# Correspondence

## Correlated Jamming on MIMO Gaussian Fading Channels

Akshay Kashyap, *Student Member, IEEE*, Tamer Başar, *Fellow, IEEE*, and R. Srikant, *Senior Member, IEEE*

***Abstract*—We consider a zero-sum mutual information game on multiple-input multiple-output (MIMO) Gaussian Rayleigh-fading channels. The players are an encoder–decoder pair as the maximizer, and a jammer as the minimizer, of the mutual information between the input and the output of the channel. There are total power constraints on both the jammer and the encoder. Also, the jammer has access to the encoder output. We find the unique saddle point of this game, and prove the somewhat surprising result that the knowledge of the channel input is useless to the jammer.**

***Index Terms*—Correlated jamming, multiple-input multiple-output (MIMO) Gaussian fading channels, zero-sum games in information theory.**

### I. Introduction

In this correspondence, we consider the problem of communication on a multiple-input multiple-output (MIMO) fading channel in the presence of a disruptive jammer. We use the mutual information between the input and the output of the channel as a measure of the effectiveness of the communication. We therefore study a zero-sum game in which the players are the encoder–decoder pair (which we refer to as the communicator from now on) as the maximizer and a jammer as the minimizer, with mutual information as the payoff (to the maximizer).

Mutual information games have been studied in various settings, for example [4] and [10]. However, we also allow the jammer perfect access to the output of the encoder. Thus, our problem is an adaptation of the correlated jamming problem studied in [9] to MIMO fading channels. Correlated jamming has also been considered with a different payoff on an additive white Gaussian noise (AWGN) channel in [2].

We find the unique saddle point of the above game. Our approach is similar to that of [9], but our main result is that for a Rayleigh-fading channel, the information about the channel input is useless to the jammer. This is different from the strategies found in [9] and [2] for constant channels, where the jammer uses the knowledge of the channel input to determine its signal.

The outline of the correspondence is as follows. In Section II, we state the problem formally and introduce notation. In Section III, we prove the main result of the correspondence, that is, the side information about the channel input is useless to the jammer, and find the saddle point of the game. In Section IV, we give a fixed analog of the Rayleigh-fading channel considered in Section III, and show that for a constant channel, side information is always useful to the jammer. We make some concluding remarks in Section V.

Manuscript received November 12, 2003; revised April 16, 2004. The work of A. Kashyap was supported by the Vodafone-US Foundation Fellowship for the year 2003-2004. The work of all authors was supported in part by the National Science Foundation under Grant CCR-0085917 ITR. The material in this correspondence was presented in part at the IEEE 2004 International Conference on Communications, Paris, France, June 2004 [7].

The authors are with the Department of Electrical Engineering and the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA.

Communicated by M. Medard, Associate Editor for Communications.
Digital Object Identifier 10.1109/TIT.2004.833358

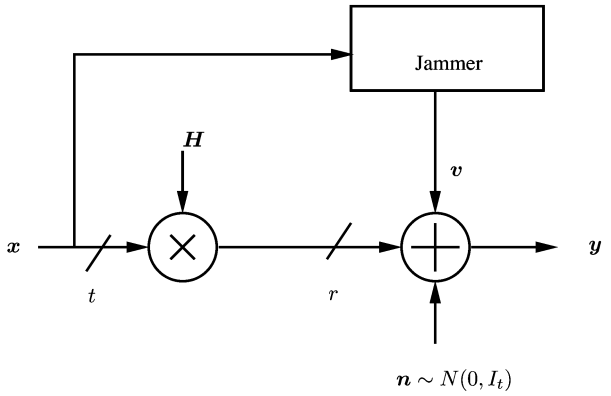


Fig. 1. Mutual information game with intelligent jammer.

### II. Problem Statement

In this correspondence, we use the linear channel model

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{n} + \boldsymbol{v} \tag{1}$$

depicted in Fig. 1. Here, $\boldsymbol{x} \in \mathbb{C}^t$ is the "legitimate" input of the channel, that is, the output of the encoder; $\boldsymbol{y} \in \mathbb{C}^r$ is the output of the channel; $\boldsymbol{n} \in \mathbb{C}^r$ is an additive complex circularly symmetric Gaussian (CSCG henceforth) noise, $\boldsymbol{n} \sim \mathcal{CN}(0, I_r)$; and $\boldsymbol{H} \in \mathbb{C}^{r \times t}$ is the (random) channel gain. Each entry of $\boldsymbol{H}$ is assumed to be distributed $\mathcal{CN}(0,1)$, and realized independently for each channel use. Such a channel model corresponds to a Rayleigh-faded wireless link. Finally, $\boldsymbol{v} \in \mathbb{C}^r$ is the jammer input.

We assume that the jammer knows the realization $x$ of the channel input $\boldsymbol{x}$ and is able to process it to generate its signal, and so $\boldsymbol{v}$ is allowed to be arbitrarily correlated with $\boldsymbol{x}$. We also impose power constraints both on the encoder and the jammer

$$\mathcal{E}[\boldsymbol{x}^\dagger \boldsymbol{x}] \le P \tag{2}$$

$$\mathcal{E}[\boldsymbol{v}^\dagger \boldsymbol{v}] \le E_J. \tag{3}$$

We assume that the decoder has perfect information about the channel state, that is, the receiver knows the realization $H$ of the channel state $\boldsymbol{H}$. The encoder and the jammer are assumed to know only the distribution of $\boldsymbol{H}$, i.e., the random variables $\boldsymbol{x}$ and $\boldsymbol{v}$ are independent of $\boldsymbol{H}$.

We also make the natural assumption that $\boldsymbol{n}$ is independent of all the other random variables.

We study a mutual information zero-sum game on this channel between the communicator and the jammer. The payoff $J$ is defined as

$$J(\boldsymbol{x}, \boldsymbol{v}) = \mathcal{I}(\boldsymbol{x}; \boldsymbol{y}) \tag{4}$$

where $\mathcal{I}(\boldsymbol{x}; \boldsymbol{y})$ denotes the mutual information between $\boldsymbol{x}$ and $\boldsymbol{y}$. We are interested in finding the saddle point of $J$. (We are abusing the notation slightly, since the strategies in this game actually are the distributions of $\boldsymbol{x}$ and $\boldsymbol{v}$, but the notation $J(\boldsymbol{x}, \boldsymbol{v})$ does not lead to any confusion and is simpler.)

### A. Notation

1) For $\boldsymbol{x}$ (real) Gaussian with mean $\mu$ and covariance matrix $Q$, we write $\boldsymbol{x} \sim \mathcal{N}(\mu, Q)$. For $\boldsymbol{x}$ CSCG with mean $\mu$ and variance $Q$, we write $\boldsymbol{x} \sim \mathcal{CN}(\mu, Q)$.

2) We use $\boldsymbol{\Sigma_x}$ to denote the covariance $\mathcal{E}[\boldsymbol{x}\boldsymbol{x}^\dagger]$ of $\boldsymbol{x}$, and $\boldsymbol{\Sigma_{xy}}$ to denote the cross-covariance $\mathcal{E}[\boldsymbol{x}\boldsymbol{y}^\dagger]$ between $\boldsymbol{x}$ and $\boldsymbol{y}$.

3) We use $h(\boldsymbol{x})$ to denote the (differential) entropy of $\boldsymbol{x}$ and $\mathcal{I}(\boldsymbol{x};\boldsymbol{y})$ to denote the mutual information between $\boldsymbol{x}$ and $\boldsymbol{y}$.

4) For $Q \in \mathbb{C}^{n\times n}$,we write $Q \succ 0$ if $Q$ is Hermitian and positive definite. We write $Q \succeq 0$ if $Q$ is Hermitian and positive semi-definite.

## III. Saddle Point for a MIMO Rayleigh-Fading Gaussian Channel

To find the saddle point of $J(\boldsymbol{x}, \boldsymbol{v})$ as defined in (4), we proceed in the following manner.

1) Reduce the game to one in which the strategies of the players are covariance matrices by proving that the saddle-point strategy of the communicator is a CSCG signal $\boldsymbol{x} \sim \mathcal{CN}(0, Q)$, and that of the jammer is of the form $\boldsymbol{v} = \xi\boldsymbol{x} + \boldsymbol{z}$ where $\xi \in \mathbb{C}^{r\times t}$ is a constant matrix and $\boldsymbol{z}$ is distributed $\mathcal{CN}(0, \boldsymbol{\Sigma_z})$ independent of $\boldsymbol{x}$.

2) Assume that the communicator uses $Q = (P/t)I_t$. Prove that it is optimal for the jammer to use $\xi = 0$ and $\boldsymbol{\Sigma_z} = (E_J/r)I_r$.

3) Assume that the jammer uses $\xi = 0$ and $\boldsymbol{z} \sim \mathcal{CN}(0, (E_J/r)I_r)$. Prove that it is optimal for the communicator to use $Q = (P/t)I_t$.

### A. Reduction to Matrix Strategies

Since we are assuming perfect channel state information at the receiver, once we condition on this information, this step of the solution is identical to one in [9]. We list the reasoning for the sake of completeness.

Assume a CSCG distribution on $\boldsymbol{x}$, $\boldsymbol{x} \sim \mathcal{CN}(0, Q)$. Now

$$\begin{aligned} J &= \mathcal{I}(\boldsymbol{x}; (\boldsymbol{y}, \boldsymbol{H})) \\ &= \mathcal{I}(\boldsymbol{x}; \boldsymbol{y}|\boldsymbol{H}) \quad \text{since } \boldsymbol{x} \text{ is independent of } \boldsymbol{H} \\ &= h(\boldsymbol{x}|\boldsymbol{H}) - h(\boldsymbol{x}|\boldsymbol{y}, \boldsymbol{H}). \end{aligned}$$

The jammer can minimize $\mathcal{I}$ only through maximizing the conditional entropy. Now, for any function $A(H) : \mathbb{C}^{r\times t} \mapsto \mathbb{C}^{t\times r}$

$$\begin{aligned} h(\boldsymbol{x}|\boldsymbol{y}, \boldsymbol{H}) &= h(\boldsymbol{x} - A(\boldsymbol{H})\boldsymbol{y}|\boldsymbol{y}, \boldsymbol{H}) \\ &\le h(\boldsymbol{x} - A(\boldsymbol{H})\boldsymbol{y}|\boldsymbol{H}) \\ &\le \mathcal{E}\left[\log\det(\boldsymbol{\Sigma_{e|H}})|\boldsymbol{H}\right] \end{aligned} \tag{5}$$

where in the last inequality we have taken $A(H) = \boldsymbol{\Sigma_{xy}}\boldsymbol{\Sigma_y}^{-1}$, and

$$\boldsymbol{\Sigma_{e|H}} = \boldsymbol{\Sigma_x} - \boldsymbol{\Sigma_{xy}}\boldsymbol{\Sigma_y}^{-1}\boldsymbol{\Sigma_{yx}},\ \boldsymbol{\Sigma_y} = \mathcal{E}[\boldsymbol{y}\boldsymbol{y}^\dagger|\boldsymbol{H}],\ \text{and } \boldsymbol{\Sigma_{xy}} = \mathcal{E}[\boldsymbol{x}\boldsymbol{y}^\dagger|\boldsymbol{H}].$$

Equality holds in (5) if and only if $\boldsymbol{y}$ is jointly CSCG with $\boldsymbol{x}$ conditioned on $\boldsymbol{H}$. However, since $\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{n} + \boldsymbol{v}$, for it to be jointly CSCG with $\boldsymbol{x}$ conditioned on $\boldsymbol{H}$ we need $\boldsymbol{v}$ to be jointly CSCG with $\boldsymbol{x}$ conditioned on $\boldsymbol{H}$. Further, $\boldsymbol{v}$ and $\boldsymbol{x}$ being independent of $\boldsymbol{H}$, we need $\boldsymbol{v}$ to be jointly CSCG with $\boldsymbol{x}$, that is, even without the conditioning on $\boldsymbol{H}$. Moreover, any value of $\mathcal{E}\left[\log\det(\boldsymbol{\Sigma_{e|H}})|\boldsymbol{H}\right]$ allowable by the constraints can also be achieved by a Gaussian distribution on $\boldsymbol{v}$.

So, without any loss of generality, we have that the jammer strategy for any Gaussian input $\boldsymbol{x}$ is of the form

$$\boldsymbol{v} = \xi\boldsymbol{x} + \boldsymbol{z}. \tag{6}$$

Now assume a jammer strategy as in (6). Then the communicator sees the equivalent channel $\boldsymbol{y} = (\boldsymbol{H} + \xi)\boldsymbol{x} + (\boldsymbol{n} + \boldsymbol{z})$, which is just a Rician channel with side information at the receiver. So we know that the input that maximizes $\mathcal{I}(\boldsymbol{x};\boldsymbol{y})$ is CSCG with a covariance $Q$ that depends only on $\xi$ and the covariance $(I + \boldsymbol{\Sigma_z})$ of additive noise in the equivalent channel.

So, we have reduced the problem to a game where the strategy of the communicator is completely described by the covariance matrix $Q$ and that of the jammer is completely described by the pair $(\xi, \boldsymbol{\Sigma_z})$, and the payoff is

$$J(Q, \xi, \boldsymbol{\Sigma_z}) = \mathcal{E}\left[\log\det\left\{(\boldsymbol{H} + \xi)Q(\boldsymbol{H} + \xi)^\dagger(\boldsymbol{\Sigma_z} + I)^{-1} + I\right\}\right]. \tag{7}$$

The power constraints (2) and (3) on the encoder and the jammer signals translate to the following constraints on the matrix strategies:

$$\begin{aligned} \mathrm{tr}(Q) &\le P \\ \mathrm{tr}\left(\xi Q\xi^\dagger + \boldsymbol{\Sigma_z}\right) &\le E_J. \end{aligned}$$

### B. Finding the Saddle Point in Matrix Strategies

In the derivation of the saddle-point strategies, we make use of the following lemmas.

*Lemma 1:* For any $\gamma \ge 0$ and any $M \in \mathbb{C}^{r\times r}$, and nonnegative principal diagonal matrices $\Lambda^{(1)}, \Lambda^{(2)} \in \mathbb{C}^{r\times t}$ such that

$$\Lambda^{(1)}(\Lambda^{(1)})^\dagger \succeq \Lambda^{(2)}(\Lambda^{(2)})^\dagger$$

the following stochastic ordering [1] holds:

$$\log\det\left\{\gamma I_r + M(\boldsymbol{H} + \Lambda^{(1)})(\boldsymbol{H} + \Lambda^{(1)})^\dagger M^\dagger\right\} \overset{\text{st}}{\ge} \log\det\left\{\gamma I_r + M(\boldsymbol{H} + \Lambda^{(2)})(\boldsymbol{H} + \Lambda^{(2)})^\dagger M^\dagger\right\}.$$

*Proof:* See Appendix I. □

*Lemma 2:* [11, Lemma 5] For $\boldsymbol{H} \in \mathbb{C}^{r\times t}$ with each entry $\boldsymbol{H}_{ij}$ independent and identically distributed (i.i.d.) $\mathcal{CN}(0,1)$, given any two unitary matrices $U \in \mathbb{C}^{r\times r}$ and $V \in \mathbb{C}^{t\times t}$, the distribution of $\boldsymbol{H}$ is the same as that of $U\boldsymbol{H}V^\dagger$.

*Lemma 3:* For $X \succeq 0$ the function

$$f(N) = \log\det\left(I + XN^{-1}\right)$$

is convex over the positive-definite cone $\mathcal{S} = \{N \in \mathbb{C}^{n\times n} | N \succ 0\}$, with strict convexity for $X \succ 0$.

*Proof:* See Appendix II .[2] □

With the three preceding lemmas at hand, we now proceed by assuming that $\boldsymbol{x} \sim \mathcal{CN}(0, (P/t)I_t)$. Putting $Q = (P/t)I_t$ in (7), we get

$$J\left(\frac{P}{t}I_t, \xi, \boldsymbol{\Sigma_z}\right) = \mathcal{E}\log\det\left\{I + \frac{P}{t}(\boldsymbol{\Sigma_z} + I)^{-\frac{1}{2}}(\boldsymbol{H} + \xi)(\boldsymbol{H} + \xi)^\dagger(\boldsymbol{\Sigma_z} + I)^{-\frac{1}{2}}\right\}. \tag{8}$$

Using singular value decomposition to write $\xi = U\Lambda_\xi V^\dagger$, where $U$ and $V$ are unitary matrices and $\Lambda_\xi$ is a nonnegative principal diagonal

[1] For two random variables $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}$ we say that $\boldsymbol{x} \overset{\text{st}}{\ge} \boldsymbol{y}$ if for every $t$, $\Pr(\boldsymbol{y} \le t) \ge \Pr(\boldsymbol{x} \le t)$.

[2] An information-theoretic proof of this lemma appeared earlier in [5, Lemma II.3].

matrix consisting of the singular values of $\xi$, and substituting in (8), we get

$$J\left(\frac{P}{t}I_t,\xi,\boldsymbol{\Sigma_z}\right)=\mathcal{E}\log\det\left\{I+\frac{P}{t}(\boldsymbol{\Sigma_z}+I)^{-\frac{1}{2}}U(\boldsymbol{H}+\Lambda_\xi)(\boldsymbol{H}+\Lambda_\xi)^\dagger U^\dagger(\boldsymbol{\Sigma_z}+I)^{-\frac{1}{2}}\right\}$$

where we have made use of Lemma 2.

So, using Lemma 1, we see that

$$J\left(\frac{P}{t}I_t,0,\boldsymbol{\Sigma_z}\right)\le J\left(\frac{P}{t}I_t,\xi,\boldsymbol{\Sigma_z}\right)$$

for any $\xi$, and so it is optimal for the jammer to use $\xi=0$.

It now remains to find the optimal $\boldsymbol{\Sigma_z}$. To that end, note that

$$J\left(\frac{P}{t}I_t,0,\boldsymbol{\Sigma_z}\right)=\mathcal{E}\left[\log\det\left(I+\frac{P}{t}\boldsymbol{H}\boldsymbol{H}^\dagger(\boldsymbol{\Sigma_z}+I)^{-1}\right)\right]$$

so that, on putting $\boldsymbol{\Sigma_z}=U\Lambda_zU^\dagger$ for a unitary matrix $U$ and a positive diagonal matrix $\Lambda_z$, we see (using Lemma 2 again) that

$$J\left(\frac{P}{t}I_t,0,U\Lambda_zU^\dagger\right)=J\left(\frac{P}{t}I_t,0,\Lambda_z\right)\tag{9}$$

so that it suffices to minimize $J$ over positive diagonal $\boldsymbol{\Sigma_z}=\Lambda_z$.

Now, for a permutation matrix $\Pi$, we see that, just as in (9)

$$J\left(\frac{P}{t}I_t,0,\Pi\Lambda_z\Pi^\dagger\right)=J\left(\frac{P}{t}I_t,0,\Lambda_z\right).$$

So, using Lemma 3 and Jensen's inequality

$$\begin{aligned}J\left(\frac{P}{t}I_t,0,\frac{1}{r!}\sum_\Pi\Pi\Lambda_z\Pi^\dagger\right)&\le\frac{1}{r!}\sum_\Pi J\left(\frac{P}{t}I_t,0,\Pi\Lambda_z\Pi^\dagger\right)\\&=J\left(\frac{P}{t}I_t,0,\Lambda_z\right)\end{aligned}$$

for any positive diagonal $\Lambda_z$.

But $\frac{1}{r!}\sum_\Pi\Pi\Lambda_z\Pi^\dagger=\theta I$, where $\theta$ is real and positive. From the jammer power constraint, it is obvious that the optimal value of $\theta$ is $(E_J/r)$.

Thus, we have proved that for the communicator strategy $\boldsymbol{x}\sim\mathcal{CN}(0,(P/t)I_t)$, the optimal jammer strategy is $\boldsymbol{v}=\boldsymbol{z}$ where $\boldsymbol{z}\sim\mathcal{CN}(0,(E_J/r)I_r)$.

Further, if the jammer strategy is $\boldsymbol{v}=\boldsymbol{z}$ with $\boldsymbol{z}\sim\mathcal{CN}(0,(E_J/r)I_r)$, then the problem of finding the optimal communicator strategy is the same as the mutual information optimization for finding the capacity of a Rayleigh-fading channel [11], and the optimal communicator response is $\boldsymbol{x}\sim\mathcal{CN}(0,(P/t)I_t)$.

Hence, we have the following result.

*Theorem 1:* The mutual information game (4) for the vector fading channel in (1) has the unique saddle point

$$\begin{aligned}\boldsymbol{x}^*&\sim\mathcal{CN}\left(0,\frac{P}{t}I_t\right)\\\boldsymbol{v}^*&\sim\mathcal{CN}\left(0,\frac{E_J}{r}I_r\right).\end{aligned}$$

The knowledge of $\boldsymbol{x}$ is useless for the jammer.

*Remark:* Uniqueness follows from the interchangeability property of saddle points [3]. If $(\tilde{\boldsymbol{x}},\tilde{\boldsymbol{v}})$ is another saddle point, then $(\tilde{\boldsymbol{x}},\boldsymbol{v}^*)$ and $(\boldsymbol{x}^*,\tilde{\boldsymbol{v}})$ should also be a saddle points. However, $\boldsymbol{x}^*$ is the only communicator strategy in equilibrium with $\boldsymbol{v}^*$ and *vice versa*. Thus, $\tilde{\boldsymbol{x}}=\boldsymbol{x}^*$ and $\tilde{\boldsymbol{v}}=\boldsymbol{v}^*$. Therefore, $x(\boldsymbol{x}^*,\boldsymbol{v}^*)$ is the unique saddle point.

## IV. Saddle Point When the Channel Gain Is a Constant

In this section, we give an example of a mutual information game similar to the one studied above, with the exception that the channel is now a constant gain channel. The purpose is to illustrate that for a constant channel, the jammer *always* makes use of the knowledge of the encoder output, as one would intuitively expect. So it is indeed the (symmetric) randomness of the channel that renders this additional information useless to the jammer.

The channel we now consider is a scalar, constant version of the channel in (1)

$$\boldsymbol{y}=H\boldsymbol{x}+\boldsymbol{n}+\boldsymbol{v}\tag{10}$$

where all quantities are now scalars, and $H\in\mathbb{C}$ is a constant. The power constraints are the same as in (2) and (3).

This is essentially the same problem considered in [9]. The solution therefore is very similar.

We start with the assumption that $\boldsymbol{x}\sim\mathcal{CN}(0,P)$. Then, just as in Section III, we can take the jammer strategy to be of the form

$$\boldsymbol{v}=\xi\boldsymbol{x}+\boldsymbol{z}\tag{11}$$

where $\xi\in\mathbb{C}$ is a constant, and $\boldsymbol{z}\sim\mathcal{CN}(0,\sigma_z^2)$ is independent of $\boldsymbol{x}$, with variance $\sigma_z^2$.

Now we fix the jammer strategy as in (11), for some $\xi$ and $\sigma_z^2$. For this jammer strategy, the communicator sees the channel

$$\boldsymbol{y}=(H+\xi)\boldsymbol{x}+(\boldsymbol{n}+\boldsymbol{z})$$

for which the mutual information maximizing input distribution is $\boldsymbol{x}\sim\mathcal{CN}(0,P)$.

So the saddle-point strategy of the communicator is $\boldsymbol{x}\sim\mathcal{CN}(0,P)$, and that of the jammer is of the form $\boldsymbol{v}=\xi\boldsymbol{x}+\boldsymbol{z}$. Since for this $\boldsymbol{v}$

$$\begin{aligned}\mathcal{I}(\boldsymbol{x};\boldsymbol{y})&=\mathcal{I}((H+\xi)\boldsymbol{x}+\boldsymbol{n}+\boldsymbol{v};\boldsymbol{x})\\&=\log\left[1+\frac{|H+\xi|^2P}{\sigma_n^2+\sigma_z^2}\right]\end{aligned}$$

the optimal $\xi$ and $\sigma_z^2$ are solutions to the following problem:

$$\begin{aligned}\text{minimize}\quad&\log\left[1+\frac{|H+\xi|^2P}{\sigma_n^2+\sigma_z^2}\right]\\\text{subject to}\quad&|\xi|^2P+\sigma_z^2\le E_J.\end{aligned}\tag{12}$$

It is easy to see that (12) would always be satisfied with equality, so we can reduce the above problem to the following single (complex) variable optimization problem:

$$\begin{aligned}\text{minimize}\quad&\frac{|H+\xi|^2P}{\sigma_n^2+E_J-|\xi|^2P}\\\text{subject to}\quad&|\xi|^2P-E_J\le0.\end{aligned}\tag{13}$$

We summarize the result of the optimization in Lemma 4.

*Lemma 4:* Define $\alpha^2:=|H|^2P$. Then, the optimal jammer strategy is $\boldsymbol{v}^*=\xi^*\boldsymbol{x}+\boldsymbol{z}^*$, where $\xi^*=-k^*H$

$$k^*=\begin{cases}1, & \text{if } E_J\ge\alpha^2\\ \min\left\{\frac{\sqrt{E_J}}{\alpha},\frac{E_J+\sigma_n^2}{\alpha^2}\right\}, & \text{if } E_J<\alpha^2.\end{cases}$$

*Proof:* First observe from (13) that the optimal $\xi$ has to be of the form $-kH$ for some positive scalar $k$. Moreover, if the power constraint of the jammer is loose enough, $E_J\ge|H|^2P=\alpha^2$, then the jammer can completely cancel the communicator's signal by using $k=-1$, and this is indeed the best for the jammer. So, assume that

$E_J < \alpha^2$. For this case, we have reduced the problem to the single (real) variable optimization

$$\begin{aligned} \text{minimize} \quad & \frac{(k-1)^2 P}{\sigma_n^2 + E_J - k^2\alpha^2} \\ \text{subject to} \quad & k^2\alpha^2 - E_J \le 0. \end{aligned} \tag{14}$$

The derivative with respect to $k$ of the objective function vanishes at $k = 1$ and $k = \frac{E_J+\sigma_n^2}{\alpha^2}$. The assumption that $E_J < \alpha^2$ rules out $k = 1$, and so the candidates for the minima are

$$\left\{ \frac{\sqrt{E_J}}{\alpha}, \frac{E_J + \sigma_n^2}{\alpha^2} \right\}$$

with $\frac{\sqrt{E_J}}{\alpha}$ determined from (14). We know that the objective function is continuously differentiable, and is decreasing at the origin. The minimum then depends on the relative positions of the above candidate solutions, and is as given in the lemma. □

Summarizing, we get the following.

*Theorem 2:* The unique saddle point of the game in (4) for the channel in (10) is given by $\boldsymbol{x}^* \sim \mathcal{CN}(0, P)$ and $\boldsymbol{v}^* = \xi^* \boldsymbol{x} + \boldsymbol{z}^*$, where $\xi^*$ is given by Lemma 4, and $\boldsymbol{z}^* \sim \mathcal{CN}(0, (\sigma_z^*)^2)$ with $(\sigma_z^*)^2$ is given by

$$(\sigma_z^*)^2 = \min\{0, E_J - |\xi^*|^2 P\}.$$

*Remark:* Again, uniqueness follows from the interchangeability property, as in Theorem 1.

## V. Conclusion

We have proved that for a MIMO Rayleigh-fading Gaussian channel, a jammer with access to the channel input can inflict only as much damage to communication as one without access to the channel input (when the effectiveness of communication is measured by the mutual information between the channel input and output). The saddle-point strategy of the encoder is to transmit a symmetric CSCG signal, and that of the jammer is to inject a symmetric CSCG signal independent of the transmitter's signal. The equivalent channel seen by the communicator is a Rayleigh-fading channel with additive Gaussian noise, the noise power being the sum of the thermal noise power and the jammer's power.

It is the symmetric nature of fading that brings about this result. We have also illustrated through an example the intuitive result that for a constant channel, the jammer always uses the knowledge of the channel input to inject a signal that is out of phase with the channel input.

## Appendix I
## Proof of Lemma 1

We list this proof here only for completeness. No credit is claimed for it, as it is a straightforward extension of [8, Theorem 1]. The proof relies on the following two lemmas.

*Lemma 5:* [8, Lemma 3] Let $X \in \mathbb{C}^{r\times t}$. Denote the columns of $X$ by $x_i, i = 1, \ldots, t$, i.e., $X = [x_1 x_2 \ldots x_t]$ with $x_i \in \mathbb{C}^r$. Denote all columns of $X$ except the $i$th by $x_{-i}$. For any fixed $x_{-i}$ and $\gamma \ge 0$, define the function $g_0$ of $x_i$ as

$$g_0(x_i|x_{-i}) = \log\det(\gamma I_r + XX^\dagger).$$

Then $g_0$ is convex and symmetric about the origin in $x_i$.

Using Lemma 5, for any fixed $x_{-i}$ and $M \in \mathbb{C}^{r\times r}$, the function of $x_i$

$$g_1(x_i|x_{-i}) = g_0(x_i|Mx_{-i})$$

is convex and symmetric about the origin in $x_i$. Therefore, for any fixed $x_{-i}$ and $\gamma \ge 0$, the function

$$\begin{aligned} g(x_i|x_{-i}) &= g_1(Mx_i|x_{-i}) = g_0(Mx_i|Mx_{-i}) \\ &= \log\det(\gamma I_r + MXX^\dagger M^\dagger) \end{aligned}$$

is convex and symmetric about the origin in $x_i$.

*Lemma 6:* [1, Theorem 1] Let $E$ be a convex set in the $r$-dimensional Euclidean space, symmetric about the origin. For an $r$-vector $\boldsymbol{x}$, let $f(\boldsymbol{x})$ be a function such that i) $f(\boldsymbol{x}) = f(-\boldsymbol{x})$, ii) the set $\{\boldsymbol{x} : f(\boldsymbol{x}) \ge u\}$ is convex for every $0 < u < \infty$, and iii) $\int_E f(\boldsymbol{x})d\boldsymbol{x} < \infty$. Then

$$\int_E f(\boldsymbol{x} + k\boldsymbol{y})d\boldsymbol{x} \ge \int_E f(\boldsymbol{x} + \boldsymbol{y})d\boldsymbol{x}$$

for every $r$-vector $\boldsymbol{y}$ and $0 \le k \le 1$.

*Proof of Lemma 1:* Let $\Lambda_{ii}^{(1)} = \sqrt{\lambda_i^{(1)}}$, $\Lambda_{ii}^{(2)} = \sqrt{\lambda_i^{(2)}}$ for $i = 1, \ldots, \min\{r, t\}$. (Since $\Lambda^{(1)}$ and $\Lambda^{(2)}$ are principal diagonal matrices by assumption, $\Lambda_{ij}^{(l)} = 0$ for all $i \ne j$, $1 \le i \le r$, $1 \le j \le t$, and $l = 1, 2$.) Without loss of generality, we prove Lemma 1 for the case $\lambda_k^{(1)} \ge \lambda_k^{(2)}$ for a particular $k \in \{1, 2, \ldots, \min\{r, t\}\}$ and $\lambda_j^{(1)} = \lambda_j^{(2)}$ for $j = 1, \ldots, (k-1), (k+1), \ldots, \min\{r, t\}$.

Since log is an increasing function, it suffices to show that for any given $u \ge 0$

$$\begin{aligned} &\text{Prob}[\det\{\gamma I_r + M(\boldsymbol{H} + \Lambda^{(2)})(\boldsymbol{H} + \Lambda^{(2)})^\dagger M^\dagger\} \le u] \\ &\qquad \ge \text{Prob}[\det\{\gamma I_r + M(\boldsymbol{H} + \Lambda^{(1)})(\boldsymbol{H} + \Lambda^{(1)})^\dagger M^\dagger\} \le u]. \end{aligned}$$

To that end, define $\Lambda'$ such that $\Lambda'_{kk} = 0$, and $\Lambda'_{ij} = \Lambda_{ij}^{(1)} = \Lambda_{ij}^{(2)}$ for all other $(i, j)$. Use $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_t)$ to denote the columns of $\boldsymbol{H}$.

Define

$$F = \left\{ X \,|\, \det(\gamma I + M(X + \Lambda')(X + \Lambda')^\dagger M^\dagger) \le u \right\}$$

for a given $u \ge 0$. Further, define the set

$$E(x_k|x_{-k}) = \{x_k | X(x_k, x_{-k}) \in F\}$$

where $X(x_k, x_{-k})$ is shorthand for $[x_1, \ldots, x_{k-1}, x_k, x_{k+1}, \ldots, x_t]$.

Then, because of the convexity and symmetry of the function $g(\cdot)$ proved in Lemma 5, the set $E$ is convex and symmetric about the origin. Further, the density of $\boldsymbol{x}_k$, which is the $k$th column of $\boldsymbol{H}$, is $f(\boldsymbol{x}_k) \sim \mathcal{CN}(0, I_r)$, which satisfies conditions i)–iii) of Lemma 6.

So, we have

$$\int_E f\left(\boldsymbol{x}_k + \sqrt{\lambda_k^{(2)}}\boldsymbol{e}_k\right) d\boldsymbol{x}_k \ge \int_E f\left(\boldsymbol{x}_k + \sqrt{\lambda_k^{(1)}}\boldsymbol{e}_k\right) d\boldsymbol{x}_k$$

where $\boldsymbol{e}_k$ is the $k$th unit vector. Multiplying both sides by the joint density of the temporarily fixed columns $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{k-1}, \boldsymbol{x}_{k+1}, \ldots, \boldsymbol{x}_t)$ and integrating with respect to them, we get the desired inequality.

## Appendix II
## Proof of Lemma 3

Proving that $f$ is convex over $\mathcal{S}$ is equivalent to proving that the function

$$g(t) = f(Z + tN)$$

defined over $\mathcal{T} = \{t \in \mathbb{R} | Z + tN \in \mathcal{S}\}$ is convex for any given $Z, N \in \mathcal{S}$.

For $X$ as in the statement of the lemma, and $Z$ and $N$ given as above, define $Y = X + Z$. Then, using [6, Corollary 4.3.3, p. 182], it can be shown that if the eigenvalues $\langle\lambda_i(NZ^{-1})\rangle$ and $\langle\lambda_i(NY^{-1})\rangle$ of the matrices $NZ^{-1}$ and $NY^{-1}$ are put in increasing order, then

$$\lambda_k(NZ^{-1}) \ge \lambda_k(NY^{-1}) \tag{15}$$

with strict inequality for $X \succ 0$.

Now

$$\begin{aligned} g(t) &= \log\det(X+Z+tN) - \log\det(Z+tN) \\ &= \sum_{i=1}^{n} \left\{ \log(1+t\lambda_i(NY^{-1})) \right. \\ &\quad \left. - \log(1+t\lambda_i(NZ^{-1})) \right\} + \log\det(ZY^{-1}) \end{aligned}$$

so that

$$\begin{aligned} \frac{d^2 g}{dt^2} &= \sum_{i=1}^{n} \left\{ \frac{1}{\left(t+\frac{1}{\lambda_i(NZ^{-1})}\right)^2} - \frac{1}{\left(t+\frac{1}{\lambda_i(NY^{-1})}\right)^2} \right\} \\ &\geq 0 \end{aligned}$$

where for the last inequality we have used (15) and the fact that $t \in \mathcal{T}$. Strict inequality holds for $X \succ 0$.

## REFERENCES


[1] T. W. Anderson, "The integral of a symmetric unimodal function over a symmetric convex set and some probability inequalities," *Proc. Amer. Math. Soc.*, vol. 6, pp. 170–176, 1955.

[2] T. Başar, "The Gaussian test channel with an intelligent jammer," *IEEE Trans. Inform. Theory*, vol. IT-29, pp. 152–157, Jan. 1983.

[3] T. Başar and G. J. Olsder, *Dynamic Noncooperative Game Theory*. Philadelphia, PA: SIAM, 1999.

[4] J. M. Borden, D. M. Mason, and R. J. McEliece, "Some information theoretic saddlepoints," *SIAM J. Control Optimiz.*, vol. 23, no. 1, pp. 129–143, 1985.

[5] S. N. Diggavi and T. M. Cover, "The worst additive noise under a covariance constraint," *IEEE Trans. Inform. Theory*, vol. 47, pp. 3072–3081, Nov. 2001.

[6] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K: Cambridge Univ. Press, 1985.

[7] A. Kashyap, T. Başar, and R. Srikant, "Correlated jamming on MIMO Gaussian fading channels," in *Proc. IEEE Int. Conf. Communications (ICC)*, Paris, France, 2004, pp. 458–462.

[8] Y. H. Kim and A. Lapidoth, "On the log-determinant of the noncentral wishart distribution," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Yokohama, Japan, 2003, p. 54.

[9] M. Medard, "Capacity of correlated jamming channels," in *Proc. 35th Allerton Conf.*, 1997, pp. 1043–1052.

[10] W. E. Stark and R. J. McEliece, "On the capacity of channels with block memory," *IEEE Trans. Inform. Theory*, vol. 34, pp. 322–324, Mar. 1988.

[11] İ. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Europ. Trans. Telecommun.*, vol. 10, pp. 585–589, Nov./Dec. 1999.


# Analysis of Multiple-Antenna Wireless Links at Low SNR

Chaitanya Rao, *Student Member, IEEE,* and Babak Hassibi

***Abstract*—Wireless channels with multiple transmit/receive antennas are known to provide a high spectral efficiency both when the channel is known to the receiver, and when the channel is not known to the receiver if the signal-to-noise ratio (SNR) is high. Here we analyze such systems at low SNR, which may find application in sensor networks and other low-power devices. The key point is that, since channel estimates are not reliable, it is often not reasonable to assume that the channel is known at the receiver at low SNR. In this unknown channel case, we show that for sensible input distributions, in particular all practical modulation schemes, the capacity is asymptotically quadratic in the SNR, $\rho$, and thus much less than the known channel case where it exhibits a linear growth in $\rho$. We show that under various signaling constraints, e.g., Gaussian modulation, unitary space–time modulation, and peak constraints, that mutual information is maximized by using a single transmit antenna. We also show that at low SNR, sending training symbols leads to a rate reduction in proportion to the fraction of training duration time so that it is best not to perform training. Furthermore, we show that the per-channel use mutual information is linear in both the number of receive antennas and the channel coherence interval.**

***Index Terms*—Low-signal-to-noise ratio (SNR) regime, multiple-antenna systems, noncoherent channels, Rayleigh fading.**

## I. INTRODUCTION

Multiple-antenna wireless systems have been shown to provide high capacity, exploiting the presence of fading in such channels. However, this is based on the premise that either the channel coefficients are known to the receiver, or that the signal-to-noise ratio (SNR) of the channel is high [1]–[3].

Wireless systems operating at low SNR (exhibiting weak signaling or in noisy environments) find increasing use in energy-efficient devices such as sensor networks. Recent work on analyzing the capacity of low-SNR multiple-antenna links, assuming that the channel is known at the receiver, has appeared in [4]. However, at low SNR, channel estimates in some circumstances are unreliable and so it is sensible to assume that the channel is unknown. In the following analysis we, therefore, assume the channel is unknown to both transmitter and receiver. As shown later, this leads to results qualitatively different from the known channel case.

We use the block-fading model of a wireless multiple-antenna system proposed by Marzetta and Hochwald in [5], expressing the mutual information between input and output as a function of the model parameter $\rho$ (proportional to the SNR) up to second order. This model is described in detail in the next section. Maximizing this expression gives us insight about desired signaling at low SNR as well as the optimal number of antennas to be used at the transmitter and receiver. It has been shown in [6] that the optimum signaling at low SNR achieves the same minimum energy per bit as the known channel cases for single transmit antenna systems. We show that the on–off optimal signaling found in [6] also generalizes to the multiple-antenna setting (a result that also follows from [7, Theorems

Manuscript received March 10, 2003; revised February 1, 2004. This work was supported in part by the National Science Foundation under Grant CCR-0133818, by the Office of Naval Research under Grant N00014-02-1-0578, and by Caltech's Lee Center for Advanced Networking.

The authors are with the Department of Electrical Engineering, California Institute of Technology, Pasadena CA 91125 USA (e-mail: rao@systems.caltech.edu; hassibi@systems.caltech.edu).

Communicated by G. Caire, Associate Editor for Communications.

Digital Object Identifier 10.1109/TIT.2004.833369