MIMO Channels in the Low SNR Regime: Communication Rate, Error Exponent and Signal Peakiness

Xinzhou Wu and R. Srikant
Department of Electrical and Computer Engineering
and
Coordinated Science Lab
University of Illinois at Urbana-Champaign
xwu@uiuc.edu, rsrikant@uiuc.edu

Abstract — **We consider non-coherent MIMO fading channels and characterize the reliability function in the low-SNR regime as a function of the number of transmit and receive antennas. We assume no CSI is available at the transmitter or the receiver. For the case when the fading matrix H has independent entries, we show that the number of transmit antennas plays a key role in reducing the peakiness in the input signal required to achieve the optimal error exponent for a given communication rate. Further, by considering a correlated channel model, we show that the maximum performance gain (in terms of the error exponent and communication rate) is achieved when the entries of the channel fading matrix are fully correlated.**

I. Introduction

In this paper, we study multiple input and multiple output (MIMO) antenna channels in the low-SNR regime. Specifically, we compute the reliability function for such channels and also characterize the signal peakiness required to achieve the maximum error exponent for a given communication rate. We use a block-fading model, which is widely used in the MIMO literature [6, 13], and assume that the fading matrix $\mathbf{H}$ remains constant for a certain time period and then changes to an independent value for the next time period. The length of this time period is usually referred to as the *coherence time* of the channel. Letting M be the number of transmit antennas, N be the number of receive antennas and T_c be the coherence time, the channel we study in this paper is

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{W}, \tag{1}$$

where $\mathbf{Y}, \mathbf{W} \in \mathcal{C}^{N\times T_c}$, $\mathbf{X} \in \mathcal{C}^{M\times T_c}$ and $\mathbf{H} \in \mathcal{C}^{N\times M}$. Here $\mathbf{W}$ is a proper complex Gaussian noise random matrix, which has i.i.d. entries, i.e., $w_{nt} \sim \mathcal{C}N(0,1)$, for $1 \leq n \leq N$ and $1 \leq t \leq T_c$, where w_{nt} denotes the $(n,t)^{\text{th}}$ entry of $\mathbf{W}$. We also assume $\mathbf{H}$ has symmetric complex Gaussian entries such that $h_{nm} \sim \mathcal{C}N(0,1)$, for $1 \leq m \leq M$ and $1 \leq n \leq N$. In other words, we consider Rayleigh fading MIMO channels in this paper. At this point, we do not make any assumptions regarding the correlation among the entries of $\mathbf{H}$. Later, we will consider two cases, one where the entries of $\mathbf{H}$ are independent and the other where they are correlated.

Throughout this paper, we assume that neither the transmitter nor the receiver has knowledge of the realization of the random matrix $\mathbf{H}$. However, we always assume the distribution of $\mathbf{H}$ is known *a priori* at both the transmitter and the receiver. We will study the reliability function of the channel (1) in the low-SNR regime under the following standard average power constraint:

$$E[\|\mathbf{X}\|^2] \leq pT_c, \tag{2}$$

where $\|\mathbf{X}\|$ is the *Frobenius norm* [2] of the matrix $\mathbf{X}$. A special case of this model is the case when $M = N = T_c = 1$, which we will refer to as a SISO fast-fading channel.

The reliability function of a channel is defined as

$$E(R,p) = \lim_{N_c\to\infty} -\frac{\ln P_e(N_c,R)}{N_cT_c}, \tag{3}$$

where $P_e(N_c,R)$ is the minimum average probability of decoding error for any block code with block length N_c and coding rate R, which satisfies the average power constraint (2). The rate R is defined in nats per symbol and is equal to $\frac{\ln M_c}{N_cT_c}$, where M_c denotes the size of the code book, and N_c is the codeword length measured in number of channel uses, where each channel use consists of the transmission of T_c symbols.

It is well known that with multiple antennas at the receiver and/or the transmitter, the channel performance can be improved. The extent of this improvement is now well understood in the high SNR regime. From a capacity point of view, when the elements of $\mathbf{H}$ are i.i.d. and known at the receiver, we can achieve a degree of freedom gain (or multiplexing gain) up to $\min(M,N)$ [6]. From a probability of decoding error point of view, we can achieve a *diversity gain* up to MN with space-time coding at the input [5]. A recent work by Zheng and Tse [13] illustrated the tradeoff between these two gains in the high-SNR regime by showing the best degree of freedom gain and diversity gain that can be achieved *simultaneously*. They also extend this result to a non-coherent MIMO channel in a later work [14]. Roughly speaking, in an error-exponent theory framework, their results characterize the following curve

$$d^*(r) = \lim_{p\to\infty} \frac{E(r\ln p, p)}{\ln p},$$

where the degree of freedom gain r can be interpreted as the normalized communication rate at high SNR and the diversity gain $d^*(r)$ is nothing but the normalized reliability function at rate r.

In the low-SNR regime, i.e., the average power constraint $p \to 0$, both the reliability function and the capacity approach 0 linearly in p. Thus, it is natural to define the normalized reliability function in the low-SNR regime as follows:

$$E^{lp}(r) = \lim_{p \to 0} \frac{E(rp, p)}{p}, \quad \textit{if the limit exists.} \tag{4}$$

Here the superscript stands for *low power*. For simplicity, from now on, we will refer to $E^{lp}(r)$ as the low-SNR reliability function, and r as the low-SNR communication rate. However, it should be clear that both quantities are normalized by p in the low-SNR regime. It turns out that having multiple antennas in the low-SNR regime, as compared to the SISO fast-fading channels $(M = N = T_c = 1)$, can provide two kinds of gains: *performance gain* and *peakiness gain*. The performance gain is defined as follows.

Definition 1 *We say a MIMO block-fading channel has a performance gain of L if*

$$E^{lp}(r) = L\bar{E}^{lp}(\frac{r}{L}),$$

where $\bar{E}^{lp}$ denotes the reliability function for a SISO fast-fading channel. ◇

Note that the performance gain as defined above provides simultaneous gains in both the error exponent and the rate. In other words, a performance gain of L indicates that the transmission rate has increased by a factor of L *and* the error exponent has increased by a factor of L as compared to a SISO channel.

In the low-SNR regime, it is well known that a "peaky" signaling scheme is required to achieve capacity [7, 3, 4, 10], in the sense that, as the average power goes to zero, the optimal signaling scheme is to keep silent most of the time and use symbols of very large amplitude (going to infinity as we approach capacity) when transmitting. In this paper, we will show that a similar scheme is required to achieve the optimal decoding error probability at an arbitrary communication rate between 0 and capacity and we define a quantity $\tau(r)$, which we call the *minimum signal peakiness*, to denote how peaky the signaling has to be to achieve the optimal decoding error probability at rate r. (We will precisely define this quantity in Section III.) The peakiness gain is defined as follows.

Definition 2 *We say a MIMO block-fading channel, which has a performance gain of L, has a peakiness gain of G if the* minimum signaling peakiness $\tau(r)$ *required to achieve the optimal decoding error probability at communication rate r satisfies*

$$\tau(r) = \frac{1}{G}\bar{\tau}(\frac{r}{L}),$$

where $\bar{\tau}(r)$ denotes the minimum signal peakiness at rate r for a SISO fast-fading channel.

The performance gain tells us how much we can improve the channel performance (both the capacity and the reliability function) by using multiple antennas. On the other hand, since signals with very large amplitudes are undesirable in practice, the peakiness gain quantifies the effect of using multiple antennas in reducing the peakiness necessary to achieve a certain performance level.

II. Preliminaries

In this section, we present a theorem characterizing the low SNR reliability function for an arbitrary memoryless channel by extending the results in [1] to allow less restrictive input distributions. This theorem will be used extensively in the rest of the paper.

We consider a general discrete-time memoryless channel with an arbitrary input alphabet A, and its output determined by the transition probability function $f(y|x)$. Further, we have a cost function $b : A \to R^+$ associated with each input symbol $x \in A$, and assume that the input to the channel is constrained by the following average cost constraint

$$E[b(x)] \leq p. \tag{5}$$

Next we specify the additional constraint on the input distributions besides the average power constraint. Throughout this paper, we only consider input distributions with a discrete and finite alphabet. Specifically, at a SNR lever p, we constrain the input distributions to be in the following set

Definition 3 *Define* $\mathcal{D}(p) = \{q(x) : E[b(x)] = p$*; support of q(x) is an arbitrary finite set of discrete points in* $A\}$.

Under this constraint on the input distribution, our goal in this section is to find the low-SNR reliability function $E^{lp}(r)$ of this channel model, as defined by (4), by borrowing tools from both [1] and [9]. For simplicity, in this paper, we assume that a *unique* zero-cost symbol always exists in the input alphabet, i.e., $\exists x \in A$ satisfying $b(x) = 0$. Without loss of generality, we let 0 be the zero-cost symbol and define $A' = A \backslash \{0\}$.

Our main theorem in this section is a generalization of Gallager's result in [1] and is stated below.

Theorem 1 *The low-SNR reliability function $E^{lp}(r)$ satisfies*

$$E^{lp}_r(r) \leq E^{lp}(r) \leq \min(E^{lp}_{sp}(r), E^{lp}_{sl}(r)). \tag{6}$$

Here E^{lp}_{sp} and E^{lp}_r are defined as the following

$$E^{lp}_r(r) = \sup_{0 \leq \rho \leq 1} -\rho r + \tilde{E}_o(\rho); \tag{7}$$

$$E^{lp}_{sp}(r) = \sup_{\rho \geq 0} -\rho r + \tilde{E}_o(\rho), \tag{8}$$

and $\tilde{E}_o(\rho)$ is characterized by the following equation

$$\tilde{E}_o(\rho) = \sup_{x \in A'} \frac{-(1+\rho)\ln \int f(y|0)^{\frac{\rho}{1+\rho}} f(y|x)^{\frac{1}{1+\rho}} dy}{b(x)}. \tag{9}$$

The low-SNR straight-line bound $E_{sl}^{lp}(r)$ is the smallest linear function of r which touches $E_{sp}^{lp}(r)$ and satisfies

$$E_{sl}^{lp}(0) = \sup_q \frac{E_x\left[-\ln\int f(y|x_1)^{\frac{1}{2}} f(y|x_2)^{\frac{1}{2}} dy\right]}{E_x[b(x)]}, \qquad (10)$$

and the $\sup$ *is over all possible probability distributions with a discrete and finite alphabet set. Further, if*

$$E_{sl}^{lp}(0) \leq \tilde{E}_o(1), \qquad (11)$$

then $E^{lp}(r) = E_r^{lp}(r)$ for all rates.

Proof: All proofs are omitted in this version of the paper. ◇

Remarks: The low-SNR bounds here have similar forms as the bounds in [1]. However, in this paper, we are allowing the input distribution to use symbols as a function of p, instead of fixing them. Surprisingly, the only difference in our results as compared to Gallager's results is that in the expression of $\tilde{E}_o(\rho)$ (9), we have a $\sup$ over all possible non-zero input symbols rather than a $\max$ over the non-zero alphabet as in [1].

III. Independent MIMO fading channels

In this section, we will assume that the entries of $\mathbf{H}$ are independent and apply Theorem 1 to find the low-SNR reliability function for the MIMO fading channel (1). Further, we will characterize the asymptotic optimal signaling schemes which minimizes the average probability of decoding error for a certain communication rate.

A. Low-SNR reliability function

Since we assume that $\mathbf{H}$ has i.i.d. entries, given an input matrix $\mathbf{X}$, the row vectors of the output matrix $\mathbf{Y}$ are independent of each other. Denote $\mathbf{Y}^T = (\mathbf{y}_1^*, \mathbf{y}_2^*, \cdots, \mathbf{y}_N^*)$, $\mathbf{H}^T = (\mathbf{h}_2, \mathbf{h}_2, \cdots, \mathbf{h}_N)$ and $\mathbf{W}^T = (\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_N)$. The covariance matrix of $\mathbf{y}_i^\dagger = \mathbf{h}_i^T\mathbf{X} + \mathbf{w}_i^T$ is

$$\mathbf{\Sigma}_i = E[\mathbf{y}_i\mathbf{y}_i^\dagger] = \mathbf{I} + \mathbf{X}^\dagger\mathbf{X}.$$

Then, the channel transition probability density function $f(\mathbf{Y}|\mathbf{X})$ is as follows

$$f(\mathbf{Y}|\mathbf{X}) = \prod_{i=1}^N \frac{1}{\pi^{T_c}\det(\mathbf{I} + \mathbf{X}^\dagger\mathbf{X})} \exp\{-\mathbf{y}_i^\dagger(\mathbf{I} + \mathbf{X}^\dagger\mathbf{X})^{-1}\mathbf{y}_i\}$$

Applying Theorem 1 here, we have the following result.

Theorem 2 *For the MIMO model (1) with M transmit antennas and N receive antennas and independent fading matrix $\mathbf{H}$, the low-SNR reliability function is given by*

$$E^{lp}(r) = \sup_{0\leq\rho\leq 1} -\rho r + \tilde{E}_o(\rho), \qquad (12)$$

where $\tilde{E}_o(\rho)$ is defined as follows

$$\tilde{E}_o(\rho) = N(1+\rho)\sup_{x>0}\frac{1}{x}\ln\frac{1+\frac{\rho}{1+\rho}x}{(1+x)^{\frac{\rho}{1+\rho}}}. \qquad (13)$$

Remarks: Consider the SISO fast fading model where $M = N = T_c = 1$. The reliability function curve corresponding to this model is

$$E^{lp}(r) = \sup_{0\leq\rho\leq 1} -\rho r + (1+\rho)\sup_{x>0}\frac{1}{x}\ln\frac{1+\frac{\rho}{1+\rho}x}{(1+x)^{\frac{\rho}{1+\rho}}}. \qquad (14)$$

Let $\bar{\tau}(r)$ be the optimizing x in (14), i.e.,

$$\bar{\tau}(r) = \arg\max_{x>0}\sup_{0\leq\rho\leq 1} -\rho r + (1+\rho)\frac{1}{x}\ln\frac{(1+\frac{\rho}{1+\rho}x)}{(1+x)^{\frac{\rho}{1+\rho}}}. \qquad (15)$$

$\bar{\tau}(r)$ establishes a direct relation between rate and how large the non-zero symbol has to be for an on-off signaling scheme to achieve the optimal reliability function at rate r. As we will show in the next section, for any signaling scheme, including on-off signaling, it is a necessary condition that a symbol with energy of $\bar{\tau}(r)$ be used for a signaling scheme to be able to achieve the optimal channel performance at rate r. Thus, $\bar{\tau}(r)$ indicates the *minimum signal peakiness* for the SISO channel at rate r.

For the more general MIMO block-fading model, our result here tells us

1. Having a longer channel coherence time T_c does not improve the decoding error probability. This is a bit surprising since one would expect that making T_c large may give us more time to estimate the channel and thus help improve the non-coherent channel performance. However, our result shows otherwise.

2. Having more than one input antenna does not improve the channel performance.

3. The performance gain, as defined in Definition 1, for a MIMO fading channel is N, which is the number of receive antennas.

B. Optimal signaling schemes in the low-SNR regime

In this section we study the conditions for an input distribution to be optimal in the low-SNR regime.

Definition 4 *A sequence of input distributions $\{q_p(\mathbf{X}) \in \mathcal{D}(p)\}$ is called first-order optimal with respect to a communication rate r if it satisfies*

$$\lim_{p\to 0}\frac{E(rp, q_p, p)}{p} = E^{lp}(r), \qquad (16)$$

where $E(R, q_p, p)$ is the reliability function of the channel when the input distribution is chosen to be q_p. ◇

In the above definition, to be consistent with prior literature, we use the term *first-order* optimal to indicate optimality in the limit $SNR \to 0$ [10, 11, 12].

From the proof of Theorem 2, it immediately follows that the on-off signaling scheme given below is first-order optimal:

Corollary 1 *A sufficient condition for an on-off signaling scheme*

$$\mathbf{X} = \begin{cases} 0 & w.p.\ 1 - \frac{pT_c}{\|\mathbf{X}_o\|^2}; \\ \mathbf{X}_o & w.p.\ \frac{pT_c}{\|\mathbf{X}_o\|^2}, \end{cases} \tag{17}$$

to be first-order optimal with respect to a transmission rate r *is that the non-zero eigenvalues of* $\mathbf{X}_o\mathbf{X}_o^\dagger$ *are all equal to* $\kappa(r)$, *where* $\kappa(r) = \bar{\tau}(\frac{r}{N})$. ◇

For the SISO fast-fading model where $M = N = T_c = 1$, a first-order optimal signaling scheme is simply to choose the input symbol to be 0 or $\kappa(r)^{\frac{1}{2}}$. However, the peak-to-average ratio of the on-off signaling scheme goes to infinity as the mean power goes to 0. A natural question to ask here is the following: can we find less "peaky" signaling schemes which are also first-order optimal, by considering other signaling schemes other than on-off signaling? It turns out that only signaling schemes in the following category *can* be first-order optimal.

Definition 5 *A sequence of input distributions* $\{q_p\}$ *defined on the input matrix* $\mathbf{X} \in C^{M\times T_c}$ *is said to be* a general flash signaling scheme with peak constraint K *if it satisfies the average power constraint*

$$E_{\mathbf{X}}[\|\mathbf{X}\|^2] = pT_c,$$

and for any $\epsilon > 0$ *and any* $l = 1, 2, \cdots, T_c$,

$$\liminf_{p\to 0} \frac{E_{\mathbf{X}}\left[\lambda_l I_{\{|\lambda_l - K| \geq \epsilon\}}\right]}{p} = 0, \tag{18}$$

where $\{\lambda_1, \lambda_2, \cdots, \lambda_l\}$ *are the eigenvalues of the input matrix* $\mathbf{X}^\dagger\mathbf{X}$ *and* I_A *denotes the indicator function of the event* A. ◇

Remarks: This condition (18) can be interpreted as follows: *almost all probability mass is assigned to those input matrices* $\mathbf{X}$ *such that the eigenvalues of* $\mathbf{X}^\dagger\mathbf{X}$ *are either* 0, *or in an arbitrarily small neighborhood of a constant* K. Now consider the case when $T_c = 1$, the input matrix $\mathbf{X}$ is simply a vector and therefore the only eigenvalue of $\mathbf{X}^\dagger\mathbf{X}$ is $\|\mathbf{X}\|^2$. When $K = \kappa(r)$, (18) becomes

$$\liminf_{p\to 0} \frac{E_{\mathbf{X}}\left[\|\mathbf{X}\|^2 I_{\{|\|\mathbf{X}\|^2 - \kappa(r)| \geq \epsilon\}}\right]}{p} = 0. \tag{19}$$

As r approaches capacity, $\kappa(r)$ goes to infinity. Then this definition is equivalent to the definition of the *flash signaling* in [10]. Thus, the signaling scheme defined in (18) is essentially a generalized flash signaling scheme. The next lemma shows that a general flash signaling scheme with peak constraint $\kappa(r)$ is necessary to achieve the best reliability function at rate r.

Lemma 1 *A necessary condition for a sequence of input distributions* $\{q_p \in \mathcal{D}(p)\}$ *to be first-order optimal with respect to a coding rate* r *is that* $\{q_p\}$ *is a general flash-signaling scheme with peak constraint* $\kappa(r)$. ◇

The necessary condition (18) gives us a constraint on the eigenvalues of $\mathbf{X}^\dagger\mathbf{X}$ for a signaling scheme to be first-order optimal. However, usually it is not that convenient to look at the eigenvalues and try to make these eigenvalues satisfy certain peakiness conditions. For this purpose, we define the following quantity, which indicates the level of peakiness we need at the entries of the input matrix $\mathbf{X}$, rather than the eigenvalues of $\mathbf{X}^\dagger\mathbf{X}$, to achieve the optimal decoding error performance for a certain rate r. Let $\mathcal{F}_1(r)$ be the set of sequences of first-order optimal distributions with respect to rate r.

Definition 6 *We define* $\tau(r)$ *to be the* minimum signal peakiness *for a communication rate* r *on the MIMO block-fading channel (1):*

$$\tau(r) = \inf_{\{q_p\}\in\mathcal{F}_1(r)} \limsup_{p\to 0} \sup_{\mathbf{X}\in\mathcal{A}_p} \|\mathbf{X}\|_\infty^2, \tag{20}$$

where $\mathcal{A}_p$ *is the alphabet for the input distribution* q_p *and* $\|\mathbf{X}\|_\infty$ *is the* l_∞ *norm of matrix* $\mathbf{X}$. ◇

From a practical point of view, it is unreasonable to allow input signals with arbitrarily large peakiness. Thus, $\tau(r)$ is an important parameter which shows how difficult in practice it is to achieve the best performance at rate r. There is a rather simple relation between $\tau(r)$ and $\kappa(r)$, which is presented in the following theorem.

Theorem 3 *For a MIMO block-fading channel (1), the minimum signal peakiness to achieve the optimum decoding error probability at coding rate* r *is*

$$\tau(r) = \frac{\kappa(r)}{MT_c} = \frac{\bar{\tau}(\frac{r}{N})}{MT_c}. \tag{21}$$

Thus, the peakiness gain achieved by the MIMO block-fading channel is MT_c.

This theorem tells us that the role of having more than one transmit antenna in the low-SNR regime is to reduce the minimum signal peakiness we need to achieve the optimal reliability function. Further, it can be shown that for a fading channel with both average and peak power constraint, having multiple transmit antennas can actually *improve* the channel performance, in terms of both capacity and reliability function.

IV. CORRELATED MIMO FADING CHANNELS

In the last section, we studied the MIMO channel model (1) with the assumption that each entry of the fading matrix $\mathbf{H}$ is an i.i.d. complex Gaussian random variable. In this section, we assume $\mathbf{H}$ to be a *proper* complex Gaussian random matrix with an arbitrary correlation matrix. Specifically, we assume that each entry of $\mathbf{H}$ is a symmetric complex Gaussian random variable $\mathcal{CN}(0, 1)$ and we allow an arbitrary correlation between any two entries of $\mathbf{H}$. This correlation is usually referred to as the *spatial correlation*.

For general Gaussian random matrix $\mathbf{H}$, we cannot always get a closed-form expression for the reliability function as in the capacity case [10]. However, we can get an upper bound on the reliability function.

Theorem 4 *For the non-coherent MIMO model (1) with a general fading matrix* $\mathbf{H}$, *we have*

$$E^{lp}(r) = \sup_{0\le\rho\le 1} -\rho r + \tilde{E}_o(\rho), \tag{22}$$

where $\tilde{E}_o(\rho)$ *is upper bounded by*

$$\tilde{E}_o(\rho) \le G(1+\rho)\sup_{x>0}\frac{1}{x}\ln\frac{1+\frac{\rho}{1+\rho}x}{(1+x)^{\frac{\rho}{1+\rho}}}, \tag{23}$$

and G *is the maximum channel gain [10] and is defined as*

$$G = \sup_{\mathbf{X}\neq\mathbf{0}}\frac{E_{\mathbf{H}}[\|\mathbf{H}\mathbf{X}\|^2]}{\|\mathbf{X}\|^2} = \lambda_{max}(E_{\mathbf{H}}[\mathbf{H}^\dagger\mathbf{H}]).$$

Here $\lambda_{max}(\mathbf{A})$ *denotes the largest eigenvalue of matrix* $\mathbf{A}$. ◇

The inequality in (23) can not always be achieved as equality. For example, when $M = 1$ and $N > 1$, it can be shown that we can not achieve equality in (23). On the other hand, for the opposite extreme case when $N = 1$, we can check that the upper bound in (23) can be achieved.

Now we study the quantity $\lambda_{max}(E_{\mathbf{H}}[\mathbf{H}^\dagger\mathbf{H}])$. We claim

$$\lambda_{max}(E_{\mathbf{H}}[\mathbf{H}^\dagger\mathbf{H}]) \ge N, \tag{24}$$

which is the channel gain under independently faded matrix $\mathbf{H}$. To see this, consider

$$\sum_{l=1}^{M}\lambda_l = Tr(E_{\mathbf{H}}[\mathbf{H}^\dagger\mathbf{H}]) = MN \le M\lambda_{max}.$$

Thus for any fading matrix $\mathbf{H}$, (24) must be true, which means the reliability function for a correlated $\mathbf{H}$ may be larger than (12). Especially for the case when $N = 1$, we can show that any fading matrix $\mathbf{H}$ will do better than the independent fading case. The independent fading channel model is widely studied because it can provide better *diversity gain* and *degree of freedom gain* in the high SNR regime, as compared with the correlated fading matrix. To get approximately independent fading paths, we have to put the antennas physically far apart (more than half the transmission wavelength), which limits the possibility of having multiple antennas on small communication devices. However, in the low-SNR regime, the *diversity gain* and *degree of freedom gain* are not more important that the *power gain*. Thus, in this regime, the design goal of utilizing the MIMO system has changed and it is might not be a good idea to *make* the fading paths independent at all.

Next we assume that we have control over the fading matrix $\mathbf{H}$ and explore the best fading matrix $\mathbf{H}$ in the sense that we can obtain the best reliability function under this fading matrix. Specifically, we need to maximize the reliability function (22) over $\mathbf{H}$ under the constraint that $\mathbf{H}$ is complex Gaussian and each entry has variance 1.

Theorem 5 *A fully correlated fading channel matrix* $\mathbf{H}$ ($h_{ij} = h;\ \forall i,j$) *provides the maximum performance gain and the reliability function for this fading model is*

$$E^{lp}(r) = \sup_{0\le\rho\le 1} -\rho r + \tilde{E}_o(\rho), \tag{25}$$

where $\tilde{E}_o(\rho)$ *is upper bounded by*

$$\tilde{E}_o(\rho) = MN(1+\rho)\sup_{x>0}\frac{1}{x}\ln\frac{1+\frac{\rho}{1+\rho}x}{(1+x)^{\frac{\rho}{1+\rho}}}. \tag{26}$$

Next, we find the minimum signal peakiness $\tau(r)$ for this fully correlated MIMO channel. With the input matrix we constructed in the proof of Theorem 5, it is straightforward to have the following theorem:

Corollary 2 *For the fully correlated MIMO channel, the minimum signaling peakiness is*

$$\tau(r) = \frac{\bar{\tau}(\frac{r}{MN})}{M^2NT_c}. \tag{27}$$

In other words, when we have a fully correlated MIMO fading channel, we can achieve a performance gain of MN *and a peakiness gain of* M^2NT_c, *as compared to a SISO fast-fading channel.* ◇

It is not necessary to require that $\mathbf{H}$ has identical entries in order to achieve the optimal performance gain in a MIMO channel. Next we consider the case that any two entries of $\mathbf{H}$ differ by a phase shift:

$$E[h_{nm}h^*_{n'm'}] = e^{j\theta_{nmn'm'}}, \tag{28}$$

for any $n, n' = 1, 2, \cdots, N$ and $m, m' = 1, 2, \cdots, M$. Further, we assume the correlation between any pair of transmitter and receive antenna is the product of the *transmit correlation* and *receiver correlation*, i.e.,

$$E[h_{nm}h^*_{n'm'}] = R^r_{mm'}R^t_{nn'},$$

where $\mathbf{R}^r = \{R^r_{mm'}\}$ and $\mathbf{R}^t = \{R^r_{nn'}\}$ are the transmit correlation matrix and the receive correlation matrix, respectively. In our case, where the only difference between any two entries of $\mathbf{H}$ is a phase shift as in (28), it is easy to see that this is only possible when all entries of $\mathbf{R}^r$ or $\mathbf{R}^t$ have magnitude 1. Denote $\mathbf{u} = (1, e^{ju_2}, \cdots, e^{ju_M})^T$ and $\mathbf{v} = (1, e^{jv_2}, \cdots, e^{jv_N})^T$. Assume $\mathbf{R}^t = \mathbf{u}\mathbf{u}^\dagger$ and $\mathbf{R}^r = \mathbf{v}\mathbf{v}^\dagger$. Physically, this form of correlation matrix means that the phase shift between the m-th and m'-th transmit antenna is $u_m - u_{m'}$ and the phase shift between the n-th and n'-th receive antenna is $v_n - v_{n'}$ ($u_1 = v_1 = 0$). Thus,

$$E[h_{nm}h^*_{n'm'}] = e^{j(u_m - u_{m'} + v_n - v_{n'})}. \tag{29}$$

Let $\mathbf{H}_h$ denote the fading matrix with identical entries, as the one we assumed in Theorem 5. The correlated fading matrix with correlation (29) can then be written as

$$\mathbf{H} = \mathbf{V}\mathbf{H}_h\mathbf{U}, \tag{30}$$

where $\mathbf{U} = diag(\mathbf{u})$ and $\mathbf{V} = diag(\mathbf{v})$.

Corollary 3 *For a MIMO fading channel with fading matrix* $\mathbf{H}$ *satisfying (30), we can also achieve a performance gain of* MN *and peakiness gain* M^2NT_c.

Using an angular domain representation [8] of the MIMO multipath fading channels, we can establish physical conditions on the transmit and receive antenna array sizes for $\mathbf{H}$ to have entries which are identical or differ by a phase shift. We refer the reader to discussions in the full version of this paper to check these conditions.

Theorem 5, Corollary 2 and Corollary 3 tell us the potential performance gain and peakiness gain we can have with MIMO channels in the low-SNR regime, if the entries of the fading matrix $\mathbf{H}$ are fully correlated. The first observation here is that it is important to make the fading correlated in the low-SNR regime. We can potentially have a reliability function which is M times better than the result with independent fading matrix. For this model, both the low-SNR capacity and reliability function increases in proportion to the *product* of the number of the transmit antennas and the receive antennas. Further, the minimum signal peakiness reduces by a factor of $\frac{1}{MN}$ as compared to the independent fading case, which makes it much easier to achieve the reliability function. Thus, in the low-SNR regime with multiple antennas, we should try to make the channel fading more correlated, (for example, by putting the antennas close to each other physically) to fully utilize the advantage of having multiple antennas.

Another observation here is that multiple transmit antennas can be beneficial. For the independent fading case, we have shown that multiple receive antennas give us a performance gain while multiple transmit antennas only provide a peakiness gain. However, with fully correlated fading, we proved that multiple transmit antennas contribute a peakiness gain of M^2 *and* a performance gain of M, while multiple receive antennas give us a performance gain of N and a peakiness gain of N. Thus, transmit antennas are very important for low-SNR communications.

V. CONCLUSIONS

In this paper, we investigated the tradeoff between communication rate and average probability of decoding error for a noncoherent multiple-antenna fading channel in a low-SNR regime, using the framework of error-exponent theory. We started with the assumption that the fading matrix $\mathbf{H}$ has i.i.d. entries. In this regime, we showed that using M transmit antennas and N receive antennas allow us to realized a performance gain of N and peakiness gain M. Further, if the channel is constant every T_c symbols, we can see a further peakiness gain of T_c. However, neither increasing M or T_c can improve the asymptotic communication rate or the reliability function of the channel. Further, when both the average and peak power are constrained, having larger M or T_c can improve both the channel capacity and the low-SNR reliability function.

In the low-SNR regime, channel correlation can actually improve the channel performance. In the extreme case where the fading is fully correlated, in the sense that the entries of the fading matrix $\mathbf{H}$ are either identical or differ by a phase shift, we can achieve a performance gain of MN and a peakiness gain of M^2NT_c. Thus, the advantage of having multiple antennas is best realized when we have fully correlated fading channels.

REFERENCES

[1] R. G. Gallager. Energy limited channels: Coding, multiaccess, and spread spectrum, November 1987. Tech. Report LIDS-P-1714, LIDS, MIT, Cambridge, Mass. A later version appeared as a book chapter in *Codes, Graphs, and Systems: A Celebration of the Life and Career of G. David Forney, Jr. on the Occasion of His Sixtieth Birthday (Kluwer International Series in Engineering and Computer Science, 670)*.

[2] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press 1985.

[3] M. Medard and R. G. Gallager. Bandwidth scaling for fading multipath channels. *IEEE Transactions on Information Theory*, 48(4):840–853, April 2002.

[4] V. G. Subramanian. and B. Hajek. Broad-band fading channels: Signal burstiness and capacity. *IEEE Transactions on Information Theory*, 48(6):809–827, Apr. 2002.

[5] V. Tarokh, H. Jafarkhani, and A. R. Calderbank. Space-time block code from orthogonal designs. *IEEE Trans. Info. Theory*, vol. 45, pp. 1456–67, July 1999.

[6] I. E. Telatar. Capacity of multi-antenna Gaussian channels. *European Trans. On Telecommunications*, vol. 10, pp. 585-95, Nov/Dec 1999.

[7] E. Telatar and D. Tse. Capacity and mutual information of wideband multipath fading channels. *IEEE Transaction on Information Theory*, 46(7):1384-1400, 2000.

[8] D. N. Tse and P. Viswanath. *Fundamentals of Wireless Communications*. Preprint.

[9] S. Verdu. On channel capacity per unit cost. *IEEE Transactions on Information Theory*, 36(9):1019–1030, Sept. 1990.

[10] S. Verdu. Spectral efficiency in the wideband regime. *IEEE Transactions on Information Theory*, 48(6):1319–1343, June 2002.

[11] X. Wu and R. Srikant. Near-optimal Input Distributions In the Wideband Regime Part I: Parallel Gaussian Channel. (A short version of this paper appeared in the Proceedings of CISS, March 2003. The longer version is available at http://www.comm.csl.uiuc.edu/˜xwu/) Submitted to *IEEE Transactions on Information Theory.*

[12] X. Wu and R. Srikant. Near-optimal Input Distributions In the Wideband Regime Part II: Coherent Fading Channel. Submitted to *IEEE Transactions on Information Theory.*

[13] L. Zheng and D. N. Tse. Diversity and Multiplexing: A Fundamental Tradeoff in Multiple Antenna Channels. *IEEE Transactions on Information Theory*, 49(5):1073–96 May, 2003.

[14] L. Zheng and D. N. Tse. The Diversity-Multiplexing Tradeoff for Noncoherent Multiple Antenna Channels. *Proc. Allerton Conference*, 2002.