Is Transmit Beamforming Robust?

Jiaheng Wang
ACCESS Linnaeus Center
KTH Royal Institute of Technology
SE-100 44 Stockholm, Sweden
Email: jiah@kth.se

Miquel Payaró
Centre Tecnològic de Telecomunicacions de Catalunya
Parc Mediterrani de la Tecnologia
Castelldefels, Barcelona, Spain
Email: miquel.payaro@cttc.es

***Abstract*—Transmit beamforming is a simple strategy that uses only one spatial direction in a multiple-input multiple-output (MIMO) channel. Due to its simplicity, beamforming is usually seen sensitive to inaccuracies of the channel state information at the transmitter (CSIT). In this paper, we investigate the robustness of beamforming with respect to the worst-case CSIT. Our result shows that beamforming is actually robust, in the sense that it achieves the maximum received signal-to-noise ratio (SNR) in the worst channel within an elliptical uncertainty region defined by the weighted spectral norm. This result implies that beamforming has the ability to combat against the imperfectness of CSIT, especially when channel dimensions or channel uncertainty are small.**

I. Introduction

The performance of a multiple-input multiple-output (MIMO) system depends, to a substantial extent, on the quality of the channel state information (CSI). Given perfect CSI at the transmitter (CSIT), the optimal transmit strategies, under various criteria, have been well studied [1]–[3]. The simplest one of them is beamforming, which consists of transmitting data only through one spatial direction, i.e., one eigenvector of the transmit covariance matrix. The simplicity of beamforming makes it attractive in practice, but at the same time sparks worries on its sensitivity to inaccuracies of CSIT [2], [3]. In practice, however, CSIT is often imperfect due to, e.g., inaccurate channel estimations, quantization of CSI, or feedback errors. Therefore, the robustness of beamforming against imperfect CSIT should be investigated.

In the literature, there are generally two kinds of models to characterize imperfect CSIT: the stochastic model and the deterministic one. The stochastic model assumes that the channel is a random quantity, with its statistics, such as the mean and/or covariance, known by the transmitter. The optimality of beamforming, in terms of maximizing the average received signal-to-noise ratio (SNR) or mutual information, has been well studied [4]–[6]. The deterministic model, more suitable for characterizing instantaneous CSIT with errors, assumes that the actual channel lies in the neighborhood, called the uncertainty region, of a nominal channel known by the transmitter. Following the philosophy of worst-case robustness, a transmit strategy is called robust if it can achieve the best performance in the worst channel within the uncertainty region [7]–[15]. So far, the only result on the worst-case robustness of beamforming was restricted to a rank-one channel structure [12].

The goal of this paper is to study the robustness of beamforming from the perspective of worst-case robustness. Specifically, we consider a robust MIMO transmit strategy design by maximizing the worst-case received SNR, which leads to a maximin problem. The channel uncertainty region is modeled as an ellipsoid centered at the nominal channel and defined by the spectral norm.

Surprisingly, we prove that beamforming along the right singular vector associated with the maximum singular value of the nominal channel is the optimal solution to the maximin problem. This result indicates that beamforming is actually robust, in the sense that it provides the maximum received SNR in the worst channel within the spectral-norm-defined uncertainty region. Furthermore, we can reasonably infer that beamforming is approximately robust, for the most common matrix-norm-defined uncertainty regions, provided that channel dimensions or channel uncertainty are small.

Notation: The operators $(\cdot)^H$, $(\cdot)^{-1}$, $(\cdot)^\dagger$, and $\mathrm{Tr}(\cdot)$ represent the Hermitian, inverse, pseudo-inverse, and trace operations, respectively. $\|\cdot\|_2$ and $\|\cdot\|_F$ denote the spectral norm and the Frobenius norm of a matrix, respectively. The maximum eigenvalue of a Hermitian matrix is represented by $\lambda_{\max}(\cdot)$. The range space of $\mathbf{A}$ is denoted by $\mathcal{R}(\mathbf{A})$ and its null space by $\mathcal{N}(\mathbf{A})$.

II. Problem Statement

Consider a MIMO communication system with N transmit and M receive antennas. The system can be represented by

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

where $\mathbf{x} \in \mathbb{C}^N$ is the transmitted signal vector, $\mathbf{y} \in \mathbb{C}^M$ is the received signal vector, $\mathbf{H} \in \mathbb{C}^{M\times N}$ is the channel matrix, and the noise vector $\mathbf{n} \in \mathbb{C}^M$ follows $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \sigma_n^2\mathbf{I})$.

A MIMO transmit strategy, for a given distribution on the input symbols, is fully characterized by the transmit covariance matrix $\mathbf{Q} \triangleq \mathrm{E}\{\mathbf{x}\mathbf{x}^H\}$. Denote the eigenvalue decomposition (EVD) of $\mathbf{Q}$ by $\mathbf{Q} = \mathbf{U}_q\mathbf{\Lambda}_q\mathbf{U}_q^H$ with eigenvalues $\{p_i\}_{i=1}^N$. Then, the eigenvectors, i.e., the columns of $\mathbf{U}_q$, have the physical meaning of the transmit directions, and p_i corresponds to

J. Wang's work was supported by the ICT/TNG Swedish government strategic research area grant under the RICSNET project and the NSFC/RGC N_HKUST604/08 research grant. M. Payaró's work was partially supported by the Catalan Government under grant 2009 SGR 1046 and the Spanish Government under project TEC2008-06327-C03-03/TEC (FBMC-SILAN).

the power allocated to the ith eigen-mode. Beamforming is the simplest transmit strategy because it uses only one direction (i.e., $p_i > 0$ for some i and $p_j = 0$ for $j \neq i$). However, this simplicity induces worries on the sensitivity of beamforming to inaccuracies of CSIT. In practice, CSIT is usually imperfect due to many factors, which makes it necessary to investigate how robust beamforming is against imperfect CSIT.

To characterize imperfect CSIT, we follow a very common model [7]–[15] and assume that the actual channel $\mathbf{H}$ can be expressed as $\mathbf{H} = \hat{\mathbf{H}} + \mathbf{\Delta}$, where $\hat{\mathbf{H}}$ is the nominal channel known by the transmitter, and $\mathbf{\Delta}$ is the channel error lying in an elliptical uncertainty region defined by the spectral norm as [8], [13]

$$\mathcal{E} \triangleq \{\mathbf{\Delta} : \|\mathbf{\Delta}\|_2 \leq \varepsilon\} = \{\mathbf{\Delta} : \mathbf{\Delta}\mathbf{\Delta}^H \preceq \varepsilon^2 \mathbf{I}\} \tag{2}$$

with a given radius ε. According to the philosophy of worst-case robustness, a transmit strategy is defined to be robust if it can guarantee a performance level for any channel (error) realization in the uncertainty region, or, equivalently, can achieve the best performance in the worst channel.

Assuming perfect CSI at the receiver (CSIR), we take $\mathrm{Tr}(\mathbf{H}\mathbf{Q}\mathbf{H}^H)$ as the performance measure. It has been shown in [15] that maximizing $\mathrm{Tr}(\mathbf{H}\mathbf{Q}\mathbf{H}^H)$ corresponds to: 1) maximizing the received SNR; 2) minimizing the pairwise error probability (PEP) if a space-time block code (STBC) [16] is used; 3) maximizing the mutual information at low SNR; 4) minimizing the mean square error (MSE) at low SNR. Then, the robust transmit strategy is given by the solution to the following maximin problem:

$$\max_{\mathbf{Q}\in\mathcal{Q}} \min_{\mathbf{\Delta}\in\mathcal{E}} \ \mathrm{Tr}\{(\hat{\mathbf{H}} + \mathbf{\Delta})\mathbf{Q}(\hat{\mathbf{H}} + \mathbf{\Delta})^H\} \tag{3}$$

where

$$\mathcal{Q} \triangleq \{\mathbf{Q} : \mathbf{Q} \succeq 0, \ \mathrm{Tr}(\mathbf{Q}) \leq P\} \tag{4}$$

represents the power constraint at the transmitter.

The most important contribution of this paper is to show that the solution to (3) is just beamforming on the nominal channel, indicating that beamforming is actually a robust transmit strategy.

III. Optimal Transformation Directions

An important step towards the beamforming solution is to find the optimal transmit directions, i.e., the eigenvectors of the optimal transmit covariance matrix. An interesting phenomenon, for perfect CSIT [2], [3] and statistical CSIT with mean or covariance feedback [5], [6], is that the optimal transmit directions diagonalize the channel and result in an eigen-mode transmission. We show that this desirable property also holds for the worst-case design.

We first need to transform the maximin problem (3) into an equivalent, but more convenient form as follows.

Proposition 1: The maximin problem (3) is equivalent to

$$\begin{array}{ll} \underset{\mathbf{Q},\mathbf{W},\mathbf{Z}}{\text{minimize}} & \mathrm{Tr}\{(\mathbf{Z} - \mathbf{Q})\,\hat{\mathbf{H}}^H\hat{\mathbf{H}}\} + \varepsilon^2\mathrm{Tr}(\mathbf{W}) \\ \text{subject to} & \mathbf{Q} \in \mathcal{Q}, \ \mathbf{W} \succeq 0 \\ & \begin{bmatrix} \mathbf{Z} & \mathbf{Q} \\ \mathbf{Q} & \mathbf{Q} + \mathbf{W} \end{bmatrix} \succeq 0. \end{array} \tag{5}$$

Proof: See Appendix A. ■

Proposition 1, in fact, provides an efficient way to compute the optimal transmit covariance matrix through convex optimization [17]. Indeed, one can easily see that (5) is a semidefinite program (SDP), i.e., a very tractable form of convex optimization, meaning that (5) can be efficiently solved through numerical methods, e.g., the interior-point method. This observation becomes particularly important when other power constraints, e.g., the maximum power constraint or the per-antenna power constraint, are considered.

In the following, we use the equivalent form in Proposition 1 to find the optimal transmit directions. Denote the EVDs of $\mathbf{W}$, $\mathbf{Z}$, and $\hat{\mathbf{H}}\hat{\mathbf{H}}^H$ by $\mathbf{W} = \mathbf{U}_w\mathbf{\Lambda}_w\mathbf{U}_w^H$ with eigenvalues $\{w_i\}_{i=1}^N$, $\mathbf{Z} = \mathbf{U}_z\mathbf{\Lambda}_z\mathbf{U}_z^H$ with eigenvalues $\{z_i\}_{i=1}^N$, and $\hat{\mathbf{H}}\hat{\mathbf{H}}^H = \mathbf{U}_h\mathbf{\Lambda}_h\mathbf{U}_h^H$ with eigenvalues $\gamma_1 \geq \cdots \geq \gamma_N$ in decreasing order, respectively. Then, we have the following result.

Theorem 1: $\mathbf{U}_q = \mathbf{U}_h$ is optimal for the maximin problem (3), i.e., the optimal transmit directions are the right singular vectors of the nominal channel.

Proof: We shall use the following lemma.

Lemma 1 ([8]): Let $\mathbf{J} \in \mathbb{R}^{N\times N}$ be a diagonal matrix with the diagonal elements being ± 1. There are $L = 2^N$ different such matrices indexed from $l = 1$ to L. Let $\mathbf{A} \in \mathbb{C}^{N\times N}$ be an arbitrary matrix and $\mathbf{D}_\mathbf{A}$ be a diagonal matrix such that $[\mathbf{D}_\mathbf{A}]_{ii} = [\mathbf{A}]_{ii}, \forall i$. Then, $\mathbf{D}_\mathbf{A} = \frac{1}{L}\sum_{l=1}^{L} \mathbf{J}_l\mathbf{A}\mathbf{J}_l$.

By introducing $\hat{\mathbf{Q}} \triangleq \mathbf{U}_h^H\mathbf{Q}\mathbf{U}_h$, $\hat{\mathbf{W}} \triangleq \mathbf{U}_h^H\mathbf{W}\mathbf{U}_h$, and $\hat{\mathbf{Z}} \triangleq \mathbf{U}_h^H\mathbf{Z}\mathbf{U}_h$, (5) can be rewritten as

$$\begin{array}{ll} \underset{\hat{\mathbf{Q}},\hat{\mathbf{W}},\hat{\mathbf{Z}}}{\text{minimize}} & \mathrm{Tr}\{(\hat{\mathbf{Z}} - \hat{\mathbf{Q}})\mathbf{\Lambda}_h\} + \varepsilon^2\mathrm{Tr}(\hat{\mathbf{W}}) \\ \text{subject to} & \hat{\mathbf{Q}} \in \mathcal{Q}, \ \hat{\mathbf{W}} \succeq 0 \\ & \begin{bmatrix} \hat{\mathbf{Z}} & \hat{\mathbf{Q}} \\ \hat{\mathbf{Q}} & \hat{\mathbf{Q}} + \hat{\mathbf{W}} \end{bmatrix} \succeq 0. \end{array} \tag{6}$$

Notice that the objective in (6), denoted by $F(\hat{\mathbf{Q}}, \hat{\mathbf{W}}, \hat{\mathbf{Z}})$, depends only on the diagonal elements of the variables $\hat{\mathbf{Q}}$, $\hat{\mathbf{W}}$ and $\hat{\mathbf{Z}}$. Thus, we have $F(\hat{\mathbf{Q}}, \hat{\mathbf{W}}, \hat{\mathbf{Z}}) = F(\mathbf{D}_{\hat{\mathbf{Q}}}, \mathbf{D}_{\hat{\mathbf{W}}}, \mathbf{D}_{\hat{\mathbf{Z}}})$, where $\mathbf{D}_{\hat{\mathbf{Q}}}$, $\mathbf{D}_{\hat{\mathbf{W}}}$ and $\mathbf{D}_{\hat{\mathbf{Z}}}$ are diagonal matrices such that $[\mathbf{D}_{\hat{\mathbf{Q}}}]_{ii} = [\hat{\mathbf{Q}}]_{ii}$, $[\mathbf{D}_{\hat{\mathbf{W}}}]_{ii} = [\hat{\mathbf{W}}]_{ii}$, and $[\mathbf{D}_{\hat{\mathbf{Z}}}]_{ii} = [\hat{\mathbf{Z}}]_{ii}, \forall i$.

Next, we show that if the constraints in (6) are satisfied by $(\hat{\mathbf{Q}}, \hat{\mathbf{W}}, \hat{\mathbf{Z}})$, they are also satisfied by $(\mathbf{D}_{\hat{\mathbf{Q}}}, \mathbf{D}_{\hat{\mathbf{W}}}, \mathbf{D}_{\hat{\mathbf{Z}}})$. First, the linear matrix inequality (LMI) in (6) amounts to

$$\begin{bmatrix} \mathbf{J}_l & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_l \end{bmatrix} \begin{bmatrix} \hat{\mathbf{Z}} & \hat{\mathbf{Q}} \\ \hat{\mathbf{Q}} & \hat{\mathbf{Q}} + \hat{\mathbf{W}} \end{bmatrix} \begin{bmatrix} \mathbf{J}_l & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_l \end{bmatrix} = \begin{bmatrix} \mathbf{J}_l\hat{\mathbf{Z}}\mathbf{J}_l & \mathbf{J}_l\hat{\mathbf{Q}}\mathbf{J}_l \\ \mathbf{J}_l\hat{\mathbf{Q}}\mathbf{J}_l & \mathbf{J}_l\hat{\mathbf{Q}}\mathbf{J}_l + \mathbf{J}_l\hat{\mathbf{W}}\mathbf{J}_l \end{bmatrix} \succeq 0 \tag{7}$$

with $\mathbf{J}_l$ defined in Lemma 1. This implies that if $(\hat{\mathbf{Q}}, \hat{\mathbf{W}}, \hat{\mathbf{Z}})$ satisfies the LMI in (6), so does $(\mathbf{J}_l\hat{\mathbf{Q}}\mathbf{J}_l, \mathbf{J}_l\hat{\mathbf{W}}\mathbf{J}_l, \mathbf{J}_l\hat{\mathbf{Z}}\mathbf{J}_l)$. Then, according to Lemma 1, the convex combination $\frac{1}{L}\sum_{l=1}^{L}(\mathbf{J}_l\hat{\mathbf{Q}}\mathbf{J}_l, \mathbf{J}_l\hat{\mathbf{W}}\mathbf{J}_l, \mathbf{J}_l\hat{\mathbf{Z}}\mathbf{J}_l) = (\mathbf{D}_{\hat{\mathbf{Q}}}, \mathbf{D}_{\hat{\mathbf{W}}}, \mathbf{D}_{\hat{\mathbf{Z}}})$ is still in the convex feasible set defined by the LMI. Second, it is easy to see that if $\hat{\mathbf{Q}} \in \mathcal{Q}$ and $\hat{\mathbf{W}} \succeq 0$, then $\mathbf{D}_{\hat{\mathbf{Q}}} \in \mathcal{Q}$ and $\mathbf{D}_{\hat{\mathbf{W}}} \succeq 0$.

Consequently, given any feasible $(\hat{\mathbf{Q}}, \hat{\mathbf{W}}, \hat{\mathbf{Z}})$ in (6), $(\mathbf{D}_{\hat{\mathbf{Q}}}, \mathbf{D}_{\hat{\mathbf{W}}}, \mathbf{D}_{\hat{\mathbf{Z}}})$ is also feasible and results in the same objective value as $(\hat{\mathbf{Q}}, \hat{\mathbf{W}}, \hat{\mathbf{Z}})$. Therefore, in the solution set of (6), there must exist a diagonal structure, which can be achieved by letting $\mathbf{U}_q = \mathbf{U}_h$, $\mathbf{U}_w = \mathbf{U}_h$ and $\mathbf{U}_z = \mathbf{U}_h$, leading to $\mathbf{D}_{\hat{\mathbf{Q}}} = \mathbf{\Lambda}_q$, $\mathbf{D}_{\hat{\mathbf{W}}} = \mathbf{\Lambda}_w$ and $\mathbf{D}_{\hat{\mathbf{Z}}} = \mathbf{\Lambda}_z$. ■

With the optimal transmit directions given by the right singular vectors of the nominal channel, Theorem 1 indicates that the worst-case robust design leads to an eigen-mode transmission, just similar to the cases of perfect CSIT [2], [3] and statistical imperfect CSIT [5], [6]. Interestingly, the same optimal directions were also found for the uncertainty region defined by the Frobenius norm in [14], [15]. A direct consequence of Theorem 1 is that the matrix-valued problem (3) or (5) can now be simplified into a scalar power allocation problem without losing any optimality, thus helping us move closer to the beamforming solution.

IV. Optimality of Beamforming

We first use the result in Theorem 1 to simplify the matrix-valued problem (5) into the following power allocation problem.

Proposition 2: With $\mathbf{U}_q = \mathbf{U}_h$, the problem (5) reduces to

$$\underset{\mathbf{p}\in\mathcal{P},\mathbf{w}\geq 0}{\text{maximize}} \ \sum_{i=1}^{N} \frac{w_i \gamma_i p_i}{w_i + p_i} - \varepsilon^2 \sum_{i=1}^{N} w_i \tag{8}$$

where $\mathcal{P} \triangleq \{\mathbf{p} : \mathbf{p} \geq 0, \sum_{i=1}^{N} p_i \leq P\}$.

Proof: See Appendix B. ■

It can be verified that the objective in (8) is jointly concave in $(\mathbf{p}, \mathbf{w})$, implying that (8) is a convex problem (since $\mathcal{P}$ is a convex set). Now, we show that the optimal solution to the power allocation (8) is beamforming over the maximum eigen-mode.

Theorem 2: The solution to the problem (8) is $p_1^\star = P$ and $p_i^\star = 0$ for $i \geq 2$, and the optimum value of (8) is $P(\sqrt{\gamma_1} - \varepsilon)^2$.

Proof: Assume without loss of generality (w.l.o.g.) that $\{p_i\}_{i=1}^N$ are ordered decreasingly and $\text{rank}(\hat{\mathbf{H}}) = r$. For fixed $\mathbf{p}$, it is not difficult to find the optimal $\mathbf{w}$ as $w_i^\star = \frac{p_i}{\varepsilon}(\sqrt{\gamma_i} - \varepsilon)$ for $i \leq m$ and $w_i^\star = 0$ for $i > m$, where $m \in \{1, \cdots, r\}$ is an integer such that $\gamma_m > \varepsilon^2 \geq \gamma_{m+1}$ with $\gamma_{r+1} \triangleq 0$. Plugging $\mathbf{w}^\star$ back into (8), the objective function can be expressed as

$$\sum_{i=1}^{m} \frac{w_i^\star \gamma_i p_i}{w_i^\star + p_i} - \varepsilon^2 \sum_{i=1}^{m} w_i^\star = \sum_{i=1}^{m} \left(\sqrt{\gamma_i} - \varepsilon\right)^2 p_i. \tag{9}$$

Therefore, the power allocation problem reduces to the following simple linear program (LP):

$$\begin{array}{ll} \underset{\mathbf{p}}{\text{maximize}} & \sum_{i=1}^{m} \left(\sqrt{\gamma_i} - \varepsilon\right)^2 p_i \\ \text{subject to} & \sum_{i=1}^{N} p_i \leq P, \ p_1 \geq \cdots \geq p_N \geq 0 \end{array} \tag{10}$$

where we have explicitly taken into account the decreasing order of $\{p_i\}_{i=1}^N$. Apparently, the optimal solution to (10) is to put all available power P on the maximum term $(\sqrt{\gamma_1} - \varepsilon)$, which completes the proof. ■

Theorem 2 (together with Theorem 1) indicates that beamforming along the right singular vector associated with the maximum singular value of the nominal channel results in the maximum received SNR in the worst channel within the elliptical uncertainty region defined by the spectral norm. This means that beamforming is a robust solution for the spectral-norm-defined uncertainty region.

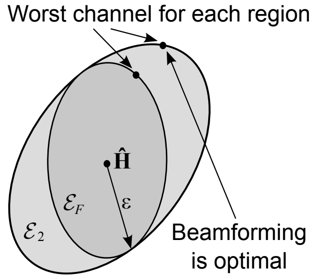

Fig. 1. General situation: given the same error radius, the Frobenius-norm-defined uncertainty region $\mathcal{E}_F$ is contained by the spectral-norm-defined uncertainty region $\mathcal{E}_2$. Observe that the worst channel lies on the boundary of the region in both cases.

Given that the maximin problem with the Frobenius-norm-defined uncertainty region also admits a closed-form solution [14], [15], we summarize the robustness results of beamforming for the spectral-norm-defined and Frobenius-norm-defined uncertainty regions in the following corollary.

Corollary 1: Beamforming is the solution to the problem of maximizing the worst-case received SNR if

1) either $\mathbf{\Delta} \in \mathcal{E}_2 \triangleq \{\mathbf{\Delta} : \mathbf{\Delta}\mathbf{\Delta}^H \preceq \varepsilon^2 \mathbf{I}\}$ for any ε;
2) or $\mathbf{\Delta} \in \mathcal{E}_F \triangleq \{\mathbf{\Delta} : \text{Tr}(\mathbf{\Delta}\mathbf{\Delta}^H) \leq \varepsilon^2\}$ and $\varepsilon \leq (\sqrt{\gamma_1} - \sqrt{\gamma_2})$ [15, Corollary 1].

In general, the optimality of beamforming is determined by the shape of the uncertainty region (e.g., $\mathcal{E}_2$ and $\mathcal{E}_F$) and its parameters such as the center $\hat{\mathbf{H}}$ and the radius ε. Interestingly, within the uncertainty region $\mathcal{E}_2$ defined by the spectral norm, beamforming is always the robust solution, regardless of the parameters $\hat{\mathbf{H}}$ and ε.

One may notice that the spectral norm is the smallest one among the most common matrix norms [18] (e.g., given $\mathbf{A} \in \mathbb{C}^{M\times N}$, we have that $\|\mathbf{A}\|_2 \leq \|\mathbf{A}\|_1$, $\|\mathbf{A}\|_2 \leq \|\mathbf{A}\|_{\max}$, and $\|\mathbf{A}\|_2 \leq \|\mathbf{A}\|_F$, where $\|\mathbf{A}\|_1 = \sum_{i,j} |[\mathbf{A}]_{ij}|$, $\|\mathbf{A}\|_{\max} = \sqrt{NM} \max_{i,j} |[\mathbf{A}]_{ij}|$, and $\|\mathbf{A}\|_F$ is the aforementioned Frobenius norm). Consequently, given the same radius ε, $\mathcal{E}_2$ is the biggest one among all ellipsoids defined by these matrix norms, e.g., $\mathcal{E}_F \subseteq \mathcal{E}_2$ (see Fig. 1), which implies that $\mathcal{E}_2$ is the most conservative uncertainty region based on these matrix norms.

Strictly speaking, beamforming is not the exact robust solution for other uncertainty regions than $\mathcal{E}_2$. Nevertheless, the gap between the spectral norm and any other matrix norm decreases as the matrix's dimension decreases. For example, when $M = 1$ or $N = 1$, i.e., a rank-one channel, we have $\mathcal{E}_2 = \mathcal{E}_F$. Meanwhile, if the radius ε is small enough, using $\mathcal{E}_2$ to cover the set defined by other norms will only add a small amount of uncertainty (see Fig. 2). Therefore, loosely speaking, one can state that beamforming is an approximately robust solution, for the most common matrix-norm-defined uncertainty regions, when the channel dimension is small or

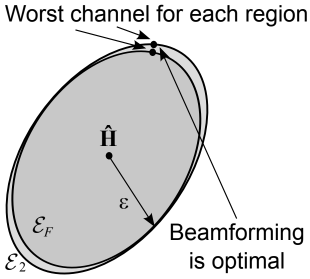

Fig. 2. Small uncertainty radius or small channel dimensions: the Frobenius-norm-defined uncertainty region $\mathcal{E}_F$ is still contained by the spectral-norm-defined uncertainty region $\mathcal{E}_2$, but the gap between the two regions is very small.

the channel uncertainty is small.

V. Numerical Results

In this section, we evaluate the robustness of beamforming through numerical examples. To be more exact, we compare beamforming, which is robust for $\mathcal{E}_2$, with the equal-power transmission and the robust strategy for $\mathcal{E}_F$ [14], [15] through their worst-case performance in the Frobenius-norm-defined uncertainty region $\mathcal{E}_F$ (see Corollary 1). The radius of $\mathcal{E}_F$ is set to be $\varepsilon = s||\hat{\mathbf{H}}||_2 = s\sqrt{\gamma_1}$ with $s \in [0, 1)$. Although beamforming is not the exact robust solution for $\mathcal{E}_F$, it will be shown later that the performance of beamforming is quite close to that of the robust strategy for $\mathcal{E}_F$ when then channel dimension or channel uncertainty is small. The elements of the nominal channel $\hat{\mathbf{H}}$ are randomly generated according to zero-mean, unit-variance, i.i.d. Gaussian distributions.

Figs. 3 and 4 depict the average worst-case symbol error rate (SER) versus SNR for different sizes of the uncertainty region (i.e., different values of s), where the average is taken over the nominal channel $\hat{\mathbf{H}}$. In Fig. 3, we use 2 antennas at both ends of the MIMO link ($M = N = 2$), while we use 2 transmit antennas ($N = 2$) and 4 receive antennas ($M = 4$) in Fig. 4. In both figures, the full-rate complex STBC [16] and QPSK modulation are adopted.

One can observe from Figs. 3 and 4 that, the robust strategy for $\mathcal{E}_F$ achieves the best worst-case performance, since we have used $\mathcal{E}_F$ in the simulation. However, the interesting part is that, for a small or even moderate amount of uncertainty, the performance of beamforming is very close to that of the robust strategy for $\mathcal{E}_F$, and better than that of the equal-power transmission, even though beamforming is robust for $\mathcal{E}_2$ but not $\mathcal{E}_F$. Therefore, the numerical results verify our inference that beamforming is approximately robust for small channel uncertainty or small numbers of antennas, rather independent of the shape of the uncertainty region.

VI. Conclusion

We have investigated the ability of beamforming to combat against the imperfectness of CSIT in a MIMO channel from the perspective of worst-case robustness. We have proved that beamforming is robust in the sense of providing the maximum received SNR in the worst channel within the spectral-norm-defined uncertainty region. It was then reasonably inferred that beamforming is a robust transmit strategy for small channel dimensions or channel uncertainty independent of the shape of the uncertainty region. Finally, our results were verified by the numerical examples.

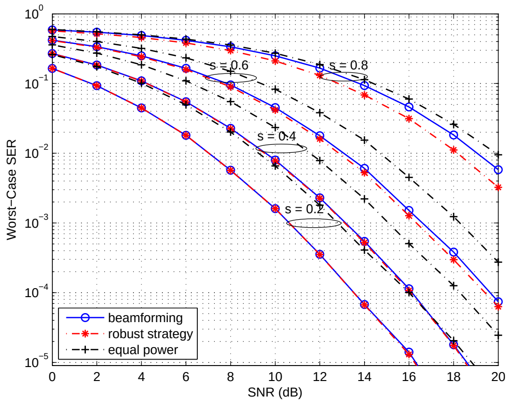

Fig. 3. Worst-case SER versus SNR for N = M= 2 and different amounts of uncertainty.

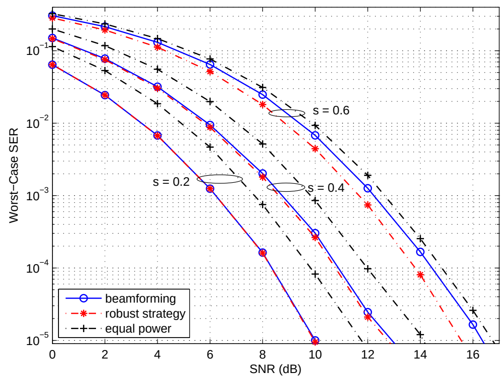

Fig. 4. Worst-case SER versus SNR for N = 2, M = 4, and different amounts of uncertainty.

Appendix A
Proof of Proposition 1

We will use the following results in the proof.

Lemma 2 ([19]): Let $\mathbf{A} \succeq \mathbf{B} \succeq 0$. Then, $\mathcal{R}(\mathbf{A}) \supseteq \mathcal{R}(\mathbf{B})$ and $\mathcal{N}(\mathbf{A}) \subseteq \mathcal{N}(\mathbf{B})$.

Lemma 3 (General Schur's Complement [17]): Let

$$\mathbf{X} = \begin{bmatrix} \mathbf{A} & \mathbf{B}^H \\ \mathbf{B} & \mathbf{C} \end{bmatrix}$$

be a Hermitian matrix. Then, $\mathbf{X} \succeq 0$ if and only if $\mathbf{C} \succeq 0$, $\mathbf{A} - \mathbf{B}^H\mathbf{C}^\dagger\mathbf{B} \succeq 0$ and $\mathbf{B}\mathbf{v} \in \mathcal{R}(\mathbf{C})$, $\forall \mathbf{v}$.

First, we note that by exploiting the Lagrangian duality on the inner minimization in the maximin problem (3), one can transform (3) into the following equivalent problem:

$$\underset{\mathbf{Q}\in\mathcal{Q},\mathbf{W}\succeq 0}{\text{minimize}}\ \text{Tr}\{\mathbf{Q}(\mathbf{Q}+\mathbf{W})^{\dagger}\mathbf{Q}\mathbf{H}^H\mathbf{H}\}-\text{Tr}(\mathbf{Q}\mathbf{H}^H\mathbf{H})+\varepsilon^2\text{Tr}(\mathbf{W}). \tag{11}$$

Then, it is easy to see that (11) is equivalent to

$$\begin{array}{ll}\underset{\mathbf{Q}\in\mathcal{Q},\mathbf{W}\succeq 0,\mathbf{Z}}{\text{minimize}} & \text{Tr}(\mathbf{Z}\mathbf{H}^H\mathbf{H})-\text{Tr}(\mathbf{Q}\mathbf{H}^H\mathbf{H})+\varepsilon^2\text{Tr}(\mathbf{W})\\ \text{subject to} & \mathbf{Q}(\mathbf{Q}+\mathbf{W})^{\dagger}\mathbf{Q}\preceq\mathbf{Z}.\end{array} \tag{12}$$

Since $\mathbf{Q}+\mathbf{W}\succeq\mathbf{Q}$, from Lemma 2, it follows that $\mathcal{R}(\mathbf{Q})\subseteq\mathcal{R}(\mathbf{Q}+\mathbf{W})$, or in other words $\mathbf{Q}\mathbf{v}\in\mathcal{R}(\mathbf{Q}+\mathbf{W})$, $\forall\mathbf{v}$. Therefore, by using Lemma 3, we can equivalently transform (12) into (5) in Proposition 1.

Appendix B
Proof of Proposition 2

Using $\mathbf{U}_q=\mathbf{U}_h$, $\mathbf{U}_w=\mathbf{U}_h$ and $\mathbf{U}_z=\mathbf{U}_h$, the problem (5) becomes

$$\begin{array}{ll}\underset{\mathbf{\Lambda}_q,\mathbf{\Lambda}_w,\mathbf{\Lambda}_z}{\text{minimize}} & \text{Tr}\{(\mathbf{\Lambda}_z-\mathbf{\Lambda}_q)\,\mathbf{\Lambda}_h\}+\varepsilon^2\text{Tr}(\mathbf{\Lambda}_w\mathbf{\Lambda}_t^{-1})\\ \text{subject to} & \mathbf{\Lambda}_q\in\mathcal{Q},\ \mathbf{\Lambda}_w\succeq 0\\ & \begin{bmatrix}\mathbf{\Lambda}_z & \mathbf{\Lambda}_q\\ \mathbf{\Lambda}_q & \mathbf{\Lambda}_q+\mathbf{\Lambda}_w\end{bmatrix}\succeq 0.\end{array} \tag{13}$$

The LMI in (13), via proper symmetric column and row permutations, can be rewritten as

$$\begin{bmatrix} z_i & p_i\\ p_i & p_i+w_i\end{bmatrix}\succeq 0,\ \forall i \tag{14}$$

which represent the same constraints as

$$z_i\,(p_i+w_i)\geq p_i^2,\ p_i+w_i\geq 0,\ z_i\geq 0,\ \forall i. \tag{15}$$

Hence, (13) is equivalent to

$$\begin{array}{ll}\underset{\mathbf{p},\mathbf{w},\mathbf{z}}{\text{minimize}} & \sum_{i=1}^N\gamma_i\,(z_i-p_i)+\varepsilon^2\sum_{i=1}^N\frac{w_i}{\tau_i}\\ \text{subject to} & \mathbf{p}\in\mathcal{P},\ \mathbf{w}\geq 0,\ \mathbf{z}\geq 0\\ & z_i\,(w_i+p_i)\geq p_i^2,\ \forall i.\end{array} \tag{16}$$

Assume w.l.o.g. that $w_i>0$ for $i\in\mathcal{I}\subseteq\{1,\cdots,N\}$, and $w_i=0$ for $i\notin\mathcal{I}$. First, the constraint $z_i\,(w_i+p_i)\geq p_i^2$ reduces to $z_i\geq p_i$ for $i\notin\mathcal{I}$, so (16) amounts to

$$\begin{array}{ll}\underset{\mathbf{p},\mathbf{w},\mathbf{z}}{\text{minimize}} & \sum_{i\in\mathcal{I}}\gamma_i\,(z_i-p_i)+\varepsilon^2\sum_{i\in\mathcal{I}}\frac{w_i}{\tau_i}\\ \text{subject to} & \mathbf{p}\in\mathcal{P},\ \mathbf{w}\geq 0,\ \mathbf{z}\geq 0\\ & z_i\,(w_i+p_i)\geq p_i^2,\ i\in\mathcal{I}.\end{array} \tag{17}$$

Second, the constraint $z_i\,(w_i+p_i)\geq p_i^2$ is equal to $p_i^2/(w_i+p_i)\leq z_i$ for $i\in\mathcal{I}$. Hence, (17) is equivalent to

$$\underset{\mathbf{p}\in\mathcal{P},\mathbf{w}\geq 0}{\text{minimize}}\ \sum_{i\in\mathcal{I}}\frac{\gamma_i p_i^2}{w_i+p_i}-\sum_{i\in\mathcal{I}}\gamma_i p_i+\varepsilon^2\sum_{i\in\mathcal{I}}\frac{w_i}{\tau_i} \tag{18}$$

whose objective amounts to

$$-\sum_{i\in\mathcal{I}}\frac{w_i\gamma_i p_i}{w_i+p_i}+\varepsilon^2\sum_{i\in\mathcal{I}}\frac{w_i}{\tau_i}=-\sum_{i=1}^N\frac{w_i\gamma_i p_i}{w_i+p_i}+\varepsilon^2\sum_{i=1}^N\frac{w_i}{\tau_i} \tag{19}$$

where the second equality is due to $w_i=0$ for $i\notin\mathcal{I}$. Finally, (8) is just the maximizing counterpart of (18).

Acknowledgment

The authors would like to thank Prof. Daniel P. Palomar for helpful discussions.

References

[1] A. Scaglione, P. Stoica, S. Barbarossa, G. B. Giannakis, and H. Sampath, "Optimal designs for space-time linear precoders and decoders," *IEEE Trans. Signal Processing*, vol. 50, no. 5, pp. 1051–1064, May 2002.

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

[5] 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.

[6] E. A. Jorswieck and H. Boche, "Optimal transmission strategies and impact of correlation in multiantenna systems with different types of channel state information," *IEEE Trans. Signal Processing*, vol. 52, no. 12, pp. 3440–3453, Dec. 2004.

[7] S. A. Vorobyov, A. B. Gershman, and Z.-Q. Luo, "Robust adaptive beamforming using worst-case performance optimization: A solution to the signal mismatch problem," *IEEE Trans. Signal Processing*, vol. 51, no. 2, pp. 313–324, Feb. 2003.

[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] 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.

[10] 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.

[11] M. Payaró, A. Pascual-Iserte, and M. A. Lagunas, "Robust power allocation designs for multiuser and multiantenna downlink communication systems through convex optimization," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 7, pp. 1390–1401, Sep. 2007.

[12] A. Wiesel, Y. C. Eldar, and S. Shamai, "Optimization of the MIMO compound capacity," *IEEE Trans. Wireless Commun.*, vol. 6, no. 3, pp. 1094–1101, Mar. 2007.

[13] N. Vucic and H. Boche, "Robust QoS-constrained optimization of downlink multiuser MISO systems," *IEEE Trans. Signal Processing*, vol. 57, no. 2, pp. 714–725, Feb. 2009.

[14] J. Wang and D. P. Palomar, "Maximin robust design for MIMO communication systems against imperfect CSIT," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Taipei, Taiwan, Apr. 2009.

[15] ——, "Worst-case robust MIMO transmission with imperfect channel knowledge," *IEEE Trans. Signal Processing*, vol. 57, no. 8, pp. 3086–3100, Aug. 2009.

[16] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inform. Theory*, vol. 45, no. 5, pp. 1456–1467, Jul. 1999.

[17] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge University Press, 2004.

[18] R. A. Horn and C. R. Johnson, *Matrix Analysis*. New York: Cambridge University Press, 1985.

[19] D. S. Bernstein, *Matrix Mathematics: Theory, Facts, and Formulas With Application to Linear Systems Theory*. Princeton, NJ: Princeton Univ. Press, 2005.