DISTRIBUTIONALLY ROBUST CHANCE-CONSTRAINED TRANSMIT BEAMFORMING FOR MULTIUSER MISO DOWNLINK

Qiang Li[⋆], Anthony Man-Cho So[†] and Wing-Kin Ma[‡]

[⋆]School of Comm. & Info. Eng., University of Electronic Science & Technology of China, P. R. China
[†]Dept. of Sys. Eng. & Eng. Management, The Chinese University of Hong Kong, Hong Kong
[‡]Dept. of Electronic Engineering, The Chinese University of Hong Kong, Hong Kong
lq@uestc.edu.cn, manchoso@se.cuhk.edu.hk, wkma@ieee.org

ABSTRACT

This paper considers robust transmit beamforming for multiuser multi-input single-output (MISO) downlink transmission, where imperfect channel state information (CSI) is assumed at the base station (BS). The imperfect CSI is captured by a moment-based random error model, in which the BS knows only the mean and covariance of each CSI error, but not the exact distribution. Under this error model, we formulate a distributionally robust beamforming (DRB) problem, in which the total transmit power at the BS is to be minimized, while each user's SINR outage probability, evaluated w.r.t. *any* distribution with the given mean and covariance, is kept below a given threshold. The DRB problem is a semi-infinite chance-constrained problem. By employing recent results in distributionally robust optimization, we show that the DRB problem admits an explicit conic reformulation, which can be conveniently turned into a convex optimization problem after semidefinite relaxation (SDR). We also consider the case where the mean and covariance are not perfectly known. We show that the resulting DRB problem still admits a conic reformulation and can be approximately solved using SDR. The robustness of the proposed designs are demonstrated by numerical simulations.

Index Terms— robust transmit beamforming, distributionally robust optimization, semidefinite relaxation

1. INTRODUCTION

Owing to its simplicity and capability of leveraging system performance, transmit beamforming, a spatial diversity technique, has been widely employed in wireless communications recently. In this paper, we consider multiuser multi-input single-output (MISO) downlink transmission using transmit beamforming. Under this setting, a classical beamformer design formulation is to minimize the total transmit power at the base station (BS) while providing certain quality of service (QoS), e.g., the signal-to-interference-plus-noise ratio (SINR), for each user. In the last decade, different approaches have been proposed to tackle this problem, such as the uplink-downlink duality approach [1,2], the semidefinite relaxation (SDR) approach [3, 4] and the second-order cone (SOCP) approach [5]. It should be noted, however, that all these approaches assume perfect channel state information (CSI) at the BS, which in practice may not be possible because of channel estimation and/or quantization errors. In view of this, there has been growing interest in beamformer designs that are robust to CSI errors. Currently, there are two main approaches to designing such beamformers. The first assumes that the errors are adversarially chosen from some (bounded) set, which results in worst-case robust beamformer designs [6–9]. In this approach, one does not utilize any distributional properties of the errors. The second assumes that the errors follow a certain fully-specified distribution, whose properties are then exploited to yield outage-constrained robust beamformer designs [10–14]. In practice, however, neither of these approaches is satisfactory, as we typically are able to obtain some, but not complete, information about the error distribution through channel estimates. This raises the natural question of whether an alternative, tractable model for robust beamformer design can be devised to better account for the available distributional information of the CSI error.

In this paper, we depart from the aforementioned error models and consider another *moment-based* random CSI error model, in which the BS has no a priori knowledge of the error distribution except for its first and second-order statistics. Such a model is motivated by the observation that it is relatively easy for a BS to have accurate estimates of the error mean and covariance from accumulated channel estimates. Under this moment-based error model, we formulate a new *distributionally robust* beamforming (DRB) problem, in which the beamformers are designed to minimize the total transmit power at the BS, while the SINR outage probability of each user, evaluated w.r.t. *any* distribution with the given mean and covariance, is kept below a given threshold. The DRB problem is a semi-infinite chance-constrained optimization problem, which is generally hard to solve. Nevertheless, by employing recent results in distributionally robust optimization [15–17], we show that the seemingly challenging DRB problem actually admits an explicit conic reformulation, which can then be conveniently approximated using the SDR technique [18]. We also consider the DRB problem for the case where the mean and covariance are not perfectly known. We show that the resulting DRB problem still admits a conic reformulation and can be approximately solved using SDR.

Before we present our setup and results, let us mention some related works. The MISO downlink transmit beamforming problem has been extensively studied in the past; see, e.g., [12] and the references therein. Some representative works include [2–5] for the case of perfect CSI at the BS, and [6–13] for the case of imperfect CSI. The robust beamformer design problem can also be tackled using the sample average approximation approach developed in the recent work [19]. Such an approach uses channel estimates to approximate the outage probability and hence does not require any knowledge of the error distribution. However, the size of the resulting optimization problem can be huge, which limits the practicality of the approach.

2. SYSTEM MODEL AND PROBLEM STATEMENT

Consider a multiuser downlink transmission, where a base station (BS) simultaneously sends K data streams to K users with each data stream exclusively for one user. The BS is equipped with N antennas and all the receivers have a single antenna, i.e., MISO downlink. Assuming that transmit beamforming is employed at the BS, the transmit signal $\boldsymbol{x}(t) \in \mathbb{C}^N$ at time t may be expressed as

$$\boldsymbol{x}(t) = \sum_{i=1}^{K} \boldsymbol{w}_i s_i(t), \tag{1}$$

where $s_i(t)$ is the data stream for user i with unit power; i.e, $\mathbb{E}[s_i(t)] = 1, \forall\, i$, and $\boldsymbol{w}_i \in \mathbb{C}^N$ is the beamforming vector for the ith data stream. Assuming frequency-flat and slow fading channels, the received signal $y_i(t)$ at the ith user is given by

$$y_i(t) = \boldsymbol{h}_i^H \boldsymbol{x}(t) + n_i(t),\; i \in \mathcal{K}, \tag{2}$$

where $\mathcal{K} \triangleq \{1, \ldots, K\}$, $\boldsymbol{h}_i \in \mathbb{C}^N$ is the channel vector from the BS to the ith user, and $n_i(t) \sim \mathcal{CN}(0, \sigma_i^2)$ is the additive white Gaussian noise with mean zero and variance σ_i^2. According to (1) and (2), the received SINR at the ith user may be calculated as

$$\mathsf{SINR}_i \triangleq \frac{|\boldsymbol{h}_i^H \boldsymbol{w}_i|^2}{\sum_{j \neq i} |\boldsymbol{h}_i^H \boldsymbol{w}_j|^2 + \sigma_i^2}, \quad \forall\, i \in \mathcal{K}. \tag{3}$$

Due to imperfect channel estimation and/or feedback errors, the BS usually has only some rough knowledge of $\boldsymbol{h}_i$. To describe the imperfect CSI at the BS, we consider the following random CSI error model:

$$\boldsymbol{h}_i = \bar{\boldsymbol{h}}_i + \Delta \boldsymbol{h}_i, \quad i \in \mathcal{K}, \tag{4}$$

where $\bar{\boldsymbol{h}}_i$ is the presumed CSI at the BS, and $\Delta \boldsymbol{h}_i$ is the associated CSI error, which is randomly distributed with mean $\boldsymbol{\xi}_i \in \mathbb{C}^N$ and covariance $\boldsymbol{\Sigma}_i \in \mathbb{H}_{++}^N$. Different from [10–13], herein we do not impose any particular distribution (such as Gaussian or uniform distribution) on $\Delta \boldsymbol{h}_i$. Instead, we allow $\Delta \boldsymbol{h}_i$ to be arbitrarily distributed as long as its distribution, denoted by ΔF_i, has mean $\boldsymbol{\xi}_i$ and covariance $\boldsymbol{\Sigma}_i$; i.e.,

$$\Delta \boldsymbol{h}_i \sim \Delta F_i \in \mathscr{D}(\boldsymbol{\xi}_i, \boldsymbol{\Sigma}_i), \tag{5}$$

where $\mathscr{D}(\boldsymbol{\xi}_i, \boldsymbol{\Sigma}_i)$ denotes the set of distributions with mean $\boldsymbol{\xi}_i$ and covariance $\boldsymbol{\Sigma}_i$. We assume for now that $\boldsymbol{\xi}_i$ and $\boldsymbol{\Sigma}_i$ are known, and the case of imperfect $\boldsymbol{\xi}_i$ and $\boldsymbol{\Sigma}_i$ will be considered in Sec. 4. The error model (5) is motivated by the fact that it is more convenient and easier for the BS to have the statistics rather than the accurate distribution of the CSI error from the accumulated CSI estimates.

Under the above system model, we consider the following distributionally robust beamforming (DRB) problem:

$$\min_{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_K \in \mathbb{C}^N} \sum_{i=1}^{K} \|\boldsymbol{w}_i\|^2 \tag{6a}$$

$$\text{s.t.} \min_{\Delta F_i \in \mathscr{D}(\boldsymbol{\xi}_i, \boldsymbol{\Sigma}_i)} \mathbb{P}_{\Delta \boldsymbol{h}_i \sim \Delta F_i} \{\mathsf{SINR}_i \geq \gamma_i\} \geq 1 - \epsilon_i, \, \forall\, i, \tag{6b}$$

where $\gamma_i > 0, \forall\, i \in \mathcal{K}$ is a given SINR threshold required by user i, and $0 < \epsilon_i < 1$ specifies the SINR outage probability; i.e., the chance of the ith receiver's SINR falling below γ_i in the presence of CSI error. Clearly, the DRB problem (6) is a chance-constrained problem, which is generally difficult to solve. In particular, the difficulty of (6) mainly arises from the following two aspects: First, the outage probability $\mathbb{P}_{\Delta \boldsymbol{h}_i \sim \Delta F_i} \{\mathsf{SINR}_i \geq \gamma_i\}$ generally has no closed form. Even if it has, the resulting constraint is likely to be non-convex. Secondly, the set $\mathscr{D}(\boldsymbol{\xi}_i, \boldsymbol{\Sigma}_i)$ typically contains infinitely many distributions, which gives rise to an infinite number of chance constraints in (6b). Therefore, the DRB problem (6) *per se* is a semi-infinite chance-constrained optimization problem.

Nevertheless, in the next section, we will develop a tractable approach to problem (6) by employing recent results in distributionally robust optimization [15–17] and the SDR technique [18]. The former gives an explicit conic reformulation of (6b), while the latter provides us with a tractable approach to obtaining an (approximate) solution to (6).

3. A TRACTABLE APPROACH TO THE DRB PROBLEM

For ease of exposition, let us denote $\boldsymbol{W}_i = \boldsymbol{w}_i \boldsymbol{w}_i^H, \forall\, i \in \mathcal{K}$. Clearly, it follows from (4) and (5) that $\boldsymbol{h}_i \sim F_i \in \mathscr{D}(\bar{\boldsymbol{h}}_i + \boldsymbol{\xi}_i, \boldsymbol{\Sigma}_i)$. Hence, the constraint (6b) can be expressed as

$$\min_{F_i \in \mathscr{D}(\bar{\boldsymbol{h}}_i + \boldsymbol{\xi}_i, \boldsymbol{\Sigma}_i)} \mathbb{P}_{\boldsymbol{h}_i \sim F_i} \{L_i(\boldsymbol{h}_i, \{\boldsymbol{W}_l\}_{l \in \mathcal{K}}) \leq 0\} \geq 1 - \epsilon_i, \forall\, i, \tag{7}$$

where $L_i(\boldsymbol{h}_i, \{\boldsymbol{W}_l\}_{l \in \mathcal{K}}) = \boldsymbol{h}_i^H (\sum_{k \neq i} \boldsymbol{W}_k - \gamma_i^{-1} \boldsymbol{W}_i)\boldsymbol{h}_i + \sigma_i^2$.

Next, we show that the semi-infinite chance constraint (7) admits an explicit conic reformulation. Specifically, we have

Theorem 1. *The constraint* (7) *is equivalent to*

$$0 \geq \min_{\beta_i \in \mathbb{R}, \boldsymbol{M}_i \in \mathbb{H}^{N+1}} \beta_i + \epsilon_i^{-1} \mathrm{Tr}(\boldsymbol{\Omega}_i \boldsymbol{M}_i) \tag{8a}$$

$$\text{s.t.}\; \boldsymbol{M}_i \succeq \begin{bmatrix} \sum_{k \neq i} \boldsymbol{W}_k - \gamma_i^{-1} \boldsymbol{W}_i & \boldsymbol{0} \\ \boldsymbol{0}^T & \sigma_i^2 - \beta_i \end{bmatrix}, \tag{8b}$$

$$\boldsymbol{M}_i \succeq \boldsymbol{0}, \tag{8c}$$

for all $i \in \mathcal{K}$, *where* $\boldsymbol{\Omega}_i = \begin{bmatrix} \boldsymbol{\Sigma}_i + (\bar{\boldsymbol{h}}_i + \boldsymbol{\xi}_i)(\bar{\boldsymbol{h}}_i + \boldsymbol{\xi}_i)^H & \bar{\boldsymbol{h}}_i + \boldsymbol{\xi}_i \\ (\bar{\boldsymbol{h}}_i + \boldsymbol{\xi}_i)^H & 1 \end{bmatrix}.$

Proof. See the Appendix. ∎

Invoking Theorem 1, we can express the DRB problem (6) as

$$\min_{\{\boldsymbol{W}_l, \boldsymbol{w}_l\}_{l \in \mathcal{K}}} \sum_{i=1}^{K} \mathrm{Tr}(\boldsymbol{W}_i) \tag{9a}$$

$$\text{s.t.} \min_{(\beta_i, \boldsymbol{M}_i) \in \mathcal{Z}_i} \beta_i + \epsilon_i^{-1} \mathrm{Tr}(\boldsymbol{\Omega}_i \boldsymbol{M}_i) \leq 0, \, \forall\, i \in \mathcal{K}, \tag{9b}$$

$$\boldsymbol{W}_i = \boldsymbol{w}_i \boldsymbol{w}_i^H, \, \forall\, i \in \mathcal{K}, \tag{9c}$$

where $\mathcal{Z}_i \triangleq \{\beta_i \in \mathbb{R}, \boldsymbol{M}_i \in \mathbb{H}^{N+1} : (8b) \text{ and } (8c) \text{ hold}\}$. It is easy to see that (9b) holds if and only if there exists a feasible point $(\beta_i, \boldsymbol{M}_i) \in \mathcal{Z}_i$ such that $\beta_i + \epsilon_i^{-1} \mathrm{Tr}(\boldsymbol{\Omega}_i \boldsymbol{M}_i) \leq 0$. Thus, problem (9) can be rewritten as

$$\min_{\{\boldsymbol{W}_l, \boldsymbol{w}_l, \beta_l, \boldsymbol{M}_l\}_{l \in \mathcal{K}}} \sum_{i=1}^{K} \mathrm{Tr}(\boldsymbol{W}_i)$$
$$\text{s.t.} \quad \beta_i + \epsilon_i^{-1} \mathrm{Tr}(\boldsymbol{\Omega}_i \boldsymbol{M}_i) \leq 0, \, \forall\, i \in \mathcal{K}, \tag{10}$$
$$(8b) - (8c) \text{ and } (9c), \, \forall\, i \in \mathcal{K}.$$

Thus far, we have derived an equivalent reformulation of (6). However, problem (10) is still non-convex due to the non-convex constraint (9c). To circumvent this difficulty, we resort to the SDR approach. Specifically, we replace $\boldsymbol{W}_i = \boldsymbol{w}_i \boldsymbol{w}_i^H$ with $\boldsymbol{W}_i \succeq \boldsymbol{0}$ and drop the non-convex rank-one constraint on $\boldsymbol{W}_i$ to get the following convex relaxation of (10):

$$\min_{\{\boldsymbol{W}_l, \beta_l, \boldsymbol{M}_l\}_{l \in \mathcal{K}}} \sum_{i=1}^{K} \mathrm{Tr}(\boldsymbol{W}_i)$$
$$\text{s.t.} \; \beta_i + \epsilon_i^{-1} \mathrm{Tr}(\boldsymbol{\Omega}_i \boldsymbol{M}_i) \leq 0, \, \forall\, i \in \mathcal{K},$$
$$\boldsymbol{M}_i \succeq \begin{bmatrix} \sum_{k \neq i} \boldsymbol{W}_k - \gamma_i^{-1} \boldsymbol{W}_i & \boldsymbol{0} \\ \boldsymbol{0}^T & \sigma_i^2 - \beta_i \end{bmatrix}, \, \forall i \in \mathcal{K}, \tag{11}$$
$$\boldsymbol{M}_i \succeq \boldsymbol{0}, \; \boldsymbol{W}_i \succeq \boldsymbol{0}, \, \forall i \in \mathcal{K}.$$

Problem (11) is a semidefinite program, which can be efficiently solved with off-the-shelf optimization softwares, such as CVX [20] and SeDuMi [21]. Let $\{\boldsymbol{W}_i^\star\}_{i \in \mathcal{K}}$ be an optimal solution to (11). If rank$(\boldsymbol{W}_i^\star) \leq 1$, $\forall i \in \mathcal{K}$, then an optimal solution $\{\boldsymbol{w}_i^\star\}_{i \in \mathcal{K}}$ for the DRB problem (6) can be obtained through eigen-decomposition. Otherwise, we can perform Gaussian randomization to generate an approximate solution to (6). Readers are referred to [18] for more details on Gaussian randomization. Curiously, our simulation results in Sec. 5 reveal that the optimal $\boldsymbol{W}_i^\star$, $\forall i \in \mathcal{K}$ of (11) returned by CVX is always rank-one. Similar observations have also been mentioned in [9, 12] when SDR is applied to handle other robust beamforming designs.

4. DRB UNDER IMPERFECT MEAN AND COVARIANCE

In the last section, we have developed an SDR approach to the DRB problem (6) when the mean $\boldsymbol{\xi}_i$ and covariance $\boldsymbol{\Sigma}_i$ of $\Delta \boldsymbol{h}_i$ are perfectly known at the BS. In practice, however, both $\boldsymbol{\xi}_i$ and $\boldsymbol{\Sigma}_i$ may subject to some uncertainty. To accommodate this, we generalize the previous CSI error model (5) by incorporating the uncertainty on $\boldsymbol{\xi}_i$ and $\boldsymbol{\Sigma}_i$ as follows [16, 17]:

$$\Delta \boldsymbol{h}_i \sim \Delta F_i \in \mathscr{D}(\bar{\boldsymbol{\xi}}_i, \bar{\boldsymbol{\Sigma}}_i, \tau_{i,1}, \tau_{i,2}), \ \forall i \in \mathcal{K}. \tag{12}$$

Here, $\bar{\boldsymbol{\xi}}_i \in \mathbb{C}^N$ and $\bar{\boldsymbol{\Sigma}}_i \in \mathbb{H}_{++}^N$ are the estimated mean and covariance of $\Delta \boldsymbol{h}_i$; $\tau_{i,1} \geq 0$ and $\tau_{i,2} \geq 1$, where $i \in \mathcal{K}$, are given constants; $\mathscr{D}(\bar{\boldsymbol{\xi}}_i, \bar{\boldsymbol{\Sigma}}_i, \tau_{i,1}, \tau_{i,2})$ denotes the set of distributions of $\Delta \boldsymbol{h}_i$ such that

$$\begin{cases} (\mathbb{E}[\Delta \boldsymbol{h}_i] - \bar{\boldsymbol{\xi}}_i)^H \bar{\boldsymbol{\Sigma}}_i^{-1} (\mathbb{E}[\Delta \boldsymbol{h}_i] - \bar{\boldsymbol{\xi}}_i) \leq \tau_{i,1}, & (13a) \\ \mathbb{E}[(\Delta \boldsymbol{h}_i - \bar{\boldsymbol{\xi}}_i)(\Delta \boldsymbol{h}_i - \bar{\boldsymbol{\xi}}_i)^H] \preceq \tau_{i,2} \bar{\boldsymbol{\Sigma}}_i. & (13b) \end{cases}$$

Roughly speaking, (13a) means that $\boldsymbol{\xi}_i$ lies in an ellipsoid of size $\tau_{i,1}$ centered at the estimate $\bar{\boldsymbol{\xi}}_i$, and (13b) requires that the centered second-moment matrix $\mathbb{E}[(\Delta \boldsymbol{h}_i - \bar{\boldsymbol{\xi}}_i)(\Delta \boldsymbol{h}_i - \bar{\boldsymbol{\xi}}_i)^H]$ should have a similar structure as $\bar{\boldsymbol{\Sigma}}_i$ [16].

With the above uncertainty model, we formulate the following distributionally robust beamforming problem with moment uncertainty (DRB-MU):

$$\min_{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_K \in \mathbb{C}^N} \sum_{i=1}^{K} \|\boldsymbol{w}_i\|^2 \tag{14a}$$

$$\text{s.t.} \min_{\Delta F_i \in \mathscr{D}\left(\substack{\bar{\boldsymbol{\xi}}_i, \bar{\boldsymbol{\Sigma}}_i, \\ \tau_{i,1}, \tau_{i,2}}\right)} \mathbb{P}_{\Delta \boldsymbol{h}_i \sim \Delta F_i} \{\text{SINR}_i \geq \gamma_i\} \geq 1 - \epsilon_i, \ \forall i. \tag{14b}$$

The DRB-MU problem appears to be more challenging than the DRB problem (6). Nevertheless, the next theorem shows that the chance constraint (14b) also admits an explicit conic reformulation.

Theorem 2. *The constraint* (14b) *is equivalent to*

$$0 \geq \min_{r_i, t_i, \boldsymbol{Q}_i, \boldsymbol{q}_i, \nu_i} r_i + t_i + \nu_i - \nu_i \epsilon_i \tag{15a}$$

$$\text{s.t.} \begin{bmatrix} \boldsymbol{Q}_i & \boldsymbol{q}_i/2 \\ \boldsymbol{q}_i^H/2 & r_i + \nu_i \end{bmatrix} \succeq \boldsymbol{0}, \quad \nu_i \geq 0, \quad \boldsymbol{Q}_i \succeq \boldsymbol{0}, \tag{15b}$$

$$\begin{bmatrix} \boldsymbol{Q}_i & \boldsymbol{q}_i/2 \\ \boldsymbol{q}_i^H/2 & r_i \end{bmatrix} \succeq \begin{bmatrix} \sum_{k \neq i} \boldsymbol{W}_k - \boldsymbol{W}_i/\gamma_i & \boldsymbol{0} \\ \boldsymbol{0} & \sigma_i^2 \end{bmatrix}, \tag{15c}$$

$$\begin{aligned} &\sqrt{\tau_{i,1}} \|\bar{\boldsymbol{\Sigma}}_i^{1/2}(\boldsymbol{q}_i + 2\boldsymbol{Q}_i \bar{\boldsymbol{\mu}}_i)\| + \Re e(\bar{\boldsymbol{\mu}}_i^H \boldsymbol{q}_i) \\ &+ \text{Tr}\left(\boldsymbol{Q}_i(\tau_{i,2} \bar{\boldsymbol{\Sigma}}_i + \bar{\boldsymbol{\mu}}_i \bar{\boldsymbol{\mu}}_i^H)\right) \leq t_i, \end{aligned} \tag{15d}$$

for all $i \in \mathcal{K}$, *where* $\bar{\boldsymbol{\mu}}_i \triangleq \bar{\boldsymbol{h}}_i + \bar{\boldsymbol{\xi}}_i$, $\forall i \in \mathcal{K}$.

The proof of Theorem 2 is omitted due to page limit. A key step in proving Theorem 2 is to apply Lemma 1 in [16]. Now, we replace (14b) with (15) and again apply SDR to obtain the following convex relaxation of (14):

$$\min_{\substack{\{\boldsymbol{W}_i, r_i, t_i, \\ \boldsymbol{Q}_i, \boldsymbol{q}_i, \nu_i\}_{i \in \mathcal{K}}}} \sum_{i=1}^{K} \text{Tr}(\boldsymbol{W}_i) \tag{16a}$$

$$\text{s.t.} \ r_i + t_i + \nu_i - \nu_i \epsilon_i \leq 0, \ \forall i \in \mathcal{K}, \tag{16b}$$

$$(15\text{b}) - (15\text{d}) \text{ satisfied, } \boldsymbol{W}_i \succeq \boldsymbol{0}, \ \forall i \in \mathcal{K}, \tag{16c}$$

Remark. In (13), one still needs to determine an appropriate pair of $(\tau_{i,1}, \tau_{i,2})$ such that the true distribution is included in the uncertainty set (12). In practice, the determination of $(\tau_{i,1}, \tau_{i,2})$ can be done using a data-driven approach, which gives a way to determine $(\tau_{i,1}, \tau_{i,2})$ and the resulting confidence level that the uncertainty set (12) contains the true distribution; see [16, 17]. More details on this data-driven approach will be discussed in the full paper.

5. SIMULATION RESULTS AND CONCLUSIONS

In this section, we demonstrate the efficacy of the proposed robust designs using Monte-Carlo simulations. The simulation settings are as follows: The number of antennas at the BS is $N = 5$. There are $K = 3$ users and the noise at each user has unit variance, i.e., $\sigma_1^2 = \cdots = \sigma_K^2 = 1$. For simplicity, we set $\boldsymbol{\xi}_i = \boldsymbol{0}$, $\boldsymbol{\Sigma}_i = 0.002\boldsymbol{I}$, $\gamma_i = \gamma$, $\epsilon_i = \epsilon$, $\tau_{i,1} = 0.5$ and $\tau_{i,2} = 1.5$ for all $i \in \mathcal{K}$. All the results were averaged over 100 feasible channel realizations.

In the first example, we consider the DRB problem (6) and compare our proposed robust design (cf. (11)) with the Bernstein-type inequality design in [12]. The Bernstein-type design deals with a different SINR outage-constrained beamforming problem, where the channel error distribution is assumed to be complex Gaussian; i.e., by replacing $\mathscr{D}(\boldsymbol{\xi}_i, \boldsymbol{\Sigma}_i)$ with the singleton $\{\mathcal{CN}(\boldsymbol{\xi}_i, \boldsymbol{\Sigma}_i)\}$ in (6b). Fig. 1 plots the average transmit power against the SINR threshold γ for $\epsilon = 0.1$ and 0.2. From the figure, we see that the Bernstein-type design yields lower power consumptions than the DRB design. This is expected because after replacing $\mathscr{D}(\boldsymbol{\xi}_i, \boldsymbol{\Sigma}_i)$ with the singleton $\{\mathcal{CN}(\boldsymbol{\xi}_i, \boldsymbol{\Sigma}_i)\}$ in (6b), the Bernstein-type design actually deals with a relaxed problem of (6). However, such a replacement or relaxation could also make the resulting Bernstein-type solution violate the distributionally robust constraint (6b); i.e., the worst-case outage probability associated with the Bernstein-type solution[1] could be larger than the threshold ϵ. To verify this, Table 1 shows the worst-case outage probability of the two methods for a target $\epsilon = 0.1$. Clearly, the Bernstein-type design cannot satisfy the worst-case outage probability requirement, while the proposed DRB design can. Moreover, the results in Fig. 1 and Table 1 reveal that the distribution achieving the minimum in (6b) should not be a complex Gaussian. In addition, from our extensive numerical tests, we found that the optimal solution to (11) returned by CVX is always rank-one, which implies that solving the SDR (11) automatically gives us the optimal beamformer of (6) for the tested scenarios.

The second example considers the DRB-MU problem (14). We compare the proposed DRB-MU design (cf. (16)) with DRB design (11). Notice that the result of DRB is obtained by solving (11) with the estimated $\bar{\boldsymbol{h}}_i$ and $\bar{\boldsymbol{\Sigma}}_i$ without considering the uncertainties associated with the mean and covariance. Fig. 2 and Table 2 show

---

[1]The worst-case SINR outage probability for Bernstein-type design is obtained by fixing $\{\boldsymbol{W}_i\}_{i \in \mathcal{K}}$ in (11) at the Bernstein-type design solution and searching for the minimum ϵ such that problem (11) is feasible. The latter can be done by performing bisection on ϵ.

Table 1: Achieved average SINR outage probability of DRB and Bernstein with target $\epsilon = 0.10$

Algorithm	SINR threshold γ (dB)					
	3	5	7	9	11	13
DRB	0.10	0.10	0.10	0.10	0.10	0.10
Bernstein	0.30	0.31	0.31	0.31	0.31	0.31

respectively the average transmit power and the achieved worst-case SINR outage probability of the two designs, when we increase the SINR requirement γ. Similar to Fig. 1 and Table 1, we see that while DRB consumes less transmit power than DRB-MU, DRB cannot fulfill the worst-case SINR-outage-probability requirement. By contrast, DRB-MU can always meet the outage probability requirement. In addition, we also found that the optimal solution obtained from the SDR (16) is always rank-one. Hence, the result of DRB-MU in Fig. 2 actually is already optimal for problem (14).

To conclude, we have considered two types of distributionally robust transmit beamforming problems and developed tractable approximate solutions for both of them by employing robust optimization methodology and SDR technique. Curiously, our extensive numerical results reveal that the proposed SDRs for the considered two robust designs are always tight. As a future work, it would be interesting to analyze why SDR performs so well in this context.
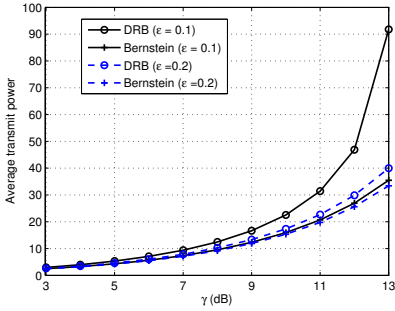

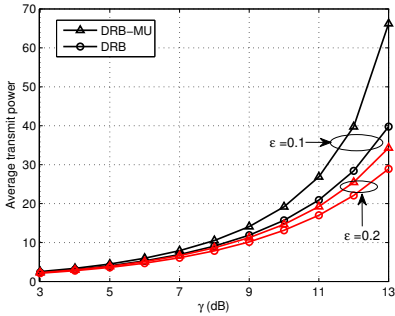Figure 1: SINR threshold γ vs. the average transmit power under perfect mean and covariance.

Figure 2: SINR threshold γ vs. the average transmit power under imperfect mean and covariance.

6. APPENDIX

The proof of Theorem 1 follows from [15]. To start, we need the following two lemmas:

Lemma 1 ([15, Theorem 2.2]). *Let $f : \mathbb{C}^N \to \mathbb{R}$ be a continuous function. Suppose that $f(\boldsymbol{\eta})$ is either concave or (possibly non-*

Table 2: Achieved average SINR outage probability of DRB-MU and DRB with target $\epsilon = 0.10$

Algorithm	SINR threshold γ (dB)					
	3	5	7	9	11	13
DRB-MU	0.10	0.10	0.10	0.10	0.10	0.10
DRB	0.16	0.16	0.16	0.16	0.16	0.15

concave) quadratic in $\boldsymbol{\eta}$. Then the following equivalence holds

$$\sup_{F \in \mathscr{D}(\boldsymbol{\mu}, \boldsymbol{\Sigma})} \mathrm{CVaR}_\epsilon(f(\boldsymbol{\eta})) \leq 0$$
$$\iff \inf_{F \in \mathscr{D}(\boldsymbol{\mu}, \boldsymbol{\Sigma})} \mathbb{P}_{\boldsymbol{\eta} \sim F} \{f(\boldsymbol{\eta}) \leq 0\} \geq 1 - \epsilon, \quad (17)$$

where $0 < \epsilon < 1$ and $\mathrm{CVaR}_\epsilon(f(\boldsymbol{\eta}))$ is the Conditional Value-at-Risk functional given by

$$\mathrm{CVaR}_\epsilon(f(\boldsymbol{\eta})) = \inf_{\beta \in \mathbb{R}} \left\{\beta + \epsilon^{-1} \mathbb{E}_F[(f(\boldsymbol{\eta}) - \beta)^+]\right\}. \quad (18)$$

Lemma 2 ([15, Lemma A.1]). *Let $f : \mathbb{C}^N \to \mathbb{R}$ be a measurable function, and define the worst-case expectation θ_{wc} as*

$$\theta_{\mathrm{wc}} = \sup_{F \in \mathscr{D}(\boldsymbol{\mu}, \boldsymbol{\Sigma})} \mathbb{E}_F[(f(\boldsymbol{\eta}))^+].$$

Then,

$$\theta_{\mathrm{wc}} = \inf_{\boldsymbol{M} \in \mathbb{H}^{N+1}, \boldsymbol{M} \succeq \mathbf{0}} \mathrm{Tr}(\boldsymbol{\Omega} \boldsymbol{M})$$
$$\text{s.t. } [\boldsymbol{\eta}^H \ 1] \boldsymbol{M} [\boldsymbol{\eta}^H, \ 1]^H \geq f(\boldsymbol{\eta}), \ \forall \ \boldsymbol{\eta} \in \mathbb{C}^N,$$

where $\boldsymbol{\Omega} = \begin{bmatrix} \boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^H & \boldsymbol{\mu} \\ \boldsymbol{\mu}^H & 1 \end{bmatrix}$.

Equipped with the above two lemmas, we are now ready to present the proof. Specifically, by noticing that $L_i(\boldsymbol{h}_i, \{\boldsymbol{W}_l\}_{l \in \mathcal{K}})$ is quadratic in $\boldsymbol{h}_i$ [cf. (7)], we invoke Lemma 1 to get

$$(7) \iff \sup_{F_i \in \mathscr{D}(\bar{\boldsymbol{h}}_i + \boldsymbol{\xi}_i, \boldsymbol{\Sigma}_i)} \mathrm{CVaR}_{\epsilon_i}(L_i(\boldsymbol{h}_i, \{\boldsymbol{W}_l\}_{l \in \mathcal{K}})) \leq 0, \ \forall \ i \in \mathcal{K}. \quad (19)$$

According to the definition of $\mathrm{CVaR}_\epsilon(\cdot)$ in (18), we have

$$\sup_{F_i \in \mathscr{D}(\bar{\boldsymbol{h}}_i + \boldsymbol{\xi}_i, \boldsymbol{\Sigma}_i)} \mathrm{CVaR}_{\epsilon_i}(L_i(\boldsymbol{h}_i, \{\boldsymbol{W}_l\}_{l \in \mathcal{K}}))$$
$$= \sup_{F_i \in \mathscr{D}(\bar{\boldsymbol{h}}_i + \boldsymbol{\xi}_i, \boldsymbol{\Sigma}_i)} \inf_{\beta_i \in \mathbb{R}} \left\{\beta_i + \frac{1}{\epsilon_i} \mathbb{E}_{F_i}[(L_i(\boldsymbol{h}_i, \{\boldsymbol{W}_l\}_{l \in \mathcal{K}}) - \beta_i)^+]\right\}$$
$$= \inf_{\beta_i \in \mathbb{R}} \left\{\beta_i + \frac{1}{\epsilon_i} \sup_{F_i \in \mathscr{D}(\bar{\boldsymbol{h}}_i + \boldsymbol{\xi}_i, \boldsymbol{\Sigma}_i)} \mathbb{E}_{F_i}[(L_i(\boldsymbol{h}_i, \{\boldsymbol{W}_l\}_{l \in \mathcal{K}}) - \beta_i)^+]\right\}. \quad (20)$$

In the last equality of (20), we have interchanged the maximization and minimization operations, which can be justified by a stochastic saddle point theorem due to Shapiro and Kleywegt [22] (see also [15]). It then follows from Lemma 2 that the supremum in (20) is equal to

$$\inf_{\boldsymbol{M}_i \in \mathbb{H}^{N+1}, \boldsymbol{M}_i \succeq \mathbf{0}} \mathrm{Tr}(\boldsymbol{\Omega}_i \boldsymbol{M}_i) \quad (21a)$$

$$\text{s.t. } [\boldsymbol{h}_i^H \ 1] \boldsymbol{M}_i [\boldsymbol{h}_i^H, \ 1]^H \geq L_i(\boldsymbol{h}_i, \{\boldsymbol{W}_l\}_{l \in \mathcal{K}}), \ \forall \ \boldsymbol{h}_i \in \mathbb{C}^N, \quad (21b)$$

where $\boldsymbol{\Omega}_i$ is defined in (8). Since $L_i(\boldsymbol{h}_i, \{\boldsymbol{W}_l\}_{l \in \mathcal{K}})$ is quadratic in $\boldsymbol{h}_i$, it is not hard to see that (21b) holds if and only if

$$\boldsymbol{M}_i \succeq \begin{bmatrix} \sum_{k \neq i} \boldsymbol{W}_k - \gamma_i^{-1} \boldsymbol{W}_i & \mathbf{0} \\ \mathbf{0}^T & \sigma_i^2 - \beta_i \end{bmatrix}. \quad (22)$$

Finally, by replacing (21b) with (22) and making using of (19) and (20), we arrive at the desired the result in (8).

7. REFERENCES

[1] F. Rashid-Farrokhi, K. J. R. Liu, and L. Tassiulas, "Transmit beamforming and power control for cellular wireless systems," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1437–1450, Oct. 1999.

[2] M. Schubert and H. Boche, "Solution of the multiuser downlink beamforming problem with individual SINR constraints," *IEEE Trans. Veh. Technol.*, vol. 53, no. 1, pp. 18–28, Jan. 2004.

[3] M. Bengtsson and B. Ottersten, "Optimal and suboptimal transmit beamforming," Chapter 18 in *Handbook of Antennas in Wireless Communications, L. C. Godara, Ed., CRC Press*, Aug. 2001.

[4] Y. Huang and D. P. Palomar, "Rank-constrained separable semidefinite programming with applications to optimal beamforming," *IEEE Trans. Signal Process.*, vol. 58, no. 2, pp. 664–678, Feb. 2010.

[5] A. Wiesel, Y. C. Eldar, and S. Shamai, "Linear precoding via conic optimization for fixed MIMO receivers," *IEEE Trans. Signal Process.*, vol. 54, no. 1, pp. 161–176, Jan. 2006.

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

[7] M. B. Shenouda and T. N. Davidson, "Convex conic formulations of robust downlink precoder designs with quality of service constraints," *IEEE J. Sel. Topics Signal Process.*, vol. 1, pp. 714–724, Dec. 2007.

[8] G. Zheng, K.-K. Wong, and T.-S. Ng, "Robust linear MIMO in the downlink: A worst-case optimization with ellipsoidal uncertainty regions," *EURASIP J. Advances Signal Process.*, vol. 2008, pp. 1–15, Jan. 2008.

[9] E. Song, Q. Shi, M. Sanjabi, R. Sun, and Z.-Q. Luo, "Robust SINR-constrained MISO downlink beamforming: When is semidefinite programming relaxation tight?" in *Proc. IEEE ICASSP*, May 2011, pp. 3096–3099.

[10] N. Vučić and H. Boche, "A tractable method for chance-constrained power control in downlink multiuser MISO systems with channel uncertainty," *IEEE Signal Process. Lett.*, vol. 16, no. 5, pp. 346–349, Apr. 2009.

[11] M. B. Shenouda and T. N. Davidson, "Outage-based designs for multi-user transceivers," in *Proc. IEEE ICASSP*, Apr. 2009, pp. 2389–2392.

[12] K.-Y. Wang, A. M.-C. So, T.-H. Chang, W.-K. Ma, and C.-Y. Chi, "Outage constrained robust transmit optimization for multiuser MISO downlinks: Tractable approximations by conic optimization," 2011, available online at `http://arxiv.org/abs/1108.0982`.

[13] K.-Y. Wang, T.-H. Chang, W.-K. Ma, A. M.-C. So, and C.-Y. Chi, "Probabilistic SINR constrained robust transmit beamforming: A Bernstein-type inequality based conservative approach," in *Proc. IEEE ICASSP*, May 2011, pp. 3080–3083.

[14] Y. Rong, S. A. Vorobyov, and A. B. Gershman, "Robust linear receivers for multi-access space-time block coded MIMO systems: A probabilistically constrained approach," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 8, pp. 1560–1570, Aug. 2006.

[15] S. Zymler, D. Kuhn, and B. Rustem, "Distributionally robust joint chance constraints with second-order moment information," *Math. Program.*, vol. 137, pp. 167–198, Feb. 2013.

[16] E. Delage and Y. Ye, "Distributionally robust optimization under moment uncertainty with application to data-driven problems," *Oper. Res.*, pp. 1–18, Jan. 2010.

[17] A. M.-C. So, "Moment inequalities for sums of random matrices and their applications in optimization," *Math. Program.*, vol. 130, no. 1, pp. 125–151, 2011.

[18] Z.-Q. Luo, W.-K. Ma, A. M.-C. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 20–34, May 2010.

[19] A. M.-C. So and Y. J. Zhang, "Distributionally robust slow adaptive OFDMA with soft QoS via linear programming," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 5, pp. 947–958, 2013.

[20] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming," Apr. 2011, http://cvxr.com/cvx.

[21] J. Sturm, "Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones," *Optim. Methods Softw.*, vol. 11, pp. 625–653, 1999, (webpage and software) http://sedumi.ie.lehigh.edu/.

[22] A. Shapiro and A. J. Kleywegt, "Minimax analysis of stochastic problems," *Optim. Methods Softw.*, vol. 17, no. 3, pp. 523–542, 2002.