On the Robustness of Transmit Beamforming

Jiaheng Wang and Miquel Payaró

Abstract

Beamforming is a simple transmit strategy that uses only one eigen-direction in multiple-input multiple-output channels. This simplicity makes beamforming a competitive strategy in practice, but at the same time poses a doubt on the sensitivity of beamforming to the imperfectness of the channel state information at the transmitter (CSIT). This paper studies beamforming from the perspective of worst-case robustness. We show that beamforming can achieve the maximum received signal-to-noise ratio (SNR), or guarantees a given received SNR with the minimum transmit power, in the worst channel within an elliptical uncertainty region defined by the weighted spectral norm. This result further implies that beamforming has the ability to combat against the imperfectness of CSIT, especially for small channel dimensions or small channel uncertainty.

I. Introduction

The full benefits of a multiple-input multiple-out (MIMO) system are achieved by employing the channel state information (CSI) at both ends of a link [1], [2]. Given perfect CSI at the transmitter (CSIT), the optimal transmit strategies have been found under various criteria [3]–[5]. The simplest one of them may be beamforming, which is defined as the transmit strategy using only one direction, i.e., one eigenvector of the transmit covariance matrix. This simplicity makes beamforming an attractive strategy in practice, but on the other hand may spark worries on its sensitivity to inaccuracies of CSIT [4], [5], which are, however, very common due to, e.g., inaccurate channel estimations, feedback errors, and quantization of CSI. The goal of this paper is to investigate how robust beamforming is against imperfect CSIT.

In the literature, imperfect CSIT can be modeled by either the stochastic or the deterministic approaches. The stochastic model assumes that the channel is a random quantity, and its statistics, such as the mean and/or the covariance, are 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 [6]–[9]. The deterministic model, which is more suitable for characterizing instantaneous CSIT with errors, assumes that the actual channel lies in the neighborhood, called the channel uncertainty region, of a nominal channel known by the transmitter. In this

J. Wang is with the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Hong Kong (e-mail: jhwangee@hotmail.com), whose work was supported by the NSFC/RGC N_HKUST604/08 research grant.

Miquel Payaró is with the Centre Tecnològic de Telecomunicacions de Catalunya, Parc Mediterrani de la Tecnologia, Castelldefels, Barcelona, Spain (e-mail: miquel.payaro@cttc.es), whose 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).

case, the robustness of a transmit strategy is determined by whether it can achieve the best performance in the worst channel within the uncertainty region, hence leading to worst-case robust designs [10]–[16]. However, to the best of our knowledge, the only result concerning the worst-case robustness of beamforming was restricted to a rank-one channel structure [14].

In this paper, we study the robustness of beamforming[1] from the worst-case robust perspective. Specifically, it is assumed that the channel uncertainty region is an ellipsoid with a given radius defined by the weighted spectral norm and centered at a known nominal channel. Then, we consider maximizing the worst-case received SNR, which leads to a maximin problem, as well as the quality of service (QoS) problem that minimizes the transmit power while keeping the received SNR above a given threshold for any channel realization in the uncertainty region. Note that similar problems were considered in the previous work [15], but with the uncertainty region defined by the weighted Frobenius norm. It turns out that the spectral-norm case is more challenging than the Frobenius-norm case (thus calling for more delicate mathematical tools) and, most importantly, leads to a quite different result.

Our main result shows that beamforming along the right singular vector associated with the maximum singular value of the nominal channel is the optimal solution to both the maximin and QoS problems under some mild conditions. Therefore, beamforming is robust in the sense that it provides the maximum received SNR, or guarantees a given received SNR with the minimum transmit power, in the worst channel within the spectral-norm-defined uncertainty region. With this result, we can further infer that beamforming is approximately robust, for the most common matrix norms defining the uncertainty region, provided that channel dimensions or channel uncertainty are small. Moreover, as a side result from the intermediate derivations, we also provide an efficient method to solve the maximin and QoS problems with a general power constraint through convex optimization [18].

Notation: $\mathbb{S}_+^n$ denotes the ensemble of all $n \times n$ positive semidefinite matrices. The operators $(\cdot)^H$, $(\cdot)^{-1}$, $(\cdot)^\dagger$, and $\mathrm{Tr}(\cdot)$ represent the Hermitian, inverse, pseudo-inverse, and trace operations, respectively. $\|\cdot\|$, $\|\cdot\|_2$, and $\|\cdot\|_F$ denote a general norm, 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 $\mathcal{N}(\mathbf{A})$. Finally, $\mathbf{d}(\mathbf{A})$ and $\boldsymbol{\lambda}(\mathbf{A})$ represent the vectors containing the diagonal and the eigenvalues of $\mathbf{A}$, respectively.

II. Signal Model and Problem Formulation

Consider a MIMO communication system with N transmit and M receive antennas. Let $\mathbf{x} \in \mathbb{C}^N$ and $\mathbf{y} \in \mathbb{C}^M$ be the transmitted and received signal vectors, respectively, Then, the system can be represented by

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

where $\mathbf{H} \in \mathbb{C}^{M \times N}$ is the channel matrix, and the noise vector $\mathbf{n} \in \mathbb{C}^M$ follows $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \sigma_n^2 \mathbf{I})$. Let $\mathbf{Q} = \mathrm{E}\{\mathbf{x}\mathbf{x}^H\}$ be the transmit covariance matrix and denote its eigenvalue decomposition (EVD) by $\mathbf{Q} = \mathbf{U}_q \boldsymbol{\Lambda}_q \mathbf{U}_q^H$ with eigenvalues $\{p_i\}_{i=1}^N$. A transmit strategy is fully characterized by $\mathbf{Q}$ in the sense that the eigenvectors, i.e., the

[1]The term *beamforming* is often used in the literature to denote generic array signal processing at the receiver. Readers are referred to [10], [17] for robust beamforming at the receiver. In this paper, we use beamforming to denote a simple transmit strategy in MIMO channels.

columns of $\mathbf{U}_q$, represent the transmit directions, and p_i corresponds to the power allocated to the ith eigen-mode. Beamforming is the simplest transmit strategy that uses only one direction (i.e., $p_i > 0$ for some i and $p_j = 0$ for $j \neq i$). As pointed out in the introduction, this simplicity induces worries on the sensitivity of beamforming to inaccuracies of CSIT. However, 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 [10]–[16] 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 $\mathcal{E} = \{\mathbf{\Delta} : \|\mathbf{\Delta}\| \leq \varepsilon\}$ with a given radius ε. In this paper, we consider $\mathcal{E}$ defined by the weighted spectral norm [11], [16] as

$$\mathcal{E} = \{\mathbf{\Delta} : \|\mathbf{\Delta}\|_2^{\mathbf{T}} \leq \varepsilon\} = \{\mathbf{\Delta} : \lambda_{\max}^{1/2}(\mathbf{\Delta}\mathbf{T}\mathbf{\Delta}^H) \leq \varepsilon\} = \{\mathbf{\Delta} : \mathbf{\Delta}\mathbf{T}\mathbf{\Delta}^H \preceq \varepsilon^2\mathbf{I}\} \tag{2}$$

where $\mathbf{T}$ is a known positive definite matrix. The general ellipsoid particularizes to a sphere when $\mathbf{T} = \mathbf{I}$, which is the most frequently used model [11], [16]. In this case, a transmit strategy is defined to be robust if it can provide a guaranteed performance for any channel (error) realization in the uncertainty region, or, equivalently, can achieve the best performance in the worst channel.

We assume perfect CSI at the receiver (CSIR), and take $\text{Tr}(\mathbf{H}\mathbf{Q}\mathbf{H}^H)$ as the performance measure, which is proportional to the received SNR. It can be verified that maximizing $\text{Tr}(\mathbf{H}\mathbf{Q}\mathbf{H}^H)$ corresponds to (see details in [15, Section II-B]): 1) maximizing the received SNR; 2) minimizing the pairwise error probability (PEP) if a space-time block code (STBC) [19] is used; 3) maximizing the mutual information at low SNR; 4) minimizing the MSE at low SNR if a minimum-mean-square-error (MMSE) equalizer is used at the receiver.

With this objective function, a robust transmit strategy is the solution to the following maximin problem:

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

where $\mathcal{Q}$ is the set of power constraints on $\mathbf{Q}$. Common power constraints include: 1) sum power constraint $\mathcal{Q}_1 \triangleq \{\mathbf{Q} : \mathbf{Q} \succeq 0,\ \text{Tr}(\mathbf{Q}) \leq P_T\}$; 2) maximum power constraint $\mathcal{Q}_2 \triangleq \{\mathbf{Q} : \mathbf{Q} \succeq 0,\ \lambda_{\max}(\mathbf{Q}) \leq P_T\}$; 3) per-antenna power constraint $\mathcal{Q}_3 \triangleq \{\mathbf{Q} : \mathbf{Q} \succeq 0,\ \max_i [\mathbf{Q}]_{ii} \leq P_T\}$ or $\mathcal{Q}_4 \triangleq \{\mathbf{Q} : \mathbf{Q} \succeq 0,\ [\mathbf{Q}]_{ii} \leq P_i,\ \forall i\}$. The most important contribution of this paper is to show that, under some mild conditions on $\mathcal{Q}$ and $\mathcal{E}$ that are formally specified in Section IV, the solution to (3) is just beamforming on the nominal channel, indicating that beamforming is actually a robust transmit strategy.

III. Problem Transformation and Diagonalization

To reach a beamforming solution, we need two intermediate steps. First, the maximin problem (3) will be transformed into an equivalent form that is easier to handle. It turns out that the equivalent problem provides an efficient way to solve (3) under a general power constraint. Second, we will prove that the optimal transmit directions, i.e., the eigenvectors of the optimal transmit covariance matrix, are the right singular vectors of the nominal channel under some mild conditions. This result leads to a channel-diagonalizing structure, thus reducing (3) to a power allocation problem.

A. Problem Transformation

Theorem 1: Let $\mathcal{Q} \subseteq \mathbb{S}_+^N$ be a nonempty compact set. Then, the maximin problem (3) is equivalent to

$$\begin{array}{ll} \underset{\mathbf{Q},\mathbf{W},\mathbf{Z}}{\text{minimize}} & \text{Tr}\{(\mathbf{Z}-\mathbf{Q})\,\hat{\mathbf{H}}^H\hat{\mathbf{H}}\} + \varepsilon^2\text{Tr}(\mathbf{W}\mathbf{T}^{-1}) \\ \text{subject to} & \mathbf{Q}\in\mathcal{Q},\ \mathbf{W}\succeq 0 \\ & \begin{bmatrix} \mathbf{Z} & \mathbf{Q} \\ \mathbf{Q} & \mathbf{Q}+\mathbf{W} \end{bmatrix} \succeq 0. \end{array} \tag{4}$$

Proof: The following results will be used.

Lemma 1 ([20]): 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 2 (General Schur's Complement [18]): 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 $\mathbf{B}\mathbf{v} \in \mathcal{R}(\mathbf{C})$, $\forall \mathbf{v}$. □

Proposition 1: Let $\mathcal{Q} \subseteq \mathbb{S}_+^N$ be a nonempty compact set. Then, the maximin problem (3) is equivalent to

$$\underset{\mathbf{Q}\in\mathcal{Q},\tilde{\mathbf{W}}\succeq 0}{\text{minimize}}\ \text{Tr}\{\tilde{\mathbf{Q}}(\tilde{\mathbf{Q}}+\mathbf{W})^\dagger\tilde{\mathbf{Q}}\tilde{\mathbf{H}}^H\tilde{\mathbf{H}}\} - \text{Tr}(\tilde{\mathbf{Q}}\tilde{\mathbf{H}}^H\tilde{\mathbf{H}}) + \varepsilon^2\text{Tr}(\tilde{\mathbf{W}}) \tag{5}$$

where $\tilde{\mathbf{Q}} \triangleq \mathbf{T}^{-1/2}\mathbf{Q}\mathbf{T}^{-1/2}$ and $\tilde{\mathbf{H}} \triangleq \hat{\mathbf{H}}\mathbf{T}^{1/2}$. □

Applying Proposition 1, which is proved in Appendix A, the maximin problem (3) can be equivalently transformed into (5). Then, it is easy to see that (5) is equivalent to

$$\begin{array}{ll} \underset{\mathbf{Q}\in\mathcal{Q},\tilde{\mathbf{W}}\succeq 0,\tilde{\mathbf{Z}}}{\text{minimize}} & \text{Tr}(\tilde{\mathbf{Z}}\tilde{\mathbf{H}}^H\tilde{\mathbf{H}}) - \text{Tr}(\tilde{\mathbf{Q}}\tilde{\mathbf{H}}^H\tilde{\mathbf{H}}) + \varepsilon^2\text{Tr}(\tilde{\mathbf{W}}) \\ \text{subject to} & \tilde{\mathbf{Q}}(\tilde{\mathbf{Q}}+\tilde{\mathbf{W}})^\dagger\tilde{\mathbf{Q}} \preceq \tilde{\mathbf{Z}}. \end{array} \tag{6}$$

Since $\tilde{\mathbf{Q}} + \mathbf{W} \succeq \tilde{\mathbf{Q}}$, from Lemma 1, it follows that $\mathcal{R}(\tilde{\mathbf{Q}}) \subseteq \mathcal{R}(\tilde{\mathbf{Q}} + \mathbf{W})$, or in other words $\tilde{\mathbf{Q}}\mathbf{v} \in \mathcal{R}(\tilde{\mathbf{Q}} + \mathbf{W})$, $\forall \mathbf{v}$. Therefore, by using Lemma 2, we can equivalently transform (6) into

$$\begin{array}{ll} \underset{\mathbf{Q}\in\mathcal{Q},\tilde{\mathbf{W}}\succeq 0,\tilde{\mathbf{Z}}}{\text{minimize}} & \text{Tr}\{(\tilde{\mathbf{Z}}-\tilde{\mathbf{Q}})\tilde{\mathbf{H}}^H\tilde{\mathbf{H}}\} + \varepsilon^2\text{Tr}(\tilde{\mathbf{W}}) \\ \text{subject to} & \begin{bmatrix} \tilde{\mathbf{Z}} & \tilde{\mathbf{Q}} \\ \tilde{\mathbf{Q}} & \tilde{\mathbf{Q}}+\tilde{\mathbf{W}} \end{bmatrix} \succeq 0. \end{array} \tag{7}$$

Then, it is not difficult to reformulate (7) to (4) with $\mathbf{Z} = \mathbf{T}^{1/2}\tilde{\mathbf{Z}}\mathbf{T}^{1/2}$ and $\mathbf{W} = \mathbf{T}^{1/2}\tilde{\mathbf{W}}\mathbf{T}^{1/2}$. ■

Theorem 1 indicates that, as long as the power constraint set is nonempty and compact, which is very general, the maximin problem (3) is equivalent to the minimization problem (4). If in addition $\mathcal{Q}$ is a convex set, then (4) is a convex problem, and thus can be efficiently solved by numerical methods, e.g., the interior-point method [18]. It is not difficult to see that the commonly used power constraint sets $\mathcal{Q}_1$, $\mathcal{Q}_2$, $\mathcal{Q}_3$, $\mathcal{Q}_4$, as well as any intersection of them, fall into the class of nonempty compact convex sets. Moreover, for any intersection of $\mathcal{Q}_1$, $\mathcal{Q}_2$, $\mathcal{Q}_3$, $\mathcal{Q}_4$, the problem (4) is or can be rewritten[2] as a semidefinite program (SDP) [18], which consists of linear matrix inequality (LMI) constraints and an affine objective function.

[2]The constraints in $\mathcal{Q}_1$, $\mathcal{Q}_3$, $\mathcal{Q}_4$ are already LMIs and, for $\mathcal{Q}_2$, $\lambda_{\max}(\mathbf{Q}) \leq P_T$ amounts to $\mathbf{Q} \preceq P_T\mathbf{I}$.

B. Problem Diagonalization

In the cases of perfect CSIT [4], [5] and statistical CSIT with mean or covariance feedback [7], [8], an interesting phenomenon is that the channel is diagonalized by the optimal transmit covariance matrix, resulting in an eigen-mode transmission. We show that this desirable property also holds for our worst-case design, under some conditions on the uncertainty region and the power constraints in addition to that in Theorem 1. Denote the EVDs of $\mathbf{W}$, $\mathbf{Z}$, $\mathbf{T}$ 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 $\{z_i\}_{i=1}^N$, $\mathbf{T} = \mathbf{U}_t \mathbf{\Lambda}_t \mathbf{U}_t^H$ with eigenvalues $\tau_1 \geq \cdots \geq \tau_N$ in decreasing order, 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 2: Let $\mathbf{U}_t = \mathbf{U}_h$, and $\mathcal{Q} = \{\mathbf{Q} : \mathbf{Q} \succeq 0, f_n(\boldsymbol{\lambda}(\mathbf{Q})) \leq P_n, \forall n\}$, where each $f_n(\mathbf{x})$ is a Schur-convex function. Then, $\mathbf{U}_q = \mathbf{U}_h$ is optimal for the maximin problem (3).

Proof: The following lemma will be used.

Lemma 3 ([11]): 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, $\mathbf{D}_{\mathbf{A}} = \frac{1}{L} \sum_{l=1}^L \mathbf{J}_l \mathbf{A} \mathbf{J}_l$. □

Using $\mathbf{U}_t = \mathbf{U}_h$ and 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 $\hat{\mathbf{Z}} \triangleq \mathbf{U}_h^H \mathbf{Z} \mathbf{U}_h$, (4) can be rewritten as

$$
\begin{array}{ll}
\underset{\hat{\mathbf{Q}}, \hat{\mathbf{W}}, \hat{\mathbf{Z}}}{\text{minimize}} & \text{Tr}\{(\hat{\mathbf{Z}} - \hat{\mathbf{Q}})\mathbf{\Lambda}_h\} + \varepsilon^2 \text{Tr}(\hat{\mathbf{W}} \mathbf{\Lambda}_h^{-1}) \\
\text{subject to} & \hat{\mathbf{Q}} \in \mathcal{Q}, \ \hat{\mathbf{W}} \succeq 0 \\
& \begin{bmatrix} \hat{\mathbf{Z}} & \hat{\mathbf{Q}} \\ \hat{\mathbf{Q}} & \hat{\mathbf{Q}} + \hat{\mathbf{W}} \end{bmatrix} \succeq 0
\end{array}
\tag{8}
$$

where the set $\mathcal{Q}$ is given in the form in Theorem 2. It is easy to see that the objective in (8), 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 (8) are satisfied by $(\hat{\mathbf{Q}}, \hat{\mathbf{W}}, \hat{\mathbf{Z}})$, then they are also satisfied by $(\mathbf{D}_{\hat{\mathbf{Q}}}, \mathbf{D}_{\hat{\mathbf{W}}}, \mathbf{D}_{\hat{\mathbf{Z}}})$. First, the LMI in (8) amounts to

$$
\begin{bmatrix} \mathbf{J}_l & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_l \end{bmatrix}
\begin{bmatrix} \hat{\mathbf{Z}} & \hat{\mathbf{Q}} \\ \hat{\mathbf{Q}} & \hat{\mathbf{Q}} + \hat{\mathbf{W}} \end{bmatrix}
\begin{bmatrix} \mathbf{J}_l & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_l \end{bmatrix}
=
\begin{bmatrix} \mathbf{J}_l \hat{\mathbf{Z}} \mathbf{J}_l & \mathbf{J}_l \hat{\mathbf{Q}} \mathbf{J}_l \\ \mathbf{J}_l \hat{\mathbf{Q}} \mathbf{J}_l & \mathbf{J}_l \hat{\mathbf{Q}} \mathbf{J}_l + \mathbf{J}_l \hat{\mathbf{W}} \mathbf{J}_l \end{bmatrix} \succeq 0
\tag{9}
$$

where $\mathbf{J}_l$ is defined in Lemma 3. This implies that if $(\hat{\mathbf{Q}}, \hat{\mathbf{W}}, \hat{\mathbf{Z}})$ satisfies the LMI in (8), 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 3, 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, since $\mathbf{d}(\hat{\mathbf{Q}})$ is majorized by $\boldsymbol{\lambda}(\hat{\mathbf{Q}})$ [5] and each $f_n(\mathbf{x})$ is a Schur-convex function, it follows that $f_n(\boldsymbol{\lambda}(\mathbf{D}_{\hat{\mathbf{Q}}})) = f_n(\mathbf{d}(\mathbf{D}_{\hat{\mathbf{Q}}})) \leq f_n(\boldsymbol{\lambda}(\hat{\mathbf{Q}}))$, implying that if $\hat{\mathbf{Q}} \in \mathcal{Q}$, then $\mathbf{D}_{\hat{\mathbf{Q}}} \in \mathcal{Q}$. Finally, $\hat{\mathbf{W}} \succeq 0$ implies that $\mathbf{D}_{\hat{\mathbf{W}}} \succeq 0$.

Consequently, given any feasible $(\hat{\mathbf{Q}}, \hat{\mathbf{W}}, \hat{\mathbf{Z}})$ in (8), $(\mathbf{D}_{\hat{\mathbf{Q}}}, \mathbf{D}_{\hat{\mathbf{W}}}, \mathbf{D}_{\hat{\mathbf{Z}}})$ not only is feasible but also results in the same objective value as $(\hat{\mathbf{Q}}, \hat{\mathbf{W}}, \hat{\mathbf{Z}})$. Therefore, we can conclude that, in the solution set of (8), 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 $\mathbf{D}_{\hat{\mathbf{Z}}} = \mathbf{\Lambda}_z$. ■

Indicated by Theorem 2, with the conditions on $\mathcal{Q}$ and $\mathbf{T}$, the optimal transmit directions for the worst-case design are just the right singular vectors of the nominal channel. Fortunately, these conditions are satisfied in the most common situation—the sum and maximum power constraints ($\mathcal{Q}_1 \cap \mathcal{Q}_2$) with a spherical uncertainty region ($\mathbf{T} = \mathbf{I}$). Therefore, the eigen-mode transmission is also optimal for the worst-case design, which is analogous to the cases of perfect CSIT [4], [5] and statistical CSIT [7], [8]. Interestingly, the same optimal directions were also obtained in [15] with the uncertainty region defined by the weighted Frobenius norm under similar conditions. Consequently, with this result, we move closer to finding a closed-form solution to the maximin problem (3).

IV. Optimality of Beamforming

The main consequence of applying the result in Theorem 2 is that the matrix-valued problem (4) can be simplified to a power allocation problem without losing any optimality.

Proposition 2: Provided the conditions in Theorem 2 are satisfied, the problem (4) reduces to

$$\underset{\mathbf{p}\in\mathcal{P},\mathbf{w}\geq 0}{\text{maximize}} \;\; \sum_{i=1}^{N} \frac{w_i \gamma_i p_i}{w_i + p_i} - \varepsilon^2 \sum_{i=1}^{N} \frac{w_i}{\tau_i} \tag{10}$$

where $\mathcal{P} = \{\mathbf{p} : \mathbf{p} \geq 0, f_n(\mathbf{p}) \leq P_n, \forall n\}$.

Proof: See Appendix B. ■

It can be verified that the objective in (10) is jointly concave in $(\mathbf{p}, \mathbf{w})$, implying that (10) is a convex problem if, in addition to being Schur-convex, each $f_n(\mathbf{p})$ is a convex function. Both Schur-convexity and convexity are met by the sum power $f_1(\mathbf{p}) = \sum_{i=1}^{N} p_i$ and the maximum power $f_2(\mathbf{p}) = \max_i\{p_i\}$. In particular, when one considers the most frequently-used sum power constraint, the solution to (10) is beamforming over the maximum eigen-mode.

Theorem 3: Let $\mathcal{P} = \{\mathbf{p} : \mathbf{p} \geq 0, \mathbf{1}^T\mathbf{p} \leq P_T\}$. Then, the solution to the problem (10) is $p_1^\star = P_T$ and $p_i^\star = 0$ for $i \geq 2$, and the optimum value of (10) is $P_T(\sqrt{\gamma_1} - \varepsilon/\sqrt{\tau_1})^2$.

Proof: Assume w.l.o.g. that $\{p_i\}$ are ordered decreasingly and $\text{rank}(\hat{\mathbf{H}}) = r$. One can easily find the optimal $\mathbf{w}$ for fixed $\mathbf{p}$ as $w_i^\star = \frac{p_i}{\varepsilon}(\sqrt{\tau_i\gamma_i} - \varepsilon)$ for $i \leq m$ and $w_i^\star = 0$ for $i > m$, where $m \in \{1, \cdots, r\}$ is an integer such that $\tau_m\gamma_m > \varepsilon^2 \geq \tau_{m+1}\gamma_{m+1}$ with $\tau_{r+1}\gamma_{r+1} \triangleq 0$. Substituting $\mathbf{w}^\star$ into (10), the objective function becomes

$$\sum_{i=1}^{m} \frac{w_i^\star \gamma_i p_i}{w_i^\star + p_i} - \varepsilon^2 \sum_{i=1}^{m} \frac{w_i^\star}{\tau_i} = \sum_{i=1}^{m} \left(\sqrt{\gamma_i} - \frac{\varepsilon}{\sqrt{\tau_i}}\right)^2 p_i. \tag{11}$$

Thus, the power allocation problem becomes a simple linear program (LP)

$$\begin{array}{ll} \underset{\mathbf{p}}{\text{maximize}} & \sum_{i=1}^{m} \left(\sqrt{\gamma_i} - \varepsilon/\sqrt{\tau_i}\right)^2 p_i \\ \text{subject to} & \sum_{i=1}^{N} p_i \leq P_T, \; p_1 \geq \cdots \geq p_N \geq 0 \end{array} \tag{12}$$

where we have explicitly taken into account the decreasing order of $\{p_i\}$. Clearly, the solution to (12) is to use all available power P_T on the maximum term $(\sqrt{\gamma_1} - \varepsilon/\sqrt{\tau_1})$. ■

Theorem 3 indicates that, under the sum power constraint, even if the transmitter knows only the nominal channel instead of the actual channel, beamforming along the right singular vector associated with the maximum singular value of the nominal channel can still achieve the best received SNR, in the worst channel within the elliptical uncertainty region defined by the weighted spectral norm. In other words, beamforming is a robust solution for the uncertainty region defined by the spectral norm. To the best of our knowledge, this is the first published result concerning the robustness of beamforming for a general MIMO channel. Given that the maximin problem with the uncertainty region defined by the weighted Frobenius norm admits a closed-form solution too, we summarize the robustness results of beamforming in this paper and in [15] as follows.

Corollary 1: Let $\mathcal{Q} = \mathcal{Q}_1$ and $\mathbf{U}_t = \mathbf{U}_h$. Then, beamforming is the solution to the problem of maximizing the worst-case received SNR if either 1) $\mathbf{\Delta} \in \mathcal{E}_2 = \{\mathbf{\Delta} : \mathbf{\Delta}\mathbf{T}\mathbf{\Delta}^H \preceq \varepsilon^2\mathbf{I}\}$ for any ε; or 2) $\mathbf{\Delta} \in \mathcal{E}_F = \{\mathbf{\Delta} : \text{Tr}(\mathbf{\Delta}\mathbf{T}\mathbf{\Delta}^H) \leq \varepsilon^2\}$ and $\varepsilon \leq \sqrt{\tau_1}(\sqrt{\gamma_1} - \sqrt{\gamma_2})$ [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 (provided the conditions in Theorem 2 are satisfied), regardless of the parameters $\hat{\mathbf{H}}$ and ε. On the other hand, we know from [21, p. 314] that, for a given matrix $\mathbf{A} \in \mathbb{C}^{M\times N}$, the spectral norm is the smallest one among the most common matrix norms such as $\|\mathbf{A}\|_1 = \sum_{i,j} |[\mathbf{A}]_{ij}|$, $\|\mathbf{A}\|_{\max} = \sqrt{NM}\max_{i,j} |[\mathbf{A}]_{ij}|$ or the above mentioned Frobenius norm, $\|\mathbf{A}\|_F$. Consequently, given the same radius ε, $\mathcal{E}_2$ is the biggest one among all ellipsoids defined by these matrix norms, e.g., $\mathcal{E}_F \subseteq \mathcal{E}_2$, which means that $\mathcal{E}_2$ is the most conservative uncertainty region based on these matrix norms. If the uncertainty region is a set smaller than $\mathcal{E}_2$, beamforming is usually not optimal for the maximin design. 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. Therefore, loosely speaking, one can state that beamforming is an approximately robust solution, for the most common norms defining the uncertainty region, when the number of the transmit or receive antennas is small or the channel uncertainty is small. Numerical support of this claim is given in Section VI.

V. The QoS Problem

The QoS problem, as the complement of the maximin problem (3), minimizes the power consumption at the transmitter while keeping the received SNR above a threshold for any channel (error) realization in the uncertainty region. Specifically, denoting by $S(\mathbf{Q})$ the power consumption function at the transmitter, the QoS problem is

$$\begin{aligned} \underset{\mathbf{Q}\succeq 0}{\text{minimize}} \quad & S(\mathbf{Q}) \\ \text{subject to} \quad & \text{Tr}\{(\hat{\mathbf{H}}+\mathbf{\Delta})\mathbf{Q}(\hat{\mathbf{H}}+\mathbf{\Delta})^H\} \geq \rho, \ \forall \mathbf{\Delta} \in \mathcal{E} \end{aligned} \tag{13}$$

where $\rho > 0$ is the QoS threshold and $\mathcal{E}$ is defined in (2). Common power consumption functions include: 1) sum power $S_1(\mathbf{Q}) = \text{Tr}(\mathbf{Q})$; 2) maximum power $S_2(\mathbf{Q}) = \lambda_{\max}(\mathbf{Q})$; 3) maximum power per antenna $S_3(\mathbf{Q}) = \max_i [\mathbf{Q}]_{ii}$.

Theorem 4: The QoS problem (13) is equivalent to

$$\begin{aligned}
\underset{\mathbf{Q},\mathbf{W},\mathbf{Z}}{\text{minimize}} \quad & S(\mathbf{Q}) \\
\text{subject to} \quad & \mathbf{Q} \succeq 0,\ \mathbf{W} \succeq 0 \\
& \mathrm{Tr}\{(\mathbf{Z}-\mathbf{Q})\,\hat{\mathbf{H}}^H\hat{\mathbf{H}}\} + \varepsilon^2 \mathrm{Tr}(\mathbf{W}\mathbf{T}^{-1}) + \rho \leq 0 \\
& \begin{bmatrix} \mathbf{Z} & \mathbf{Q} \\ \mathbf{Q} & \mathbf{Q}+\mathbf{W} \end{bmatrix} \succeq 0.
\end{aligned} \tag{14}$$

Proof: The robust constraint in (13) can be written as $\min_{\mathbf{\Delta}\in\mathcal{E}} \mathrm{Tr}\{(\hat{\mathbf{H}}+\mathbf{\Delta})\mathbf{Q}(\hat{\mathbf{H}}+\mathbf{\Delta})^H\} \geq \rho$, where the minimization can be replaced by its dual maximization, leading to $\max_{\mathbf{W}\succeq 0} G(\mathbf{W}) \geq \rho$ (see (19) in Appendix A). Therefore, the robust constraint can be equivalently replaced by $\mathbf{W} \succeq 0$, $G(\mathbf{W}) \geq \rho$ with an auxiliary variable $\mathbf{W}$. Then, the remainder of the proof follows that of Theorem 1. ■

Theorem 5: Let $\mathbf{U}_t = \mathbf{U}_h$ and $S(\mathbf{Q}) = \sum_n \alpha_n f_n(\lambda(\mathbf{Q}))$, where $f_n(\mathbf{x})$ is a Schur-convex function and $\alpha_n \geq 0$, $\forall n$. Then, $\mathbf{U}_q = \mathbf{U}_h$ is optimal for the QoS problem (13).

Proof: Similar to that of Theorem 2. ■

Theorem 4 provides an efficient way to directly solve the QoS problem, since (14) is a convex problem if $S(\mathbf{Q})$ is a convex function. Moreover, when $S(\mathbf{Q})$ is any positive weighted sum of $S_1(\mathbf{Q})$, $S_2(\mathbf{Q})$ and $S_3(\mathbf{Q})$, (14) is or can be transformed into an SDP. Theorem 5 is the counterpart of Theorem 2. In fact, as the QoS problem (13) and the maximin problem (3) are complementary, (13) can be alternatively solved by solving (3) with $\mathcal{Q} = \{\mathbf{Q} : \mathbf{Q} \succeq 0,\ S(\mathbf{Q}) \leq P\}$, where P is chosen such that the optimum value of (3) is equal to ρ. Therefore, when $S(\mathbf{Q}) = S_1(\mathbf{Q})$, i.e., minimizing the sum power, beamforming over the maximum eigen-mode of the nominal channel is also the optimal solution to the QoS problem.

VI. Numerical Results

We evaluate the robustness of beamforming via several numerical examples in a common situation of the sum power constraint ($\mathcal{Q} = \mathcal{Q}_1$) and a spherical channel uncertainty region ($\mathbf{T} = \mathbf{I}$). The elements of the nominal channel $\hat{\mathbf{H}}$ are randomly generated according to zero-mean, unit-variance, i.i.d. Gaussian distributions. Beamforming is compared with the equal-power transmission and the robust strategy (see [15]) in the Frobenius-norm-defined uncertainty region (FUR) through their worst-case performance. We assume that the channel uncertainty region is FUR (see $\mathcal{E}_F$ in Corollary 1), and set the radius of FUR as $\varepsilon = s||\hat{\mathbf{H}}||_2 = s\sqrt{\gamma_1}$ with $s \in [0, 1)$. In this case, beamforming, which is robust for the spectral-norm-defined uncertainty region (SUR), is generally not a robust solution. However, it will be shown later that the worst-case performance of beamforming is quite close to that of the robust strategy in FUR for small channel dimension or uncertainty.

Figs. 1 and 2 depict the worst-case symbol error rate (SER) and the average minimum transmit power to satisfy a QoS threshold $\rho = 6$dB (see (13)) versus SNR for different values of s, respectively. The average is taken over the nominal channel $\hat{\mathbf{H}}$. The transmit and receive antenna numbers are set to $M = N = 2$, with the full-rate complex STBC [19] and QPSK modulation. As can be observed, for small or even moderate uncertainty (represented by s), the worst-case performance of beamforming is very close to that of the robust strategy in FUR, and better than that of the equal-power transmission, even though beamforming is robust for SUR, but not FUR.

Fig. 3 shows, in the case of $M = N$, the relation between the number of antennas and the average minimum transmit power to satisfy a QoS threshold $\rho = 6$dB with a given uncertainty region size $s = 0.5$. The minimum transmit power of the three transmit strategies is normalized by that of the robust strategy in FUR. It can be observed that the smaller the channel dimension is, the closer the worst-case performance of beamforming is to that of the robust strategy in FUR. Consequently, the numerical results have verified our conclusion that beamforming is approximately robust for small channel uncertainty or small numbers of antennas, rather independent of the norm that defines the uncertainty region.

VII. Conclusion

We have investigated the ability of beamforming to cope with imperfect CSIT in MIMO channels from the perspective of worst-case robustness. It has been shown that beamforming is robust in terms of providing the maximum received SNR or guaranteeing a given received SNR with the minimum transmit power in the worst channel within the spectral-norm-defined uncertainty region. Therefore, beamforming can combat against the imperfectness of CSIT when the channel dimension or the channel uncertainty is small. In addition, we have provided an efficient way to solve the maximin and QoS problems, stemming from the worst-case robust design, under a general transmit power constraint.

Appendix A
Proof of Proposition 1

The main idea is to replace the inner minimization in (3) by its dual maximization, hence transforming a maximin problem to a maximization problem. Defining $\tilde{\mathbf{\Delta}} \triangleq \mathbf{\Delta}\mathbf{T}^{1/2}$, it follows that $\mathcal{E} = \{\mathbf{\Delta} : \mathbf{\Delta}\mathbf{T}\mathbf{\Delta}^H \preceq \varepsilon^2\mathbf{I}\} = \{\tilde{\mathbf{\Delta}} : \tilde{\mathbf{\Delta}}\tilde{\mathbf{\Delta}}^H \preceq \varepsilon^2\mathbf{I}_M\} = \{\tilde{\mathbf{\Delta}} : \tilde{\mathbf{\Delta}}^H\tilde{\mathbf{\Delta}} \preceq \varepsilon^2\mathbf{I}_N\}$. With the definitions of $\tilde{\mathbf{Q}}$ and $\tilde{\mathbf{H}}$ in Proposition 1, the inner minimization in (3) can be rewritten as

$$\underset{\tilde{\mathbf{\Delta}}\in\mathcal{E}}{\text{minimize}} \;\; \text{Tr}\{(\tilde{\mathbf{H}} + \tilde{\mathbf{\Delta}})\tilde{\mathbf{Q}}(\tilde{\mathbf{H}} + \tilde{\mathbf{\Delta}})^H\} \tag{15}$$

whose Lagrangian is given by

$$\begin{aligned} L(\tilde{\mathbf{\Delta}}, \tilde{\mathbf{W}}) &= \text{Tr}\{(\tilde{\mathbf{H}} + \tilde{\mathbf{\Delta}})\tilde{\mathbf{Q}}(\tilde{\mathbf{H}} + \tilde{\mathbf{\Delta}})^H\} + \text{Tr}\{(\tilde{\mathbf{\Delta}}^H\tilde{\mathbf{\Delta}} - \varepsilon^2\mathbf{I})\tilde{\mathbf{W}}\} \\ &= \text{Tr}\{\tilde{\mathbf{\Delta}}(\tilde{\mathbf{Q}} + \tilde{\mathbf{W}})\tilde{\mathbf{\Delta}}^H\} + \text{Tr}(\tilde{\mathbf{H}}\tilde{\mathbf{Q}}\tilde{\mathbf{\Delta}}^H + \tilde{\mathbf{\Delta}}\tilde{\mathbf{Q}}\tilde{\mathbf{H}}^H) + \text{Tr}(\tilde{\mathbf{H}}\tilde{\mathbf{Q}}\tilde{\mathbf{H}}^H) - \varepsilon^2\text{Tr}(\tilde{\mathbf{W}}) \end{aligned} \tag{16}$$

where $\tilde{\mathbf{W}} \succeq 0$ is the Lagrange multiplier. Since $L(\tilde{\mathbf{\Delta}}, \tilde{\mathbf{W}})$ is convex in $\tilde{\mathbf{\Delta}}$ (for fixed $\tilde{\mathbf{W}}$), its minimizer can be achieved by setting $\partial L(\tilde{\mathbf{\Delta}}, \tilde{\mathbf{W}})/\partial\tilde{\mathbf{\Delta}} = 0$, leading to $\tilde{\mathbf{\Delta}}(\tilde{\mathbf{Q}} + \tilde{\mathbf{W}}) = -\tilde{\mathbf{H}}\tilde{\mathbf{Q}}$. Assume w.l.o.g. $\text{rank}(\tilde{\mathbf{Q}} + \tilde{\mathbf{W}}) = r$. Then, it can be verified that the minimizer of $L(\tilde{\mathbf{\Delta}}, \tilde{\mathbf{W}})$ is given by

$$\tilde{\mathbf{\Delta}}^* = -\tilde{\mathbf{H}}\tilde{\mathbf{Q}}(\tilde{\mathbf{Q}} + \tilde{\mathbf{W}})^\dagger + \mathbf{M}\mathbf{P} \tag{17}$$

where $\mathbf{M} \in \mathbb{C}^{M\times N}$ is an arbitrary matrix and $\mathbf{P} \in \mathbb{S}_+^N$ is the orthogonal projector into $\mathcal{N}(\tilde{\mathbf{Q}} + \tilde{\mathbf{W}})$.

Now, the dual function of (15) can be obtained by substituting (17) into $L(\tilde{\mathbf{\Delta}}, \mathbf{W})$, leading to

$$G(\tilde{\mathbf{W}}) = L(\tilde{\mathbf{\Delta}}^*, \tilde{\mathbf{W}}) = \mathrm{Tr}(\tilde{\mathbf{H}}\tilde{\mathbf{Q}}\tilde{\mathbf{H}}^H) - \mathrm{Tr}\{\tilde{\mathbf{H}}\tilde{\mathbf{Q}}(\tilde{\mathbf{Q}} + \tilde{\mathbf{W}})^\dagger \tilde{\mathbf{Q}}\tilde{\mathbf{H}}^H\} - \varepsilon^2 \mathrm{Tr}(\tilde{\mathbf{W}}) + \mathrm{Tr}(\mathbf{M}\mathbf{P}\tilde{\mathbf{Q}}\tilde{\mathbf{H}}^H). \tag{18}$$

Since $\tilde{\mathbf{Q}} + \tilde{\mathbf{W}} \succeq \tilde{\mathbf{Q}}$, it follows from Lemma 1 that $\mathcal{N}(\tilde{\mathbf{Q}} + \tilde{\mathbf{W}}) \subseteq \mathcal{N}(\tilde{\mathbf{Q}})$, implying that $\mathbf{P}\tilde{\mathbf{Q}} = 0$. Hence, the dual problem of (15) is

$$\underset{\tilde{\mathbf{W}} \succeq 0}{\text{maximize}} \;\; G(\tilde{\mathbf{W}}) = \mathrm{Tr}(\tilde{\mathbf{H}}\tilde{\mathbf{Q}}\tilde{\mathbf{H}}^H) - \mathrm{Tr}\{\tilde{\mathbf{H}}\tilde{\mathbf{Q}}(\tilde{\mathbf{Q}} + \tilde{\mathbf{W}})^\dagger \tilde{\mathbf{Q}}\tilde{\mathbf{H}}^H\} - \varepsilon^2 \mathrm{Tr}(\tilde{\mathbf{W}}). \tag{19}$$

Given that (15) is a convex problem with a compact convex feasible set containing a nonempty interior (Slater's condition [18] is satisfied), there is no gap between (15) and (19), so (15) can be equivalently replaced by (19) in the maximin problem (3) leading to a maximization problem, which can be easily rewritten as the minimization problem (5).

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 (4) becomes

$$\begin{aligned} \underset{\mathbf{\Lambda}_q, \mathbf{\Lambda}_w, \mathbf{\Lambda}_z}{\text{minimize}} \quad & \mathrm{Tr}\{(\mathbf{\Lambda}_z - \mathbf{\Lambda}_q)\,\mathbf{\Lambda}_h\} + \varepsilon^2 \mathrm{Tr}(\mathbf{\Lambda}_w \mathbf{\Lambda}_t^{-1}) \\ \text{subject to} \quad & \mathbf{\Lambda}_q \in \mathcal{Q}, \;\; \mathbf{\Lambda}_w \succeq 0 \\ & \begin{bmatrix} \mathbf{\Lambda}_z & \mathbf{\Lambda}_q \\ \mathbf{\Lambda}_q & \mathbf{\Lambda}_q + \mathbf{\Lambda}_w \end{bmatrix} \succeq 0. \end{aligned} \tag{20}$$

The LMI in (20), via proper symmetric column and row permutations, can be rewritten as

$$\begin{bmatrix} z_i & p_i \\ p_i & p_i + w_i \end{bmatrix} \succeq 0, \;\; \forall i \tag{21}$$

which represent the same constraints as

$$z_i\,(p_i + w_i) \geq p_i^2, \;\; p_i + w_i \geq 0, \;\; z_i \geq 0, \;\; \forall i. \tag{22}$$

Hence, (20) is equivalent to

$$\begin{aligned} \underset{\mathbf{p}, \mathbf{w}, \mathbf{z}}{\text{minimize}} \quad & \textstyle\sum_{i=1}^N \gamma_i\,(z_i - p_i) + \varepsilon^2 \sum_{i=1}^N \frac{w_i}{\tau_i} \\ \text{subject to} \quad & \mathbf{p} \in \mathcal{P}, \;\; \mathbf{w} \geq 0, \;\; \mathbf{z} \geq 0 \\ & z_i\,(w_i + p_i) \geq p_i^2, \;\; \forall i. \end{aligned} \tag{23}$$

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 (23) amounts to

$$\begin{aligned} \underset{\mathbf{p}, \mathbf{w}, \mathbf{z}}{\text{minimize}} \quad & \textstyle\sum_{i \in \mathcal{I}} \gamma_i\,(z_i - p_i) + \varepsilon^2 \sum_{i \in \mathcal{I}} \frac{w_i}{\tau_i} \\ \text{subject to} \quad & \mathbf{p} \in \mathcal{P}, \;\; \mathbf{w} \geq 0, \;\; \mathbf{z} \geq 0 \\ & z_i\,(w_i + p_i) \geq p_i^2, \;\; i \in \mathcal{I}. \end{aligned} \tag{24}$$

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, (24) is equivalent to

$$\underset{\mathbf{p} \in \mathcal{P}, \mathbf{w} \geq 0}{\text{minimize}} \;\; \sum_{i \in \mathcal{I}} \frac{\gamma_i p_i^2}{w_i + p_i} - \sum_{i \in \mathcal{I}} \gamma_i p_i + \varepsilon^2 \sum_{i \in \mathcal{I}} \frac{w_i}{\tau_i} \tag{25}$$

whose objective amounts to

$$-\sum_{i\in\mathcal{I}}\frac{w_i\gamma_i p_i}{w_i+p_i}+\varepsilon^2\sum_{i\in\mathcal{I}}\frac{w_i}{\tau_i}=-\sum_{i=1}^{N}\frac{w_i\gamma_i p_i}{w_i+p_i}+\varepsilon^2\sum_{i=1}^{N}\frac{w_i}{\tau_i} \tag{26}$$

where the second equality is due to $w_i = 0$ for $i \notin \mathcal{I}$. Finally, (10) is just the maximizing counterpart of (25).

Acknowledgment

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

References

[1] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, Nov.-Dec. 1999.

[2] G. Foschini and M. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, no. 3, pp. 311–335, 1998.

[3] 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.

[4] 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.

[5] 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.

[6] 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.

[7] 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.

[8] 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.

[9] S. Srinivasa and S. A. Jafar, "The optimality of transmit beamforming: A unified view," *IEEE Trans. Inform. Theory*, vol. 53, no. 4, pp. 1558–1564, Apr. 2007.

[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] 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.

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

[16] 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.

[17] J. Li and P. Stoica, *Robust Adaptive Beamforming*. Hoboken, NJ: John Wiley & Sons, 2006.

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

[19] 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.

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

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

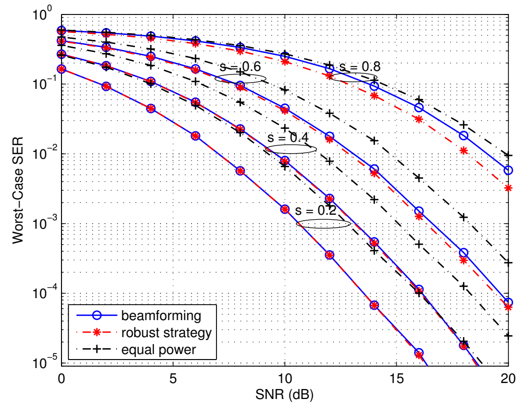

Fig. 1. Worst-case SER versus SNR for different sizes (i.e., s) of the uncertainty region and $M = N = 2$.

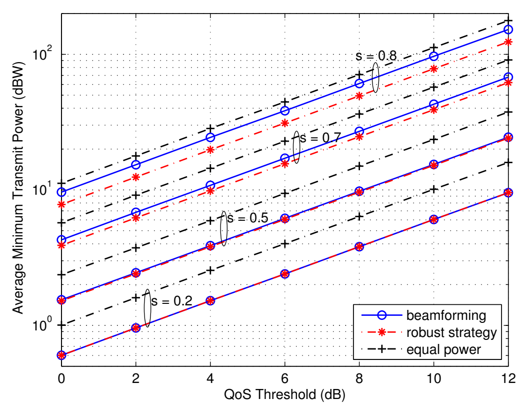

Fig. 2. Average minimum transmit power versus the QoS threshold for different sizes (i.e., s) of the uncertainty region and $M = N = 2$.

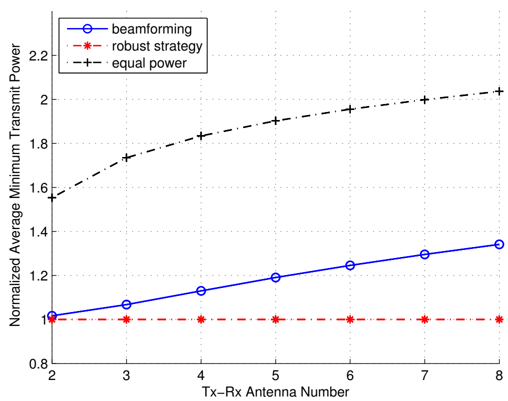

Fig. 3. Normalized average minimum transmit power versus the number of antennas for the uncertainty region size $s = 0.5$ and QoS threshold $\rho = 6$dB.