MAXIMIN ROBUST DESIGN FOR MIMO COMMUNICATION SYSTEMS AGAINST IMPERFECT CSIT

Jiaheng Wang and Daniel P. Palomar

Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Hong Kong

ABSTRACT

This paper considers robust transmit strategies, against the imperfectness of CSIT, for MIMO communication systems. Following a deterministic model that assumes the actual channel inside an ellipsoid centered at a nominal channel, we maximize the worst-case received SNR. It is shown that, for a general class of power constraints, the resulting maximin problem can be equivalently transformed into a convex problem, or even further into a semidefinite program. The most important result is that the optimal transmit directions are just the right singular vectors of the nominal channel under some mild conditions. This result reduces the complicated matrix-valued problems to scalar power allocation problems, for which the closed-form solutions are provided.

***Index Terms*—** MIMO, imperfect CSIT, worst-case robust designs, maximin, convex optimization, SDP.

1. INTRODUCTION

The performance of Multi-input multi-output (MIMO) systems depends, to a substantial extent, on the quality of the channel state information (CSI). In case of no CSI at the transmitter (CSIT), space-time coding techniques [1] can be used to harvest the diversity gain. When the transmitter knows the channel perfectly, on the other hand, the full benefit of CSIT is exploited by precoding techniques [2, 3]. Instead of these two extreme assumptions on CSIT, a practical communication system typically has to confront an intermediate case, i.e., CSIT available but imperfect.

There are two classes of models frequently used to characterize imperfect CSI: the stochastic and the deterministic (or worst-case) models. In the stochastic model, the channel is usually modeled as a complex random matrix with the mean and/or the covariance available at the transmitter. The system design is then based on optimizing the average or outage performance [4–7]. On the other hand, the deterministic model assumes that the instantaneous channel, though not exactly known, lies in a known set of possible values. In this case, the robust design [8–11] aims at optimizing the worst-case performance, and achieves a guaranteed performance level for any channel realization in the set.

In this paper, we consider robust transmit strategies for MIMO communication systems, based on the worst-case optimization, using the deterministic model of imperfect CSIT that is similar to (but more general than) that used in [8, 11]. Specifically, while assuming perfect CSI at the receiver, for the transmitter, we assume that the actual channel lies in an ellipsoid centered at a nominal channel. The design objective is to maximize the worst-case received signal-to-noise ratio (SNR) [10, 11], or to minimize the worst-case Chernoff bound of the pairwise error probability (PEP) [6] if a space-time block code (STBC) [1] is used. Our first main result is that, for a general class of power constraints, the formulated maximin problem can be equivalently transformed into a convex optimization problem that can be efficiently solved in polynomial time. For some power constraints, the problem simplifies further to a semidefinite program (SDP) [13], a very tractable form of convex optimization.

In light of the optimality of the channel-diagonalizing structure in the cases of perfect CSIT [2, 3] and statistical CSIT with mean or covariance feedback [4–7], one may wonder whether it still holds in the deterministic model. As the second main contribution of this paper, we answer affirmatively this question by proving that, for the worst-case design, the optimal transmit directions are the right singular vectors of the nominal channel under some mild conditions. As a special case of our framework, it follows that the transmit directions imposed in [10, 11] (without any justification) are actually optimal. Consequently, the complicated matrix-valued problems can be simplified to scalar power allocation problems without any loss of optimality. Our third main result consists of providing the closed-form solutions to the resulting power allocation problems.

2. PROBLEM STATEMENT

We consider a point-to-point communication system equipped with N transmit and M receive antennas. Mathematically, the system can be represented by a linear model as $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}$, where $\mathbf{x} \in \mathbb{C}^N$ is the transmitted signal vector, $\mathbf{H} \in \mathbb{C}^{M\times N}$ is the channel matrix, $\mathbf{y} \in \mathbb{C}^M$ is the received signal vector, and $\mathbf{n} \in \mathbb{C}^M$ is a circularly symmetric complex Gaussian noise vector with zero mean and covariance matrix $\sigma_n^2\mathbf{I}$, i.e., $\mathbf{n} \sim \mathcal{CN}\left(\mathbf{0}, \sigma_n^2\mathbf{I}\right)$.

To model imperfect CSI deterministically, we assume that $\mathbf{H}$ belongs to the elliptical uncertainty region $\mathcal{H} \triangleq \left\{\mathbf{H} : \left\|\mathbf{H} - \hat{\mathbf{H}}\right\| \le \varepsilon\right\}$ centered at the nominal channel $\hat{\mathbf{H}}$. Furthermore, by defining the channel error $\mathbf{\Delta} \triangleq \mathbf{H} - \hat{\mathbf{H}}$, $\mathbf{H} \in \mathcal{H}$ can be equally described by $\mathbf{\Delta} \in \mathcal{E} = \{\mathbf{\Delta} : \|\mathbf{\Delta}\| \le \varepsilon\}$. In this paper, we consider $\mathcal{E}$ defined by two matrix norms: the weighted Frobenius norm $\|\cdot\|_F^{\mathbf{T}}$ and the weighted spectral norm $\|\cdot\|_2^{\mathbf{T}}$. To be more specific, $\mathcal{E}$ could be:

$$\mathcal{E}_F \triangleq \left\{\mathbf{\Delta} : \|\mathbf{\Delta}\|_F^{\mathbf{T}} \le \varepsilon\right\} = \left\{\mathbf{\Delta} : \mathrm{Tr}\left(\mathbf{\Delta}\mathbf{T}\mathbf{\Delta}^H\right) \le \varepsilon^2\right\} \quad (1)$$

$$\mathcal{E}_2 \triangleq \left\{\mathbf{\Delta} : \|\mathbf{\Delta}\|_2^{\mathbf{T}} \le \varepsilon\right\} = \left\{\mathbf{\Delta} : \mathbf{\Delta}\mathbf{T}\mathbf{\Delta}^H \preceq \varepsilon^2\mathbf{I}\right\} \quad (2)$$

where $\mathbf{T}$ is a known positive definite matrix.

Let $\mathbf{Q} = E\left[\mathbf{x}\mathbf{x}^H\right]$ be the covariance matrix of the transmitted signal vector. Regarding the power constraints at the transmitter, we will start by considering a very general constraint $\mathbf{Q} \in \mathcal{Q}$, where $\mathcal{Q} \subseteq \mathbb{S}_+^N$ is a nonempty compact convex set. In other words, $\mathbf{Q}$ must be positive semidefinite and within a nonempty compact convex set. This constraint includes all commonly used power constraints as special cases. Here we list some of them: (1) Sum power constraint $\mathcal{Q}_1 \triangleq \{\mathbf{Q} : \mathbf{Q} \succeq 0,\ \mathrm{Tr}(\mathbf{Q}) \le P_s\}$; (2) Maximum power constraint $\mathcal{Q}_2 \triangleq \{\mathbf{Q} : \mathbf{Q} \succeq 0,\ \lambda_{\max}(\mathbf{Q}) \le P_m\}$; (3) Per-antenna power constraint $\mathcal{Q}_3 \triangleq \{\mathbf{Q} : \mathbf{Q} \succeq 0,\ \max_i [\mathbf{Q}]_{ii} \le P_a\}$ or $\mathcal{Q}_4 \triangleq \{\mathbf{Q} : \mathbf{Q} \succeq 0,\ [\mathbf{Q}]_{ii} \le P_{a,i},\ i = 1, \ldots, N\}$.

Given that the system performance is determined by both the transmit strategy and the channel, we adopt the following perfor-

This work was supported by the RGC 618008 research grant.

mance measure:

$$\Psi(\mathbf{Q},\mathbf{H}) = \operatorname{Tr}\left(\mathbf{H}\mathbf{Q}\mathbf{H}^H\right). \tag{3}$$

It can be verified that maximizing $\Psi(\mathbf{Q},\mathbf{H})$ have several meanings [12]: (1) maximizing the received SNR; (2) minimizing the Chernoff Bound of the PEP when a STBC is used; (3) approximately maximizing the mutual Information at low SNR; (4) approximately minimizing the mean square error (MSE) at low SNR if a minimum MSE (MMSE) equalizer is used at the receiver.

Based on the worst-case philosophy, our robust transmitter design can be formulated as the following maximin problem:

$$\max_{\mathbf{Q}\in\mathcal{Q}} \min_{\mathbf{\Delta}\in\mathcal{E}} \operatorname{Tr}\left[\left(\hat{\mathbf{H}}+\mathbf{\Delta}\right)\mathbf{Q}\left(\hat{\mathbf{H}}+\mathbf{\Delta}\right)^H\right]. \tag{4}$$

Note that the optimum value of (4) is zero if and only if $\mathbf{\Delta} = -\hat{\mathbf{H}}$. For $\mathcal{E} = \mathcal{E}_F$ or $\mathcal{E} = \mathcal{E}_2$, this can only happen if $\varepsilon \geq \left\|\hat{\mathbf{H}}\right\|_F^{\mathbf{T}}$ or $\varepsilon \geq \left\|\hat{\mathbf{H}}\right\|_2^{\mathbf{T}}$, respectively, i.e., when the nominal channel is very small. In such cases, there is no guarantee of performance. To avoid the trivial solution, we assume that $\varepsilon < \left\|\hat{\mathbf{H}}\right\|_F^{\mathbf{T}}$ for $\mathcal{E} = \mathcal{E}_F$ and $\varepsilon < \left\|\hat{\mathbf{H}}\right\|_2^{\mathbf{T}}$ for $\mathcal{E} = \mathcal{E}_2$.

3. CONVEX REFORMULATION OF THE MAXIMIN PROBLEM

In this section, we consider the general power constraint $\mathbf{Q} \in \mathcal{Q}$ where $\mathcal{Q} \subseteq \mathbb{S}_+^N$ is a nonempty compact convex set. In this case, the solution to the maximin problem (4) is the saddle point of the objective function, and the question is how to characterize and compute it. We show that this can be achieved by reformulating (4) as a convex problem, which can be globally solved by efficient polynomial-time numerical algorithms. Moreover, when some specific power constraints are considered, the resulting convex problem simplifies to an SDP [13].

Proposition 1 ([12]) *Let $\mathcal{E} = \mathcal{E}_F$ and $\mathcal{Q} \subseteq \mathbb{S}_+^N$ be a nonempty compact convex set. Then, the maximin problem (4) is equivalent to the following convex problem:*

$$\begin{array}{ll} \underset{\mathbf{Q},\mu,\mathbf{Z}}{\text{minimize}} & \operatorname{Tr}\left[(\mathbf{Z}-\mathbf{Q})\hat{\mathbf{H}}^H\hat{\mathbf{H}}\right] + \varepsilon^2\mu \\ \text{subject to} & \mathbf{Q} \in \mathcal{Q} \\ & \mu \geq 0 \\ & \begin{bmatrix} \mathbf{Z} & \mathbf{Q} \\ \mathbf{Q} & \mathbf{Q}+\mu\mathbf{T} \end{bmatrix} \succeq 0. \end{array} \tag{5}$$

Proposition 2 ([12]) *Let $\mathcal{E} = \mathcal{E}_2$ and $\mathcal{Q} \subseteq \mathbb{S}_+^N$ be a nonempty compact convex set. Then, the maximin problem (4) is equivalent to the following convex problem:*

$$\begin{array}{ll} \underset{\mathbf{Q},\mu,\mathbf{Z}}{\text{maximize}} & \operatorname{Tr}(\mathbf{Z}) \\ \text{subject to} & \mathbf{Q} \in \mathcal{Q} \\ & \mu \geq 0 \\ & \begin{bmatrix} \hat{\mathbf{H}}\mathbf{Q}\hat{\mathbf{H}}^H - \mathbf{Z} - \mu\mathbf{I} & \hat{\mathbf{H}}\mathbf{Q} \\ \mathbf{Q}\hat{\mathbf{H}}^H & \mathbf{Q}+\frac{\mu}{\varepsilon^2}\mathbf{T} \end{bmatrix} \succeq 0. \end{array} \tag{6}$$

Remark 1. Propositions 1 and 2 are based on *S-procedure* [14]. It is straightforward to see that (5) and (6) become SDPs if $\mathcal{Q} = \mathcal{Q}_1$ or $\mathcal{Q} = \mathcal{Q}_4$. When $\mathcal{Q} = \mathcal{Q}_2$, we can easily transform (5) and (6) into SDPs, since the constraint $\lambda_{\max}(\mathbf{Q}) \leq P_m$ is equivalent to $\mathbf{Q} \preceq P_m\mathbf{I}$. Similarly, when $\mathcal{Q} = \mathcal{Q}_3$, the constraint $\max_i[\mathbf{Q}]_{ii} \leq P_{a,i}$ can be replaced by $[\mathbf{Q}]_{ii} \leq P_{a,i}$, $i = 1,\ldots,N$. In fact, for any combination of $\mathcal{Q}_1$, $\mathcal{Q}_2$, $\mathcal{Q}_3$, and $\mathcal{Q}_4$ (i.e., any intersection among them), the convex problems (5) and (6) become SDPs.

4. OPTIMAL TRANSMIT DIRECTIONS

Although the jointly optimal transmit directions and power allocation can be achieved by decomposing the optimal transmit covariance matrix obtained through solving (5) or (6), one may wonder whether they can be obtained independently. Even further, one may ask whether there exist optimal channel-diagonalizing transmit directions, just like the cases of perfect CSIT or some statistical CSIT, that can reduce the complicated matrix-valued optimization to a simple power allocation problem. As an important result, we show that the optimal transmit directions are just the right singular vectors of the nominal channel matrix, provided some conditions are satisfied. We start by giving the following result.

Proposition 3 *Let $\mathcal{E} = \mathcal{E}_F$ and $\mathcal{Q} \subseteq \mathbb{S}_+^N$ be a nonempty compact set. Then, the maximin problem (4) is equivalent to*

$$\underset{\mathbf{Q}\in\mathcal{Q},\mu\geq 0}{\text{maximize}} \quad \mu \operatorname{Tr}\left[\mathbf{Q}\left(\mathbf{Q}+\mu\mathbf{T}\right)^{-1}\mathbf{T}\hat{\mathbf{H}}^H\hat{\mathbf{H}}\right] - \varepsilon^2\mu \tag{7}$$

where the objective is defined as 0 for $\mu = 0$.

Proof: The basic idea is to replace the inner minimization of (4) by its dual maximization, hence transforming the maximin problem to a maximization problem. See details in [12]. ■

Let the eigenvalue decomposition (EVD) of $\mathbf{Q}$ be $\mathbf{Q} = \mathbf{U}_q\mathbf{\Lambda}_q\mathbf{U}_q^H$ with the eigenvalues $p_1 \geq \cdots \geq p_N$, the EVD of $\hat{\mathbf{H}}^H\hat{\mathbf{H}}$ be $\hat{\mathbf{H}}^H\hat{\mathbf{H}} = \mathbf{U}_h\mathbf{\Lambda}_h\mathbf{U}_h^H$ with the eigenvalues $\gamma_1 \geq \cdots \geq \gamma_N$, and the EVD of $\mathbf{T}$ be $\mathbf{T} = \mathbf{U}_t\mathbf{\Lambda}_t\mathbf{U}_t^H$ with the eigenvalues $\tau_1 \geq \cdots \geq \tau_N$. Denote by $\boldsymbol{\lambda}(\mathbf{A})$ the vector consisting of the eigenvalues of a square matrix $\mathbf{A}$. We are now ready to state the main results.

Theorem 1 *Let $\mathcal{E} = \mathcal{E}_F$ with $\mathbf{T} = \tau\mathbf{I}$ and $\tau > 0$, and $\mathcal{Q} \subseteq \mathbb{S}_+^N$ be a nonempty compact set defined by constraining $\mathbf{Q}$ only through its eigenvalues. Then, $\mathbf{U}_q = \mathbf{U}_h$ is optimal for the maximin problem (4).*

Proof: From Proposition 3, when $\mu = 0$, (4) has a zero objective value and any $\mathbf{Q} \in \mathcal{Q}$ is optimal, so in particular $\mathbf{U}_q = \mathbf{U}_h$ is optimal as well. When $\mu > 0$, the maximin problem (4) with $\mathbf{T} = \tau\mathbf{I}$ is equivalent to

$$\underset{\mathbf{Q}\in\mathcal{Q},\mu>0}{\text{maximize}} \quad \mu\tau \operatorname{Tr}\left[\mathbf{Q}\left(\mathbf{Q}+\mu\tau\mathbf{I}\right)^{-1}\hat{\mathbf{H}}^H\hat{\mathbf{H}}\right] - \varepsilon^2\mu. \tag{8}$$

Lemma 1 (9.H.1.g, 9.H.1.h [15]) *Let $\mathbf{A}$ and $\mathbf{B}$ be two $N \times N$ positive semidefinite matrices, with eigenvalues $\alpha_1 \geq \cdots \geq \alpha_N$ and $\beta_1 \geq \cdots \geq \beta_N$, respectively. Then,*

$$\sum_{i=1}^N \alpha_i\beta_{N-i+1} \leq \operatorname{Tr}(\mathbf{AB}) \leq \sum_{i=1}^N \alpha_i\beta_i.$$

According to Lemma 1, it follows that

$$\begin{aligned} & \mu\tau \operatorname{Tr}\left[\mathbf{Q}\left(\mathbf{Q}+\mu\tau\mathbf{I}\right)^{-1}\hat{\mathbf{H}}^H\hat{\mathbf{H}}\right] \\ = \; & \mu\tau \operatorname{Tr}\left[\mathbf{\Lambda}_q\left(\mathbf{\Lambda}_q+\mu\tau\mathbf{I}\right)^{-1}\mathbf{U}_q^H\mathbf{U}_h\mathbf{\Lambda}_h\mathbf{U}_h^H\mathbf{U}_q\right] \\ \leq \; & \mu\tau \sum_{i=1}^N \frac{p_i}{\mu\tau+p_i}\gamma_i = \sum_{i=1}^N \frac{\mu\tau\gamma_i p_i}{\mu\tau+p_i} \end{aligned} \tag{9}$$

where the equality holds when $\mathbf{U}_q = \mathbf{U}_h$. Since the power constraint does not depend on $\mathbf{U}_q$, we can always choose $\mathbf{U}_q = \mathbf{U}_h$ to maximize the objective without affecting the constraints. ■

Theorem 2 ([12]) *Let $\mathcal{E} = \mathcal{E}_F$ with $\mathbf{T}$ such that $\mathbf{U}_t = \mathbf{U}_h$, and*

$$\mathcal{Q} = \{\mathbf{Q} : \mathbf{Q} \succeq 0, f_k(\boldsymbol{\lambda}(\mathbf{Q})) \le P_k, k = 1, \cdots, K\}$$

where each $f_k(\mathbf{x})$ is a Schur-convex function. Then, $\mathbf{U}_q = \mathbf{U}_h$ is optimal for the maximin problem (4).

Remark 2. Theorem 1 indicates that $\mathbf{U}_q = \mathbf{U}_h$ is optimal with a general power constraint relying only on the eigenvalues of $\mathbf{Q}$, provided the uncertainty region is a sphere defined by the Frobenius norm, which is the most frequently used deterministic model [8, 11]. When more restrictions are added to the eigenvalues of $\mathbf{Q}$, Theorem 2 shows that $\mathbf{U}_q = \mathbf{U}_h$ is optimal for an ellipsoid uncertainty region if $\hat{\mathbf{H}}^H\hat{\mathbf{H}}$ and $\mathbf{T}$ can be simultaneously diagonalized. Given that both $\mathrm{Tr}(\mathbf{Q})$ and $\lambda_{\max}(\mathbf{Q})$ are Schur-convex functions of the eigenvalues of $\mathbf{Q}$ [3], Theorem 2 is applicable to the two most common constraints $\mathcal{Q}_1$ (the sum power constraint) and $\mathcal{Q}_2$ (the maximum power constraint) as well as their intersection. Therefore, in most cases, the optimality of the eigenmode transmission (over the nominal channel) still holds for the worst-case design, which complies with the cases of perfect CSIT and statistical CSIT with mean or covariance feedback. Note that the problems considered in [10] (for the spherical uncertainty region) and [11] are a special case of our framework with $\mathcal{Q} = \mathcal{Q}_1$ and $\mathbf{T} = \mathbf{I}$. However, they assumed $\mathbf{U}_q = \mathbf{U}_h$ without knowing whether this is optimal or not, even if [11] offered some sufficient conditions. Consequently, by using Theorem 1 or 2, the complicated matrix-valued maximin problem (4) can be simplified to a scalar power allocation problem.

5. OPTIMAL POWER ALLOCATION

Let $r = \mathrm{rank}(\hat{\mathbf{H}})$. When $\mathcal{E} = \mathcal{E}_F$ with $\mathbf{T}$ such that $\mathbf{U}_t = \mathbf{U}_h$ and $\mathcal{Q}$ is a nonempty compact convex set satisfying the condition in Theorem 2, from (7), the maximin problem (4) can be simplified to

$$\underset{\{p_i\}:\mathbf{Q}\in\mathcal{Q},\mu\ge 0}{\text{maximize}} \quad \sum_{i=1}^{r} \frac{\mu\tau_i\gamma_i p_i}{\mu\tau_i + p_i} - \varepsilon^2\mu. \tag{10}$$

It can be verified that (10) is a convex problem, thus admitting globally optimal solutions that can be efficiently found. Similarly, when $\mathcal{E} = \mathcal{E}_F$ with $\mathbf{T} = \tau\mathbf{I}$ and $\tau > 0$ and $\mathcal{Q}$ is a nonempty compact convex set satisfying the condition in Theorem 1, the resulting power allocation problem is a convex problem as well.

5.1. Sum Power Constraint $\mathcal{Q} = \mathcal{Q}_1$ and $\mathcal{E} = \mathcal{E}_F$ with $\mathbf{T}$ such that $\mathbf{U}_t = \mathbf{U}_h$

In this case, the maximin problem (4) can be simplified to

$$\begin{aligned} \underset{\{p_i\},\mu}{\text{maximize}} \quad & \sum_{i=1}^{r} \frac{\mu\tau_i\gamma_i p_i}{\mu\tau_i + p_i} - \varepsilon^2\mu \\ \text{subject to} \quad & \sum_{i=1}^{N} p_i = P_s \\ & p_1 \ge \cdots \ge p_N \ge 0 \\ & \mu \ge 0 \end{aligned} \tag{11}$$

where we explicitly write the decreasing order of $\{p_i\}$. The solution to this problem is given by the following theorem:

Theorem 3 ([12]) *The solution to the problem (11) is*

$$p_i^\star = \begin{cases} \tau_i\left[\sqrt{\frac{\gamma_i}{b_k}}\left(P_s + c_k\mu^\star\right) - \mu^\star\right], & \text{for } i = 1, \cdots, k \\ 0, & \text{for } i > k \end{cases} \tag{12}$$

with

$$\mu^\star = \frac{P_s}{c_k}\left(\sqrt{\frac{b_k}{b_k - a_k c_k}} - 1\right) \tag{13}$$

where $a_m \triangleq \sum_{j=1}^{m}\tau_j\gamma_j - \varepsilon^2$, $b_m \triangleq \left(\sum_{j=1}^{m}\tau_j\sqrt{\gamma_j}\right)^2$, $c_m \triangleq \sum_{j=1}^{m}\tau_j$ and $\beta_m \triangleq \sum_{j=1}^{m}\tau_j\left(\sqrt{\gamma_j} - \sqrt{\gamma_m}\right)^2$ for $m = 1, \cdots, r$, and $\beta_{r+1} \triangleq +\infty$, and k is such an integer that $\beta_k < \varepsilon^2 \le \beta_{k+1}$. The optimum value of (11) is

$$\frac{P_s}{c_k^2}\left(\sqrt{b_k} - \sqrt{b_k - a_k c_k}\right)^2. \tag{14}$$

Corollary 1 *For $\mathcal{Q} = \mathcal{Q}_1$ and $\mathcal{E} = \mathcal{E}_F$ with $\mathbf{T}$ such that $\mathbf{U}_t = \mathbf{U}_h$, the robust transmitter uses only one eigenmode if*

$$\varepsilon \le \sqrt{\tau_1}\left(\sqrt{\gamma_1} - \sqrt{\gamma_2}\right). \tag{15}$$

Corollary 2 *As $\varepsilon \to \left\|\hat{\mathbf{H}}\right\|_F^{\mathbf{T}}$, the solution to the problem (11) becomes*

$$p_i^\star = \frac{\tau_i\sqrt{\gamma_i}}{\sum_{j=1}^{r}\tau_j\sqrt{\gamma_j}}P_s, \; i = 1, \cdots, r. \tag{16}$$

Remark 3. According to Theorem 3, the robust transmitter will use multiple eigenmodes to increase the reliability in the worst-case channel. The larger the error radius ε is, i.e., the more uncertainty, the more eigenmodes will be used. Corollary 1 indicates that beamforming along one direction is robust if ε is very small, or if the difference between the largest two singular values of the nominal channel is very large, which implies a nearly rank-one channel. Interestingly, the similar result on the number of used eigenmodes was also obtained in [10, 11]. However, In contrast to the semi-closed-form solutions in [10, 11], we offer the fully analytical solution in a more general case. Furthermore, from Corollary 2, as ε approaches $\left\|\hat{\mathbf{H}}\right\|_F^{\mathbf{T}}$, the worst-case design tends to allocate the transmit power according to the weighted proportion of a singular value of the nominal channel, instead of a uniform distribution. This has been observed in [10, 11] through numerical simulations, but no proper explanation was given. The fundamental reason is that the deterministic model adopted in this paper is not an isotropically unconstrained set [9], but an ellipsoid with the center, i.e., the nominal channel, away from the origin.

5.2. Maximum Power Constraint $\mathcal{Q} = \mathcal{Q}_2$ and $\mathcal{E} = \mathcal{E}_F$ with $\mathbf{T}$ such that $\mathbf{U}_t = \mathbf{U}_h$

The corresponding power allocation problem is

$$\begin{aligned} \underset{\{p_i\},\mu}{\text{maximize}} \quad & \sum_{i=1}^{r} \frac{\mu\tau_i\gamma_i p_i}{\mu\tau_i + p_i} - \varepsilon^2\mu \\ \text{subject to} \quad & P_m \ge p_1 \ge \cdots \ge p_N \ge 0 \\ & \mu \ge 0. \end{aligned} \tag{17}$$

Due to the monotonicity, it is easy to see that the optimal power allocation is $p_i^\star = P_m$, $i = 1, \cdots, r$ and $p_i^\star = 0$, $i > r$. That is, a uniform distribution with the maximum power on each nonzero eigenmode.

To obtain the optimal objective value, we need to solve

$$\underset{\mu\ge 0}{\text{maximize}} \quad \sum_{i=1}^{r} \frac{\mu\tau_i\gamma_i P_m}{\mu\tau_i + P_m} - \varepsilon^2\mu \tag{18}$$

which is a convex problem since the objective function is strictly concave in μ. The optimal μ can be found by setting the derivative of the objective function to be zero, which turns to finding the positive root of the following equation:

$$w(\mu) = \sum_{i=1}^{r} \frac{\tau_i \gamma_i P_m^2}{(\mu \tau_i + P_m)^2} = \varepsilon^2. \tag{19}$$

Unfortunately, this equation does not admit an analytical root. But we can resort to the bisection method by exploiting the monotonicity of $w(\mu)$.

In the case of $\mathbf{T} = \tau \mathbf{I}$, a closed-form solution to (18) is available

$$\mu^\star = \frac{P_m}{\tau} \left(\frac{\sqrt{\tau \sum_{i=1}^{r} \gamma_i}}{\varepsilon} - 1 \right) \tag{20}$$

which leads to the optimum value $P_m \left(\sqrt{\sum_{i=1}^{r} \gamma_i} - \frac{\varepsilon}{\sqrt{\tau}} \right)^2$.

6. NUMERICAL RESULTS

The philosophy of robustness in this paper is to guarantee a performance level for any channel realization in the uncertainty region. In other words, we are interested in the worst-case behavior. Therefore, we compare the worst-case performance of the different transmit strategies, including the robust approach, the beamforming strategy which transmits only over the maximum eigenvalue of the nominal channel, and the equal-power transmission which allocates the transmit power equally over all eigenmodes. For simplicity, we consider $\mathcal{Q} = \mathcal{Q}_1$ and $\mathcal{E} = \mathcal{E}_F$ with $\mathbf{T} = \mathbf{I}$. To take into account different channels, the elements of the nominal channel $\hat{\mathbf{H}}$ are randomly generated according to zero-mean, unit-variance, i.i.d. Gaussian distributions.

One thing worth pointing out is that even if we assume $\varepsilon < \|\hat{\mathbf{H}}\|_F$, there is still a probability that beamforming could not guarantee any performance in the uncertainty region because of $\|\hat{\mathbf{H}}\|_2 \leq \|\hat{\mathbf{H}}\|_F$. Therefore, to make sure that all three transmit strategies can work in their worst-case situations, we set $\varepsilon = s \|\hat{\mathbf{H}}\|_2$ with the parameter $s \in [0, 1)$. Nevertheless, it is possible that ε is a small proportion to $\|\hat{\mathbf{H}}\|_F$ even if s tends to 1.

In Fig. 1, we plot the symbol error rates (SERs) of the three transmit strategies in their worst-case channels versus SNR for different values of s. With four antennas equipped at both ends of the link, i.e., $M = N = 4$, the QPSK modulated symbols are encoded at the transmitter according to a 3/4-rate complex OSTBC introduced in [1], and decoded by a ML detector at the receiver. The worst-case SER is averaged over $\hat{\mathbf{H}}$. As observed from Fig. 1, the robust approach offers the lowest worst-case SER among all transmit strategies, which complies with our design objective. When s is small, i.e., the channel error is small, the performance of beamforming is close to that of the robust approach. This is consistent with Corollary 1 which says that when ε is small the robust strategy coincides with beamforming. On the other hand, as s increases (so does ε), the performance gap between beamforming and the robust approach becomes larger, and eventually beamforming is outperformed by the equal-power transmission.

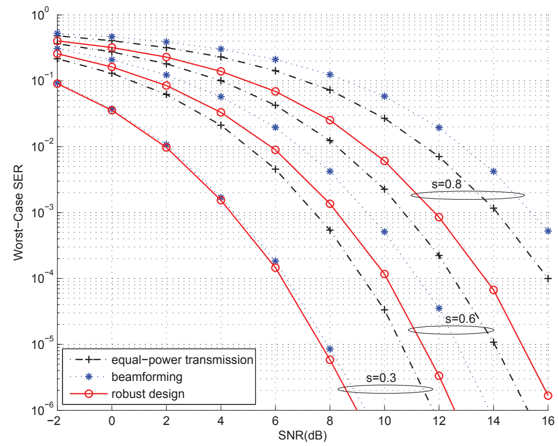

Fig. 1. Worst-case SER versus SNR with different values of s for $M = N = 4$.

7. REFERENCES

[1] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block coding for wireless communications: Performance results," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 3, pp. 451–460, Mar. 1999.

[2] D. P. Palomar, J. M. Cioffi, and M. A. Lagunas, "Joint Tx-Rx beamforming design for multicarrier MIMO channels: A unified framework for convex optimization," in *IEEE Trans. Signal Processing*, vol. 51, no. 9, pp. 2381–2401, Sep. 2003.

[3] D. P. Palomar and Y. Jiang, "MIMO transceiver design via majorization theory," *Foundations and Trends in Communications and Information Theory*, vol. 3, no. 4-5, pp. 331–551, 2006.

[4] S. A. Jafar and A. Goldsmith, "Transmitter optimization and optimality of beamforming for multiple antenna systems," *IEEE Trans. Wireless Commun.*, vol. 3, no. 4, pp. 1165–1175, Jul. 2004.

[5] S. Zhou and G. B. Giannakis, "Optimal transmitter eigen-beamforming and space-time block coding based on channel mean feedback," *IEEE Trans. Signal Processing*, vol. 50, no. 10, pp. 2599–2613, Oct. 2002.

[6] G. Jöngren, M. Skoglund, and B. Ottersten, "Combining beamforming and orthogonal space-time block coding," *IEEE Trans. Inform. Theory*, vol. 48, no. 3, pp. 611–627, Mar. 2002.

[7] E. A. Jorswieck, D. Hammarwall, and B. Ottersten, "Ergodic capacity achieving transmit strategy in MIMO systems with statistical and short-term norm CSI," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2006.

[8] Y. C. Eldar and N. Merhav, "A competitive minimax approach to robust estimation of random parameters," *IEEE Trans. Signal Processing*, vol. 52, no. 7, pp. 1931–1946, Jul. 2004.

[9] D. P. Palomar, J. M. Cioffi, and M. A. Lagunas, "Uniform power allocation in MIMO channels: A game theoretic approach," *IEEE Trans. Inform. Theory*, vol. 49, no. 7, pp. 1707–1727, Jul. 2003.

[10] A. Pascual-Iserte, D. P. Palomar, A. I. Pérez-Neira, and M. A. Lagunas, "A robust maximin approach for MIMO communications with partial channel state information based on convex optimization," *IEEE Trans. Signal Processing*, vol. 54, no. 1, pp. 346–360, Jan. 2006.

[11] A. Abdel-Samad, T. N. Davidson, and A. B. Gershman, "Robust transmit eigen beamforming based on imperfect channel state information," *IEEE Trans. Signal Processing*, vol. 54, no. 5, pp. 1596–1609, May 2006.

[12] J. Wang and D. P. Palomar, "Worst-case robust MIMO transmission with imperfect channel knowledge," accepted in *IEEE Trans. Signal Processing*, submitted Aug. 2008.

[13] A. Ben-Tal and A. Nemirovski, *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*, MPS-SIAM Series on Optimization, 2001.

[14] S. Boyd, L. E. Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*, SIAM Studies in Applied Mathematics, Philadelphia, 1994.

[15] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*, Academic Press, New York, 1979.