]. Knowledge of channel state information (CSI) at the transmitter plays an important role in these designs [3]. However, because of the time varying nature of the underlying wireless channels, the available CSI at the transmitter suffers from various imperfections, such as delay, estimation error and bandwidth constraints on the feedback link. It is thus more realistic to deal with partial, or incomplete CSI at the transmitter, where various CSI imperfections are explicitly taken into account.

In certain cases, partial CSI can be modeled as Gaussian distributed. Within this class, two special models have been studied extensively. One is the so-called channel mean information (CMI) model, that assumes knowledge of the channel mean and views the residual uncertainty as white Gaussian noise. The other is the so-termed channel covariance information (CCI) model, where the Gaussian channel is assumed to be zero mean with a nonwhite covariance matrix. Impact of partial CSI has been extensively studied under these two models [5]–[8], [10], [14], [17].

However, the Gaussian assumption of the partial CSI is not always valid [13]. Physical measurements confirm that in most wireless systems, the fading channel amplitudes are best modeled as Nakagami distributed [1], [9], while the channel phases are generally treated as independent and uniformly distributed [2]. It should further be noticed that certain feedback strategies provide the transmitter with information about channel amplitudes, but not channel phases. For example, when antenna subset selection is employed at the transmitter, the antenna subset index chosen depends on SNR (i.e., channel amplitude) information only [4].

We are therefore motivated to exploit statistical knowledge of channel amplitudes at the transmitter. In this paper, we introduce a partial channel amplitude information (CAI) model, which models channel amplitudes as a group of Nakagami random variables. Based on this partial CAI model, we study the optimal transmission strategy that maximizes the average mutual information. Due to the high computational complexity required for implementing the optimal transmission strategy, we focus on certain practical transmission schemes, and develop necessary and sufficient conditions for these practical alternatives to achieve maximal average mutual information. Capacity-achieving transmission strategy is also investigated in [13], where the channels are assumed zero mean with arbitrary fading distribution.

Notation: Bold upper and lower case letters denote matrices and column vectors, respectively; $(\cdot)^T$ and $(\cdot)^{\mathcal{H}}$ stand for transpose and Hermitian transpose, respectively; $\mathrm{trace}(\cdot)$ denotes the matrix trace operation; $[\mathbf{A}]_{pq}$ stands for the (p, q)th element of matrix $\mathbf{A}$; $\mathbb{E}[\cdot]$ denotes ensemble average; $\sim$ denotes probability distribution equivalence; and $\mathcal{G}(a, b)$ stands for the Gamma distribution with parameters a and b.

II. SYSTEM AND CHANNEL MODEL

With M transmit-antennas and a single receive-antenna, we introduce the following partial CAI model for Multi-Input Single-Output (MISO) systems in the downlink. Let $\mathbf{h} := [h_1, \ldots, h_M]^T = [a_1 e^{j\theta_1}, \ldots, a_M e^{j\theta_M}]^T$ be the vector of channel coefficients, where a_i's are channel amplitudes and θ_i's are channel phases. We assume that:

(as1) $a_i \sim$ Nakagami $(m_i, \Omega_i), \forall\, i$,
(as2) $\theta_i \sim$ Uniform $[0, 2\pi), \forall\, i$,
(as3) a_i is independent of $\theta_{i'}, \forall\, i, i'$,
(as4) θ_i is independent of $\theta_{i'}, \forall\, i \neq i'$.

Notice that the channel amplitudes $\{a_i\}$ are allowed to be correlated and their distribution parameters $\{m_i, \Omega_i\}$ could be different from antenna to antenna. The Nakagami distribution captures a wide range of wireless fading channels and often gives the best fit to urban and indoor multipath propagation. With probability density function

$$p(a_i) = 2\left(\frac{m_i}{\Omega_i}\right)^{m_i} \frac{a_i^{2m_i-1}}{\Gamma(m_i)} e^{-m_i a_i^2/\Omega_i},$$

Manuscript received November 20, 2004; revised August 9, 2005 and February 13, 2006; accepted February 20, 2006. The associate editor coordinating the review of this paper and approving it for publication was J. Tugnait. Prepared through collaborative participation in the Communications and Networks Consortium sponsored by the U. S. Army Research Laboratory under the Collaborative Technology Alliance Program, Cooperative Agreement DAAD19-01-2-0011. The U. S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation thereon. Part of this work appeared in the *Proc. of the 38th Annual Conference on Information Sciences and Systems*, Princeton University, NJ, March 17-19, 2004.

P. Xia was with the Department of Electrical and Computer Engineering, University of Minnesota, and is now with Samsung Electronics, 75 W. Plumeria Dr., San Jose, CA 95134 (email: pengfei.xia@samsung.com).

G. B. Giannakis is with the Department of Electrical and Computer Engineering, University of Minnesota, 200 Union Street SE, Minneapolis, Minnesota 55455 (email: georgios@ece.umn.edu).

Digital Object Identifier 10.1109/TWC.2006.04765.

it also includes Rayleigh distribution as a special case corresponding to $m_i = 1$.

As in [2], we here assume that the channel phases are mutually independent and uniformly distributed across $[0, 2\pi)$, and the channel statistics remain invariant over the transmission period. Under these assumptions, the parameters $\{m_i, \Omega_i\}$, together with channel amplitude correlations if any, can be estimated in advance. Letting $\mathbf{x} := [x_1, \ldots, x_M]^T$ denote the channel input vector containing information symbols to be transmitted, the received symbol is

$$y = \sqrt{P_0}\ \mathbf{x}^{\mathcal{H}}\mathbf{h} + \nu, \tag{1}$$

where additive noise ν is complex circularly Gaussian distributed with zero-mean and variance N_0, and P_0 is the transmit power. We assume that perfect channel knowledge is available at the receiver, through e.g., training-based channel estimation. The instantaneous receive signal-to-noise-ratio (SNR) can be written as $\gamma = P_0\ \mathbf{h}^{\mathcal{H}}\boldsymbol{\Sigma}_{\mathbf{x}}\mathbf{h}/N_0$, where $\boldsymbol{\Sigma}_{\mathbf{x}} := \mathbb{E}[\mathbf{x}\mathbf{x}^{\mathcal{H}}]$ is the normalized input covariance matrix with power constraint $\mathrm{trace}(\boldsymbol{\Sigma}_{\mathbf{x}}) = 1$. The average received SNR in (1) can be expressed as

$$\bar{\gamma} := \mathbb{E}_{\mathbf{h}}[\gamma] = \rho \cdot \mathbb{E}_{\mathbf{x}}[\mathbf{x}^{\mathcal{H}}\mathbf{R}_{\mathbf{h}}\mathbf{x}], \tag{2}$$

where $\mathbf{R}_{\mathbf{h}} := \mathbb{E}_{\mathbf{h}}[\mathbf{h}\mathbf{h}^{\mathcal{H}}]$ is the channel correlation matrix and $\rho := P_0/N_0$ is the transmit SNR. Based on (as1) – (as4), it is straightforward to verify that $\mathbb{E}[\mathbf{h}] = \mathbf{0}$, and $\mathbb{E}[\mathbf{h}\mathbf{h}^{\mathcal{H}}] = \mathrm{diag}[\Omega_1, \ldots, \Omega_M]$. Particularly, $\mathbb{E}[h_i] = \mathbb{E}[a_i]\mathbb{E}[e^{j\theta_i}] = 0$, $\mathbb{E}[h_i h_i^*] = \mathbb{E}[a_i^2] = \Omega_i, \forall\ i$, and $\mathbb{E}[h_i h_\ell^*] = \mathbb{E}[a_i e^{j\theta_i} a_\ell e^{-j\theta_\ell}] = \mathbb{E}[a_i a_\ell] \cdot \mathbb{E}[e^{j\theta_i} e^{-j\theta_\ell}] = \mathbb{E}[a_i a_\ell] \cdot 0 = 0, \forall\ i \neq \ell$. Notice that $\mathbb{E}[e^{j\theta_i} e^{-j\theta_\ell}] = 0, \forall\ i \neq \ell$ due to (as4).

III. Optimizing Average Mutual Information

The average mutual information in our setup is known to be maximized by inputs that are zero mean, complex circularly Gaussian distributed [12]. For such inputs, it suffices to find the input covariance matrix $\boldsymbol{\Sigma}_{\mathbf{x}}$ to fully characterize the transmission strategy. Let $\mu(\boldsymbol{\Sigma}_{\mathbf{x}}, \mathbf{h}) = \log(1 + \rho\mathbf{h}^{\mathcal{H}}\boldsymbol{\Sigma}_{\mathbf{x}}\mathbf{h})$ denote the mutual information for a channel realization $\mathbf{h}$ and a fixed input covariance matrix $\boldsymbol{\Sigma}_{\mathbf{x}}$. To maximize the average mutual information, one needs to solve the following constrained optimization problem:

$$\begin{aligned} \max_{\boldsymbol{\Sigma}_{\mathbf{x}}} \quad & \mathbb{E}_{\mathbf{h}}\left[\mu(\boldsymbol{\Sigma}_{\mathbf{x}}, \mathbf{h})\right] = \mathbb{E}_{\mathbf{h}}\left[\log(1 + \rho\mathbf{h}^{\mathcal{H}}\boldsymbol{\Sigma}_{\mathbf{x}}\mathbf{h})\right] \\ \text{s.t.} \quad & \mathrm{trace}(\boldsymbol{\Sigma}_{\mathbf{x}}) = 1 \end{aligned} \tag{3}$$

where $\boldsymbol{\Sigma}_x \geq \mathbf{0}$ is positive semi-definite.

Direct optimization of (3) is difficult, since $\boldsymbol{\Sigma}_{\mathbf{x}}$ can have arbitrary structure. However, it is possible to simplify the optimization task with partial CAI at the transmitter. Specifically, we have:

Proposition 1: *Under (as1-as4), the average mutual information in MISO systems with partial CAI at the transmitter is maximized when* $\boldsymbol{\Sigma}_{\mathbf{x}} = \mathrm{diag}(\lambda_1, \ldots \lambda_M)$, *with* $\lambda_i \geq 0, \forall\ i$.

Proof: Consider the spectral decompositions $\boldsymbol{\Sigma}_{\mathbf{x}} = \mathbf{U}_{\mathbf{x}}\boldsymbol{\Lambda}_{\mathbf{x}}\mathbf{U}_{\mathbf{x}}^{\mathcal{H}}$, and $\boldsymbol{\Sigma}_{\mathbf{h}} = \mathbf{U}_{\mathbf{h}}\boldsymbol{\Lambda}_{\mathbf{h}}\mathbf{U}_{\mathbf{h}}^{\mathcal{H}}$, where $\mathbf{U}_{\mathbf{x}}, \mathbf{U}_{\mathbf{h}}$ are unitary, and $\boldsymbol{\Lambda}_{\mathbf{x}}, \boldsymbol{\Lambda}_{\mathbf{h}}$ are diagonal. Similar problems have been considered for complex Gaussian channels, where it has been shown that $\mathbf{U}_{\mathbf{x}} = \mathbf{U}_{\mathbf{h}}$ is required to maximize average mutual information, when a certain inequality is satisfied [14, eq. (10)]. In fact, this condition also holds true under partial CAI, which can be verified from the fact that $\mathbb{E}[h_i h_j^*] = 0,\ \forall\ i \neq j$. The proof in [14] can thus be applied to non-Gaussian channels as well. Therefore, to maximize average mutual information under partial CAI, $\mathbf{U}_{\mathbf{x}} = \mathbf{U}_{\mathbf{h}}$ is also required. On the other hand, the fact that $\boldsymbol{\Sigma}_{\mathbf{h}}$ is diagonal simply implies that $\boldsymbol{\Sigma}_{\mathbf{x}}$ has to be diagonal to maximize average mutual information. The condition $\lambda_i \geq 0, \forall\ i$ is obvious as $\boldsymbol{\Sigma}_{\mathbf{x}}$ must be positive semi-definite. ■

Thanks to Proposition 1, the original optimization problem (3) simplifies to

$$\begin{aligned} \max_{\{\lambda_1, \ldots, \lambda_M\}} \quad & \bar{\mu} = \mathbb{E}_{\mathbf{h}}[\log(1 + \rho \textstyle\sum_{i=1}^M \lambda_i \gamma_i)] \\ \text{s.t.} \quad & \textstyle\sum_{i=1}^M \lambda_i = 1, \lambda_i \geq 0, \forall\ i, \end{aligned} \tag{4}$$

where we have defined $\gamma_i := a_i^2, \forall\ i$, and the power constraint translates to the simpler form $\sum_{i=1}^M \lambda_i = 1$. We will call $[\lambda_1, \ldots, \lambda_M]$ a *valid* solution if it satisfies the constraints in (4). Notice that γ_i is Gamma distributed, $\gamma_i \sim \mathcal{G}(m_i, \Omega_i/m_i)$ [9], where Ω_i/m_i captures the average received power and m_i parameterizes the degree of fading. In the following, we will investigate first the simple case where channel amplitudes are independent before the correlated case is addressed.

A. Independent channel amplitudes

For notation brevity, let us assume in this subsection that $\mathbb{E}[\gamma_1] = \Omega_1 > \mathbb{E}[\gamma_2] = \Omega_2 > \ldots > \mathbb{E}[\gamma_M] = \Omega_M$.

A-1) Necessary and sufficient condition for strongest antenna selection to be optimal:

The objective function in (4) can be expressed as:

$$\bar{\mu} = \int_{\gamma_1} \ldots \int_{\gamma_M} \log(1 + \rho \textstyle\sum_{i=1}^M \lambda_i \gamma_i) \times \\ f_1(\gamma_1) \ldots f_M(\gamma_M) d\gamma_1 \ldots d\gamma_M,$$

where $f_i(\gamma_i)$ denotes the pdf of γ_i. To find the optimal solution, an $M-1$ dimensional numerical search is required with complexity increasing exponentially.

We are therefore motivated to investigate simple yet practical transmission alternatives. The simple scheme of selecting the strongest antenna is overall optimal in terms of maximizing average SNR [10]. Notice that at low SNR, optimizing average mutual information is approximately equivalent to maximizing average SNR, since $\mathbb{E}[\log(1 + \rho\mathbf{h}^{\mathcal{H}}\boldsymbol{\Sigma}_{\mathbf{x}}\mathbf{h})] \approx \mathbb{E}[\rho\mathbf{h}^{\mathcal{H}}\boldsymbol{\Sigma}_{\mathbf{x}}\mathbf{h}]$ when ρ is small enough. It is thus of interest to investigate conditions under which strongest antenna selection achieves maximal average mutual information. Regarding this, we can show that:

Proposition 2: *For MISO systems under (as1-as4), the necessary and sufficient condition for strongest antenna selection to be overall optimal with CAI available at the transmitter is:*

$$\mathbb{E}\left[\frac{1 + \rho\gamma_2}{1 + \rho\gamma_1}\right] \leq 1. \tag{5}$$

Proof: Recall that for the strongest antenna selection, we have $\lambda_1 = 1$ and $\lambda_2 = \ldots = \lambda_M = 0$. For this simple solution to be optimal, it is necessary for $\bar{\mu}$ to decrease as we begin to shift power from the first antenna to any other antenna. Mathematically, it is necessary to have

$$\begin{aligned}&\left.\frac{d\bar{\mu}}{d\lambda_i}\right|_{(\lambda_1=1,\lambda_2=\ldots=\lambda_M=0)}\\&=\left.\left(\frac{\partial\bar{\mu}}{\partial\lambda_i}-\frac{\partial\bar{\mu}}{\partial\lambda_1}\right)\right|_{(\lambda_1=1,\lambda_2=\ldots=\lambda_M=0)}\leq 0,\forall\ i\geq 2,\end{aligned}\tag{6}$$

where the partial derivatives are to be evaluated by setting $\lambda_1 = 1$ and all other $\lambda_{i\neq 1} = 0$.

Evaluating (6), we arrive at the necessary condition:

$$\mathbb{E}\left[\frac{1+\rho\gamma_i}{1+\rho\gamma_1}\right]\leq 1,\forall\ i\geq 2.\tag{7}$$

On the other hand,

$$\begin{aligned}&\frac{\partial^2\bar{\mu}}{\partial\lambda_i^2}=-\mathbb{E}\left[\frac{\rho^2\gamma_i^2}{(1+\rho(\lambda_1\gamma_1+\ldots+\lambda_M\gamma_M))^2}\right]\leq 0,\\&\quad i=2,\ldots,M,\end{aligned}\tag{8}$$

are always true for any *valid* solution $[\lambda_1,\ldots,\lambda_M]$, because the numerator and denominator inside the expectation operator are always non-negative. Therefore, Eq. (7) is not only necessary, but also sufficient. Recalling that $\mathbb{E}[\gamma_1]\geq\mathbb{E}[\gamma_2]\geq\ldots\geq\mathbb{E}[\gamma_M]$ and the fact that γ_i's are independent, we can easily simplify (7) to (5). ■

The necessary and sufficient condition (5) can be explicitly expressed as:

$$\left(\frac{\rho\Omega_1}{m_1}\right)^{-m_1}e^{\frac{m_1}{\rho\Omega_1}}\Gamma(1-m_1,\frac{m_1}{\rho\Omega_1})(1+\rho\Omega_2)\leq 1,\tag{9}$$

where $\Gamma(\ \cdot,\ \cdot)$ is the incomplete Gamma function. Actually when $m_1 = 1$, (9) simplifies to [5, Eq. (3)], which is exactly the necessary and sufficient condition for transmit beamforming to maximize average mutual information with partial channel covariance information at the transmitter. Therefore, (9) can be deemed as a generalization of [5, eq.(3)] from Rayleigh fading channels with $m_i = 1,\forall\ i$, to Nakagami fading channels with generally unequal m_i's. Notice that condition (9) depends only on the system parameters $\rho, m_1, \Omega_1, \Omega_2$, but not on other distribution parameters such as $m_3, \Omega_3, \ldots, m_M, \Omega_M$; and particularly, not on m_2 of the second strongest antenna.

The necessary and sufficient condition (9) is depicted in Fig. 1 for different values of m_1. With $m_1 = 1$ for example, strongest antenna selection is optimal if and only if the $(\rho\Omega_1, \rho\Omega_2)$ pair lies in the area between the solid separating curve and the horizontal axis. We observe that as m_1 increases, the optimal region expands accordingly; and therefore, the strongest antenna selection scheme becomes increasingly attractive. In fact, the optimal region expands to the entire feasible set $\{\rho\Omega_1\geq\rho\Omega_2\geq 0\}$ as $m_1\to\infty$. We also notice that the separating curve behaves more and more like a straight line as m_1 increases, especially at high SNR. Dependence of the necessary and sufficient condition on ρ can then be approximately removed at high SNR.

A-2) Necessary and sufficient condition for strongest k-antenna selection to be optimal:

When the necessary and sufficient condition in (9) is not satisfied, sticking to the simple solution of strongest antenna

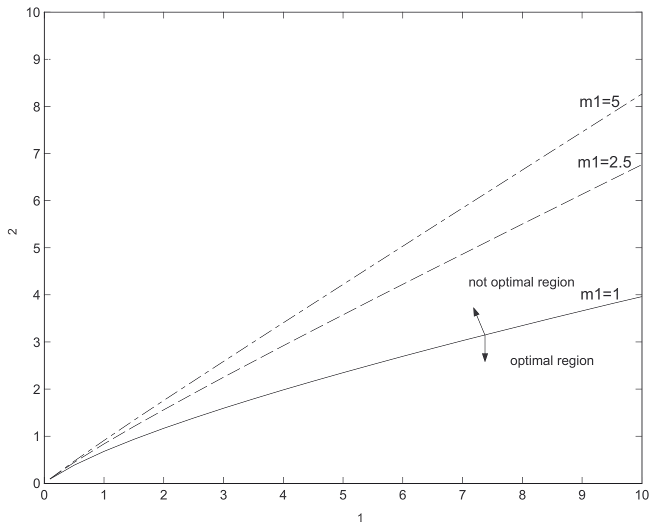

Fig. 1. Necessary and sufficient conditions for strongest antenna selection to be overall optimal (independent channels).

selection may degrade performance. One possibility is to restrict the transmitter to use the strongest k ($2\leq k\leq M$) antennas for transmission, which amounts to forcing $\lambda_{k+1} = \ldots = \lambda_M = 0$. A $(k-1)$-dimensional numerical search is then needed to find the *locally optimal solution*, i.e., the transmission strategy that maximizes average mutual information under the additional constraint that $\lambda_{k+1} = \ldots = \lambda_M = 0$. Similarly, we expect that the k-antenna-selection scheme also has potential to be overall optimal. Notice that for it to be optimal, successive decoding is needed at the receiver side. To obtain the necessary and sufficient condition for k-antenna-selection to be overall optimal, we need the following lemma.

Lemma 1: *When the transmitter is restricted to use only the strongest k antennas, the solution $\boldsymbol{\lambda}_k^o := [\lambda_1^o,\ldots,\lambda_k^o,0,\ldots,0]$ is locally optimal if and only if*

$$\mathbb{E}\left[\frac{1+\rho\sum_{i=1}^k\lambda_i\gamma_i}{1+\rho\sum_{i=1}^k\lambda_i^o\gamma_i}\right]\leq 1\tag{10}$$

holds true for any valid k-antenna-selection solution $[\lambda_1,\ldots,\lambda_k,0,\ldots,0]$.

Proof: The proof is given in [6].

Intuitively, if we look at the random ratio of the receive power required by any power allocation over that required by the optimal power allocation (slightly adjusted due to the sum operations in the numerator and the denominator), the average of this ratio should be less than or equal to 1. Based on Lemma 1, we have the following result:

Proposition 3: *For MISO systems under (as1-as4), let $\boldsymbol{\lambda}_k^o := [\lambda_1^o,\ldots,\lambda_k^o,0,\ldots,0]$ be the locally optimal solution when the transmitter chooses at most k antennas. The necessary and sufficient condition for $\boldsymbol{\lambda}_k^o$ to be overall optimal is*

$$\mathbb{E}\left[\frac{1+\rho\gamma_{k+1}}{1+\rho\sum_{i=1}^k\lambda_i^o\gamma_i}\right]\leq 1.\tag{11}$$

Proof: Suppose $\boldsymbol{\lambda}_k^o$ is indeed the overall optimal solution. Let $\boldsymbol{\lambda} := [\lambda_1,\ldots,\lambda_M]$ denote any other *valid* solution, and $\mathbf{e} := [e_1,\ldots,e_M] = \boldsymbol{\lambda}-\boldsymbol{\lambda}_k^o$ with $\sum_{i=1}^M e_i = 0$. We call $\mathbf{e} = \boldsymbol{\lambda}-\boldsymbol{\lambda}_k^o$

a valid difference vector if $\boldsymbol{\lambda}$ is a valid solution. Now, consider $\boldsymbol{\lambda}_k^o + q\mathbf{e}$, with q being an arbitrary real number between 0 and 1. It is easy to verify that $\boldsymbol{\lambda}_k^o + q\mathbf{e}$ is also a *valid* solution. On the other hand, since $\boldsymbol{\lambda}_k^o$ is overall optimal, we have

$$\left.\frac{d\,\mathbb{E}[\log(1+\rho\sum_{i=1}^M(\lambda_i^o+qe_i)\gamma_i)]}{d\,q}\right|_{q=0} \le 0, \quad (12)$$

for all valid difference vectors $\{\mathbf{e}\}$. Evaluating (12), we arrive at

$$\mathbb{E}\left[\frac{\rho\sum_{i=1}^M \lambda_i\gamma_i}{1+\rho\sum_{i=1}^k \lambda_i^o\gamma_i}\right] \le \mathbb{E}\left[\frac{\rho\sum_{i=1}^k \lambda_i^o\gamma_i}{1+\rho\sum_{i=1}^k \lambda_i^o\gamma_i}\right], \quad (13)$$

which can be further simplified as

$$\mathbb{E}\left[\frac{\rho\sum_{i=k+1}^M \lambda_i\gamma_i}{1+\rho\sum_{i=1}^k \lambda_i^o\gamma_i}\right] \le \mathbb{E}\left[\frac{\rho\sum_{i=1}^k (\lambda_i^o-\lambda_i)\gamma_i}{1+\rho\sum_{i=1}^k \lambda_i^o\gamma_i}\right]. \quad (14)$$

Notice that (14) is true for all *valid* solutions. The necessary condition (11) follows directly by setting $\lambda_{k+1} = 1$, which implies $\lambda_1 = \ldots = \lambda_k = 0$, and $\lambda_{k+2} = \ldots = \lambda_M = 0$.

We next show that condition (11) is also sufficient, i.e., the locally optimal solution $\boldsymbol{\lambda}_k^o$ is indeed overall optimal when eq. (11) is satisfied. From Lemma 1, we have

$$\mathbb{E}\left[\frac{1+\rho\gamma_j}{1+\rho\sum_{i=1}^k \lambda_i^o\gamma_i}\right] \le 1, \forall\ 1 \le j \le k, \quad (15)$$

as $\boldsymbol{\lambda}_k^o$ is locally optimal. Furthermore, we have from (11) that

$$\mathbb{E}\left[\frac{1+\rho\gamma_{k+l}}{1+\rho\sum_{i=1}^k \lambda_i^o\gamma_i}\right] \le 1, \forall\ 1 \le l \le M-k, \quad (16)$$

because $\mathbb{E}[\gamma_{k+1}] \ge \mathbb{E}[\gamma_{k+2}] \ge \ldots \ge \mathbb{E}[\gamma_M]$, and γ_i's are independent. Basically, equations (15) and (16) consider a special scheme allocating all power to only one antenna, say antenna j where $1 \le j \le M$. If we look at the random ratio of the receive power by this scheme over that of the optimal power allocation, the average ratio can not exceed 1.

Therefore, for any *valid* solution $\boldsymbol{\lambda} = [\lambda_1, \ldots, \lambda_M]$, we have

$$\begin{aligned} &\mathbb{E}\left[\frac{1+\rho\sum_{j=1}^M \lambda_j\gamma_j}{1+\rho\sum_{i=1}^k \lambda_i^o\gamma_i}\right] \\ &= \sum_{j=1}^M \lambda_j \mathbb{E}\left[\frac{1+\rho\gamma_j}{1+\rho\sum_{i=1}^k \lambda_i^o\gamma_i}\right] \le \sum_{j=1}^M \lambda_j = 1. \end{aligned} \quad (17)$$

According to Lemma 1, the solution $\boldsymbol{\lambda}_k^o = [\lambda_1^o, \ldots, \lambda_k^o,\ 0, \ldots, 0]$ is indeed overall optimal. ■

As before, the necessary and sufficient condition in (11) depends only on parameters $\rho, m_1, \Omega_1, \ldots, m_k, \Omega_k, \Omega_{k+1}$, but not on $m_{k+1}, m_{k+2}, \Omega_{k+2}, \ldots, m_M, \Omega_M$. Furthermore, by setting $m_1 = \ldots = m_M = 1$ in (11), we can easily obtain the necessary and sufficient condition for k-dimensional transmit beamforming to achieve maximal average mutual information with partial CCI at the transmitter.

B. Correlated channel amplitudes

In this subsection, we investigate the more complicated case where channel amplitudes are correlated. To find the optimal power allocation across all eigen-beams, an $M-1$ dimensional search is still required, which is not desirable from a complexity point of view. This again motivates us to investigate how strongest antenna selection performs in the correlated scenario. For notational brevity, we assume in this subsection $m_1 \ge m_2 \ge \ldots \ge m_M$, and define $n_i := m_i - m_{i+1}, \forall\ i = 1, \ldots, M-1$ and $n_M := m_M$. Let $\boldsymbol{\Sigma}_\gamma$ be the covariance matrix of $[\gamma_1, \gamma_2, \ldots, \gamma_M]^T$, with $[\boldsymbol{\Sigma}_\gamma]_{pq} := \mathbb{E}[\gamma_p\gamma_q] - \mathbb{E}[\gamma_p]\mathbb{E}[\gamma_q]$ and define matrix $\mathbf{A}$ such that $[\boldsymbol{\Sigma}_\gamma]_{pq} = \min(m_p, m_q)[A]_{pq}^2$.

Lemma 2: *For a group of correlated Gamma random variables $\gamma_1, \ldots, \gamma_M$ whose covariance matrix is $\boldsymbol{\Sigma}_\gamma$, and a group of non-negative real coefficients $\lambda_1, \ldots, \lambda_M$, the moment generating function (MGF) of $\sum_{i=1}^M \lambda_i\gamma_i$ is*

$$\Phi(s) = \prod_{k=1}^M \det(\mathbf{I}_k - s\tilde{\mathbf{A}}_k)^{-n_k} = \prod_{k=1}^M \prod_{l=1}^k (1 - s\alpha_{kl})^{-n_k},$$

where $\tilde{\mathbf{A}}_k := \boldsymbol{\Lambda}_k^{1/2}\mathbf{A}_k\boldsymbol{\Lambda}_k^{1/2}$, $\boldsymbol{\Lambda}_k :=$ diag$(\lambda_1, \ldots, \lambda_k)$, $\mathbf{A}_k := \mathbf{A}[1:k, 1:k]$ is the $k\times k$ leading submatrix of $\mathbf{A}$ in the upper left corner, and α_{kl} is the lth eigenvalue of matrix $\tilde{\mathbf{A}}_k$.

Proof: This lemma is a direct extension of [16, Property 4]. ■

Lemma 2 shows that the weighted sum of M correlated Gamma random variables can be written equivalently as a weighted sum of $M(M+1)/2$ independent Gamma random variables, i.e., $\sum_{i=1}^M \lambda_i\gamma_i \sim \sum_{k=1}^M\sum_{l=1}^k \alpha_{kl}w_{kl}$, where random variables $\{w_{kl}\} \sim \mathcal{G}(n_k, 1)$ are independent, and $\{\alpha_{kl}\}$ are functions of $\lambda_1, \ldots, \lambda_M$. Therefore, the objective function $\bar{\mu}$ in (4) can be written as:

$$\begin{aligned}\bar{\mu} = \int_{w_{11}} \ldots \int_{w_{MM}} &\log(1+\rho\textstyle\sum_{k=1}^M\sum_{l=1}^k \alpha_{kl}w_{kl}) \times \\ &\textstyle\prod_{k=1}^M\prod_{l=1}^k f_{kl}(w_{kl})\, dw_{11}\ldots dw_{MM},\end{aligned} \quad (18)$$

where $f_{kl}(w_{kl})$ is the pdf of w_{kl}. The major difficulty with (18) is that $\{\alpha_{kl}\}$ may not necessarily be expressible in closed form in terms of $\lambda_1, \ldots, \lambda_M$.

In the following, we will investigate a two-input single-output setup. Such a study provides helpful insights on how various system parameters influence the overall optimal solution in the correlated scenario. Define the correlation coefficient between γ_1 and γ_2 as

$$t = \frac{\mathbb{E}[\gamma_1\gamma_2] - \mathbb{E}[\gamma_1]\mathbb{E}[\gamma_2]}{\sqrt{(\mathbb{E}[\gamma_1^2] - (\mathbb{E}[\gamma_1])^2)\,(\mathbb{E}[\gamma_2^2] - (\mathbb{E}[\gamma_2])^2)}}, \quad (19)$$

where $0 \le t \le 1$. It then follows that

$$\mathbf{A} = \begin{bmatrix} \frac{\Omega_1}{m_1} & \sqrt{\frac{t\Omega_1\Omega_2}{\sqrt{m_1m_2^3}}} \\ \sqrt{\frac{t\Omega_1\Omega_2}{\sqrt{m_1m_2^3}}} & \frac{\Omega_2}{m_2} \end{bmatrix}.$$

Furthermore, we have $\alpha_{11} = \lambda_1 A_{11}$, and

$$\begin{aligned}\alpha_{21}, \alpha_{22} = (1/2)\Big[&\lambda_1A_{11} + \lambda_2A_{22} \pm \\ &\sqrt{(\lambda_1A_{11}+\lambda_2A_{22})^2 - 4\lambda_1\lambda_2(A_{11}A_{22} - A_{12}A_{21})}\Big].\end{aligned} \quad (20)$$

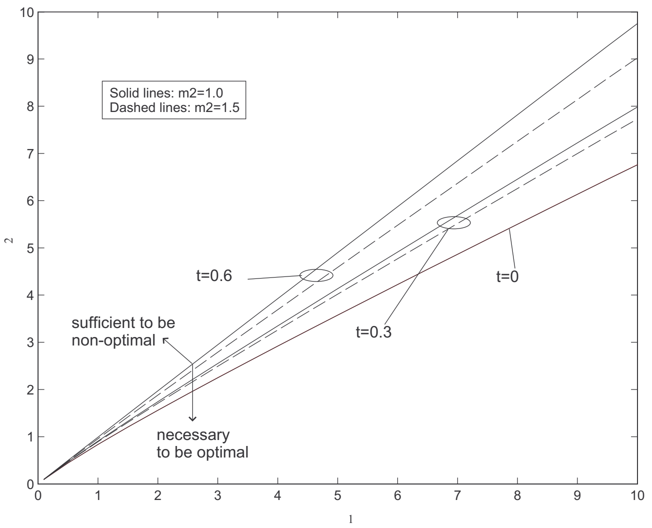

Fig. 2. Necessary conditions for strongest antenna selection to be overall optimal (correlated channels).

To proceed, we need to distinguish between two possible cases: $\Omega_1 \geq \Omega_2$ and $\Omega_1 < \Omega_2$. Here we consider only the former, while the latter case follows similarly. When $\Omega_1 \geq \Omega_2$, strongest antenna selection amounts to setting $\lambda_1 = 1, \lambda_2 = 0$. It is therefore necessary to have

$$\left.\frac{d\bar{\mu}}{d\lambda_2}\right|_{\lambda_1=1,\lambda_2=0} = \left.\frac{\partial\bar{\mu}}{\partial\lambda_2} - \frac{\partial\bar{\mu}}{\partial\lambda_1}\right|_{\lambda_1=1,\lambda_2=0} \leq 0, \qquad (21)$$

to ensure that strongest antenna selection is optimal. Evaluating (21), we obtain

$$\begin{aligned} &\rho\Omega_2 \cdot t\sqrt{\frac{m_1}{m_2^3}}\mathbb{E}\left[\frac{w_{21}-w_{22}}{1+\rho\frac{\Omega_1}{m_1}(w_{11}+w_{21})}\right] \\ &+\frac{\rho\Omega_2}{m_2}\mathbb{E}\left[\frac{w_{22}}{1+\frac{\rho\Omega_1}{m_1}(w_{11}+w_{21})}\right] \qquad (22)\\ &-\frac{\rho\Omega_1}{m_1}\mathbb{E}\left[\frac{w_{11}+w_{21}}{1+\frac{\rho\Omega_1}{m_1}(w_{11}+w_{21})}\right] \leq 0. \end{aligned}$$

When $t = 0$, i.e., when γ_1, γ_2 are independent from each other, (22) boils down to (5).

Different from the independent channel amplitude scenario in subsection III-A, the condition (22) now depends on m_2, as well as other distribution parameters. Also notice that the condition now is necessary for strongest antenna selection to be optimal, but may not be sufficient.

In Fig. 2, we take $m_1 = 2.5$, $\Omega_1 \geq \Omega_2$, and investigate the impact of m_2 and the correlation coefficient t. For strongest antenna selection to be overall optimal in this case, it is necessary for the $(\rho\Omega_1, \rho\Omega_2)$ pair to lie below the separating curve. It is observed that the necessary region grows as t increases, which suggests that it is more likely for strongest antenna selection to be optimal when the channel amplitudes are correlated rather than independent. On the other hand, the necessary region shrinks as m_2 increases for a certain fixed t. This reveals that it is more attractive to use strongest antenna selection when the distribution parameter m_2 of the weaker channel is smaller.

IV. Conclusions

Relying on a partial CAI model, we investigated the optimal transmission strategy that maximizes the average mutual information of a MISO system. We focused on the scheme of multiple antenna selection, and developed necessary and sufficient conditions for these practical alternatives to maximize the average mutual information. Both independent and correlated cases were investigated, and revealed helpful insights on how various system parameters influence the overall optimal transmission strategy [1].

References

[1] A. Abouraddy, and S. Elnoubi, "Statistical modeling of the indoor radio channel at 10 GHz through propagation measurements–part I: narrow-band measurements and modeling," *IEEE Trans. Veh. Technol.*, vol. 49, no. 5, pp. 1491–1507, Sept. 2000.

[2] N. Beaulieu and J. Cheng, "Precise error-rate analysis of bandwidth-efficient BPSK in Nakagami fading and co-channel interference," *IEEE Trans. Commun.*, vol. 52, no. 1, pp. 149–158, Jan. 2004.

[3] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Sel. Areas Commun.*, vol. 21 no. 5, pp. 684–702, June 2003.

[4] D. A. Gore and A. J. Paulraj, "MIMO antenna subset selection with space-time coding," *IEEE Trans. Signal Processing*, vol. 50, no. 10, pp. 2580–2588, Oct. 2002.

[5] S. A. Jafar and A. Goldsmith, "On optimality of beamforming for multiple antenna systems with imperfect feedback," in *Proc. IEEE International Symposium on Information Theory 2001*, pp. 321–321.

[6] E. Jorswieck and H. Boche, "Optimal transmission with imperfect channel state information at the transmit antenna array," *Kluwer Wireless Personal Commun.*, vol. 27, no. 1, pp. 33-56, Oct. 2003.

[7] E. Jorswieck and H. Boche, "Channel capacity and capacity-range of beamforming in MIMO wireless systems under correlated fading with covariance feedback," *IEEE Trans. Wireless Commun.*, vol. 3, no. 5, pp. 1543-1553, Sept. 2004.

[8] A. Moustakas and S. Simon, "Optimizing multiple-input single-output (MISO) communication systems with general Gaussian channels: non-trivial covariance and nonzero mean," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2770–2780, Oct. 2003.

[9] M. Nakagami, "The *m*-distribution–a general formula of intensity distribution of rapid fading," *Statistical Methods in Radio Wave Propagation*, W. C. Hoffman, Ed. New York: Elmsford, 1960.

[10] A. Narula, M. J. Lopez, M. D. Trott, and G. W. Wornell, "Efficient use of side information in multiple-antenna data transmission over fading channels," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1423–1436, Oct. 1998.

[11] M. Payaro, X. Mestre, and M. A. Lagunas, "Optimum transmit architecture of a MIMO system under modulus channel knowledge at the transmitter," in *Proc. IEEE Information Theory Workshop 2004*.

[12] I. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. Telecom.*, vol. 10, no. 6, pp. 585–596, Nov. 1999.

[13] A. Tulino, A. Lozano, and S. Verdú, "Capacity-achieving input covariance for correlated multi-antenna channels," in *Proc. 41st Annual Allerton Conference on Communication, Control, and Computing 2003*, pp. 242–251.

[14] E. Visotsky and U. Madhow, "Space-time transmit precoding with imperfect feedback," *IEEE Trans. Inf. Theory*, vol. 47, no. 6, pp. 2632–2639, Sept. 2001.

[15] P. Xia and G. B. Giannakis, "Conditions for multi-antenna selection to be optimal given channel amplitude-only information," in *Proc. 38th Conference on Information Sciences and Systems 2004*.

[16] Q. T. Zhang, "A generic correlated Nakagami fading model for wireless communications," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1745–1748, Nov. 2003.

[17] S. Zhou and G. B. Giannakis, "Optimal transmitter Eigen-beamforming and space-time block coding based on channel correlations," *IEEE Trans. Inf. Theory*, vol. 49, no. 7, pp. 1673–1690, July 2003.

[1]The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U. S. Government.