A SIMPLE WAY TO APPROXIMATE AVERAGE ROBUST MULTIUSER MISO TRANSMIT OPTIMIZATION UNDER COVARIANCE-BASED CSIT

Mingjie Shao and Wing-Kin Ma

Department of Electronic Engineering, The Chinese University of Hong Kong, Hong Kong S.A.R. of China

ABSTRACT

This paper focuses on an average robust transmit beamforming optimization problem for the multiuser multiple-input-single-output (MISO) downlink scenario. In this problem, the channels are modeled as Gaussian variables with mean zero and with known covariance at the transmitter. The design criterion is to maximize the sum of the users' average rates with respect to the channels, subject to the total transmission power constraint. The challenge of this problem is that the average rate function generally admits a complex expression. Such an issue can be tackled by stochastic approximation (SA) approaches, but SA may require a large number of samples, or iterations, to provide reasonable performance. In this work, a simple deterministic approximation scheme is proposed. First, we propose a closed-form surrogate of the per-user average rate function. The proposed surrogate function is shown to have an approximation accuracy within 0.8314 bits from the true average rate. Then, we utilize the proposed surrogate function to establish an algebraically simple alternating optimization algorithm for the beamforming problem. Simulation results show that the proposed algorithm is computationally much more efficient than an SA-based state-of-the-art algorithm when they are compared under similar sum rate performance levels.

***Index Terms*—** robust transmit beamforming, multiuser MIMO

1. INTRODUCTION

It has been known for more than a decade that multiuser multi-input multi-output (MIMO) transceiver techniques such as transmit beamforming can lead to tremendous improvement of the spectral efficiency of a wireless network [1]. In recent frontier developments such as massive MIMO and full-duplex MIMO, transmit beamforming and related concepts are still seen to play an eminent role. When one performs transmit beamforming design, it is common to assume that channel state information at the transmitter (CSIT) is available in a perfect and instantaneous manner. However, practical systems rarely satisfy such an assumption because of issues such as inaccurate channel estimation, limited channel feedback, outdated CSIT effects, to name a few [2,3]. Thus, a classic and fundamentally important problem is to study how we can robustify the beamforming designs when only imperfect or partially-informed CSIT is available—that is, robust transmit beamforming.

The robust beamforming topic has drawn much attention in the community, and we have seen a plethora of designs based on various robust quality-of-service (QoS) performance metrics and various imperfect CSIT settings; see the literature [4–10] and the references therein for details. In this paper our focus is on an average robust design for the multiuser multiple-input single-output (MISO) downlink scenario. We adopt a statistical CSIT model where the channels are modeled as zero-mean Gaussian random variables, and the base station (BS) knows only the covariances of the channels. This model may be used in situations where the BS is unable to acquire CSIT in an instantaneous sense, but can estimate the second-order statistics of the channels (e.g., via uplink signal measurements) [11–13]. Then, we choose the per-user average rate with respect to the channel as the robust QoS performance metric, and seek to maximize the sum of the per-user average rates subject to the total transmission power constraint. This average robust sum-rate maximization problem is difficult to deal with because the per-user average rate function generally does not exhibit a desirable structure for us to perform optimization. The average robust problem was tackled in [8] via a stochastic successive upper bound approximation (SSUM) framework; we should also mention the stochastic parallel decomposition framework [14] which was used to deal with a very similar average robust problem. These two frameworks were essentially developed for a wider class of problems, namely, stochastic optimization, and they are theoretically sound in the sense that they guarantee convergence to a stationary point under some mild assumptions. The drawback of these frameworks is that they rely on stochastic approximation, which uses randomly generated samples to handle the average rate function and may require a large amount of samples to perform well.

The contribution of this work lies in exploiting the underlying problem structures to establish a simple and deterministic scheme for approximating the average robust sum-rate maximization problem. We will first propose a simple surrogate function for the per-user average rate function. The approximation accuracy of the proposed surrogate function will be analyzed. Then, we will use our surrogate function to establish an alternating optimization algorithm that shows a similar flavor, and also efficiency, as a well-known beamforming design algorithm in the perfect CSIT case [15]. Simulation results will show that our approach outperforms the SSUM approach in terms of the number of iterations required to achieve good performance. We should highlight that our development also leads to an interesting connection between the so-called estimated signal-to-interference-and-noise ratio (SINR) performance metric [11] and the average robust design, which was not reported in previous work.

2. PROBLEM FORMULATION AND BACKGROUND

Consider a single-cell multiuser MISO downlink scenario where a base station (BS) employs linear beamforming to transmit data streams to users, with one stream per user, in a simultaneous fashion and over the same frequency band. We follow the standard signal model in this context, and readers are referred to the literature [11, 16, 17] for a complete exposition of the model. Let $\boldsymbol{h}_i \in \mathbb{C}^N$ and $\boldsymbol{w}_i \in \mathbb{C}^N$ denote the channel from the BS to user i

and the beamformer for user i, respectively, where N is the number of transmit antennas at the BS. Also, let K be the number of users. Given information of the channels $\boldsymbol{h}_i$'s at the BS, i.e., perfect CSIT, the achievable rate of user i can be characterized as

$$\mathsf{R}_i(\boldsymbol{W}, \boldsymbol{h}_i) = \log(1 + \mathsf{SINR}_i(\boldsymbol{W}, \boldsymbol{h}_i)),$$

where $\boldsymbol{W} = [\, \boldsymbol{w}_1, \ldots, \boldsymbol{w}_K \,]$,

$$\mathsf{SINR}_i(\boldsymbol{W}, \boldsymbol{h}_i) = \frac{|\boldsymbol{h}_i^H \boldsymbol{w}_i|^2}{\sum_{j \neq i} |\boldsymbol{h}_i^H \boldsymbol{w}_j|^2 + \sigma_i^2}$$

is the SINR of user i, and $\sigma_i^2 > 0$ is the noise power corresponding to user i.

We are interested in an imperfect CSIT setting where every channel $\boldsymbol{h}_i$ is modeled as a random variable, and the BS has information of the channel distributions only. The problem is to design the beamformers under channel distribution information. Particularly, the following average robust sum-rate maximization (SRM) problem will be our focus:

$$\begin{aligned} \max_{\boldsymbol{W}} \; & \textstyle\sum_{i=1}^K \mathbb{E}_{\boldsymbol{h}_i \sim \mathcal{D}_i}[\mathsf{R}_i(\boldsymbol{W}, \boldsymbol{h}_i)] \\ \text{s.t.} \; & \textstyle\sum_{i=1}^K \|\boldsymbol{w}_i\|_2^2 \leq P, \end{aligned} \tag{1}$$

where $\mathcal{D}_i$ denotes the distribution of $\boldsymbol{h}_i$, and P is the maximum total transmission power of the BS (note that $\boldsymbol{W} \in \mathbb{C}^{N \times K}$, and we make it implicit in problem (1) for brevity). It should be mentioned that problem (1) uses the average rate function $\mathbb{E}_{\boldsymbol{h}_i \sim \mathcal{D}_i}[\mathsf{R}_i(\boldsymbol{W}, \boldsymbol{h}_i)]$ as the robust QoS performance metric, and thus it is an average robust design.

Problem (1) is an instance of the stochastic optimization problem class, which is known to be challenging to solve in general. For the case of problem (1), the challenge lies in the fact that the average rate function $\mathbb{E}_{\boldsymbol{h}_i \sim \mathcal{D}_i}[\mathsf{R}_i(\boldsymbol{W}, \boldsymbol{h}_i)]$ does not admit an explicit expression in general. To tackle the aforementioned challenge, one direction is to apply a stochastic approximation (SA) approach [8, 14]. Roughly speaking, an SA algorithm works in the following way: randomly generate samples from the channel distributions $\mathcal{D}_i$'s, use those samples to perform some kind of incremental update on $\boldsymbol{W}$, and repeat such an update until a stopping rule is satisfied. As discussed in the Introduction, some sophisticated SA algorithms such as [8,14] can guarantee convergence to a stationary point under fairly mild assumptions on the channel distributions $\mathcal{D}_i$'s. However, by the sample approximation nature of SA, one would expect that an SA algorithm would require many samples, or iterations, to provide reasonable performance; this will be shown to be true by simulation results in Section 5.

3. A SIMPLE APPROXIMATION OF THE AVERAGE RATES

Our approach for tackling the average robust SRM problem in (1) is to develop a simple approximation, or surrogate, of the average rate function $\mathbb{E}_{\boldsymbol{h}_i \sim \mathcal{D}_i}[\mathsf{R}_i(\boldsymbol{W}, \boldsymbol{h}_i)]$. Let us assume the following channel distribution model:

$$\mathcal{D}_i = \mathcal{CN}(\boldsymbol{0}, \boldsymbol{R}_i), \quad i = 1, \ldots, K, \tag{2}$$

where every covariance $\boldsymbol{R}_i \succeq \boldsymbol{0}$ is known to the BS (the notation $\mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{C})$ stands for a circular complex Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{C}$, and $\boldsymbol{X} \succeq \boldsymbol{0}$ means that $\boldsymbol{X}$ is positive semidefinite (PSD)). The above model is commonly used in situations where the BS is unable to acquire the channels $\boldsymbol{h}_i$'s in an instantaneous sense, but can estimate the second-order statistics of the channels, e.g., by uplink signal measurement [11–13]. In the above model, an explicit expression of the average rate function can actually be derived [18, Theorem 1]. However, the expression is too complex and may not be easy to handle from an optimization viewpoint. Our proposed surrogate of the average rate function is as follows:

$$\bar{\mathsf{R}}_i(\boldsymbol{W}) = \log\left(1 + \frac{\boldsymbol{w}_i^H \boldsymbol{R}_i \boldsymbol{w}_i}{\sum_{j \neq i} \boldsymbol{w}_j^H \boldsymbol{R}_i \boldsymbol{w}_j + \sigma_i^2}\right). \tag{3}$$

As can be seen, Eq. (3) admits a simple expression. In the next section, we will leverage on the simplicity of (3) to establish an efficient algorithm for problem (1). For now, let us consider this question: does (3) have any guarantee in approximation accuracy? We answer this question in the following proposition:

Proposition 1 *Consider the average rate surrogate function in* (3) *under the statistical CSIT model in* (2)*. It holds true that*

$$\left|\mathbb{E}_{\boldsymbol{h}_i \sim \mathcal{D}_i}[\mathsf{R}_i(\boldsymbol{W}, \boldsymbol{h}_i)] - \bar{\mathsf{R}}_i(\boldsymbol{W})\right| \leq \gamma, \quad \textit{for any } \boldsymbol{W},$$

where $\gamma \simeq 0.5772$ is the Euler constant.

Proposition 1 reveals that the surrogate function (3) is non-heuristic in the sense that its rate approximation error is no worse than $\gamma / \log(2) = 0.8314$ bits.

Proof of Proposition 1: First, note that $\mathsf{R}_i(\boldsymbol{W}, \boldsymbol{h}_i)$ can be decomposed as

$$\mathsf{R}_i(\boldsymbol{W}, \boldsymbol{h}_i) = \log\left(1 + \frac{\sum_j |\boldsymbol{h}_i^H \boldsymbol{w}_j|^2}{\sigma_i^2}\right) - \log\left(1 + \frac{\sum_{j \neq i} |\boldsymbol{h}_i^H \boldsymbol{w}_j|^2}{\sigma_i^2}\right). \tag{4}$$

Likewise, we can write

$$\bar{\mathsf{R}}_i(\boldsymbol{W}) = \log\left(1 + \frac{\sum_j \boldsymbol{w}_j^H \boldsymbol{R}_i \boldsymbol{w}_j}{\sigma_i^2}\right) - \log\left(1 + \frac{\sum_{j \neq i} \boldsymbol{w}_j^H \boldsymbol{R}_i \boldsymbol{w}_j}{\sigma_i^2}\right). \tag{5}$$

Second, we prove a lower bound on $\mathbb{E}_{\boldsymbol{h}_i \sim \mathcal{D}_i}[\mathsf{R}_i(\boldsymbol{W}, \boldsymbol{h}_i)]$ by considering the two terms on the right-hand side of (4). By Jensen's inequality, we have

$$\begin{aligned} & \mathbb{E}_{\boldsymbol{h}_i \sim \mathcal{D}_i}\left[\log\left(1 + \frac{\sum_{j \neq i} |\boldsymbol{h}_i^H \boldsymbol{w}_j|^2}{\sigma_i^2}\right)\right] \\ & \leq \log\left(1 + \frac{\mathbb{E}_{\boldsymbol{h}_i \sim \mathcal{D}_i}\left[\sum_{j \neq i} |\boldsymbol{h}_i^H \boldsymbol{w}_j|^2\right]}{\sigma_i^2}\right) \\ & = \log\left(1 + \frac{\sum_{j \neq i} \boldsymbol{w}_j^H \boldsymbol{R}_i \boldsymbol{w}_j}{\sigma_i^2}\right) \end{aligned} \tag{6}$$

Also, consider the following lemma:

Lemma 1 *Let $\boldsymbol{\Phi} \in \mathbb{C}^{n \times n}$ be any Hermitian PSD matrix. We have*

$$\mathbb{E}_{\boldsymbol{\xi} \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{I})}[\log(1 + \boldsymbol{\xi}^H \boldsymbol{\Phi} \boldsymbol{\xi})] \geq \log(1 + \mathrm{Tr}(\boldsymbol{\Phi})) - \gamma,$$

where γ is again the Euler constant.

Lemma 1 is a consequence of stochastic majorization [19] and certain integral results [20]; the proof is omitted here owing to the limit of space. Using Lemma 1, we have

$$\begin{aligned}
&\mathbb{E}_{\boldsymbol{h}_i\sim\mathcal{D}_i}\left[\log\left(1+\frac{\sum_j|\boldsymbol{h}_i^H\boldsymbol{w}_j|^2}{\sigma_i^2}\right)\right]\\
=&\mathbb{E}_{\boldsymbol{\xi}_i\sim\mathcal{CN}(\boldsymbol{0},\boldsymbol{I})}\left[\log\left(1+\frac{\boldsymbol{\xi}^H\boldsymbol{R}_i^{1/2}\left(\sum_j\boldsymbol{w}_j\boldsymbol{w}_j^H\right)\boldsymbol{R}_i^{1/2}\boldsymbol{\xi}}{\sigma_i^2}\right)\right]\\
\geq&\log\left(1+\frac{\mathrm{Tr}\left(\boldsymbol{R}_i^{1/2}\left(\sum_j\boldsymbol{w}_j\boldsymbol{w}_j^H\right)\boldsymbol{R}_i^{1/2}\right)}{\sigma_i^2}\right)-\gamma\\
=&\log\left(1+\frac{\sum_j\boldsymbol{w}_j^H\boldsymbol{R}_i\boldsymbol{w}_j}{\sigma_i^2}\right)-\gamma, \qquad (7)
\end{aligned}$$

where $\boldsymbol{R}_i^{1/2}$ denotes the PSD square root of $\boldsymbol{R}_i$. By taking expectation on (4) with respect to $\boldsymbol{h}_i$, and applying (6)–(7) and (5), we obtain a lower bound

$$\mathbb{E}_{\boldsymbol{h}_i\sim\mathcal{D}_i}[\mathsf{R}_i(\boldsymbol{W},\boldsymbol{h}_i)]\geq\bar{\mathsf{R}}_i(\boldsymbol{W})-\gamma.$$

Moreover, by following the same proof as above, one can also show that

$$\mathbb{E}_{\boldsymbol{h}_i\sim\mathcal{D}_i}[\mathsf{R}_i(\boldsymbol{W},\boldsymbol{h}_i)]\leq\bar{\mathsf{R}}_i(\boldsymbol{W})+\gamma.$$

The two above bounds complete the proof. ■

A remark is as follows.

Remark 1 In the transmit beamforming context, there is a widely-used QoS performance metric called the *estimated SINR* [11]. The estimated SINR is defined as

$$\begin{aligned}
\overline{\mathsf{SINR}}_i(\boldsymbol{W})&=\frac{\mathbb{E}_{\boldsymbol{h}_i\sim\mathcal{D}_i}[|\boldsymbol{h}_i^H\boldsymbol{w}_i|^2]}{\mathbb{E}_{\boldsymbol{h}_i\sim\mathcal{D}_i}[\sum_{j\neq i}|\boldsymbol{h}_i^H\boldsymbol{w}_j|^2]+\sigma_i^2}\\
&=\frac{\boldsymbol{w}_i^H\boldsymbol{R}_i\boldsymbol{w}_i}{\sum_{j\neq i}\boldsymbol{w}_j^H\boldsymbol{R}_i\boldsymbol{w}_j+\sigma_i^2},
\end{aligned}$$

which is the ratio of the average signal power to the average interference-plus-noise power. It was introduced as a way to evaluate the (instantaneous) SINR approximately when perfect CSIT is unavailable. The use of the estimated SINR is particularly common when one considers the SINR-constrained formulation, that is,

$$\begin{aligned}
\min_{\boldsymbol{W}}\ &\textstyle\sum_{i=1}^K\|\boldsymbol{w}_i\|_2^2\\
\text{s.t. }&\overline{\mathsf{SINR}}_i(\boldsymbol{W})\geq\beta_i,\quad i=1,\ldots,K,
\end{aligned}\qquad (8)$$

where β_i's are the minimum QoS requirements; see [11, 16, 21] and many subsequent works that follow the above references. While the estimated SINR is a reasonable QoS measure intuitively, it is unclear in the existing literature how and in what way the estimated SINR is linked with achievable rates. Interestingly, our result provides such a connection: One can see that our rate surrogate in (3) is simply

$$\bar{\mathsf{R}}_i(\boldsymbol{W})=\log(1+\overline{\mathsf{SINR}}_i(\boldsymbol{W})).$$

Thus, under the statistical CSIT model in (2), the estimated SINR leads to a non-heuristic approximation of the average rate function $\mathbb{E}_{\boldsymbol{h}_i\sim\mathcal{D}_i}[\mathsf{R}_i(\boldsymbol{W},\boldsymbol{h}_i)]$ — an interpretation not seen in the literature. Also, for the SINR-constrained design in (8), we have the following implication by Proposition 1:

$$\overline{\mathsf{SINR}}_i(\boldsymbol{W})\geq\beta_i\quad\Longrightarrow\quad\mathbb{E}_{\boldsymbol{h}_i\sim\mathcal{D}_i}[\mathsf{R}_i(\boldsymbol{W},\boldsymbol{h}_i)]\geq\log(1+\beta_i)-\gamma.$$

4. THE PROPOSED OPTIMIZATION ALGORITHM

Having proposed a simple average rate approximation in the last section, we are now ready to establish a new algorithm for the average robust SRM problem in (1). Specifically, we apply the approximation in (3) and consider the following approximate version of problem (1):

$$\begin{aligned}
\max_{\boldsymbol{W}}\ &\textstyle\sum_{i=1}^K\bar{\mathsf{R}}_i(\boldsymbol{W})\\
\text{s.t. }&\textstyle\sum_{i=1}^K\|\boldsymbol{w}_i\|_2^2\leq P.
\end{aligned}\qquad (9)$$

We tackle problem (9) via the alternating optimization approach in [15], which was previously proposed for the perfect-CSIT SRM problem and is commonly known as the weighted minimum-mean-square-error (WMMSE) approach. Our development is as follows. Let

$$g_i(\boldsymbol{W},\boldsymbol{u}_i)=\left(\textstyle\sum_j\boldsymbol{w}_j^H\boldsymbol{R}_i\boldsymbol{w}_j+\sigma_i^2\right)\|\boldsymbol{u}_i\|_2^2-2\mathrm{Re}[\boldsymbol{u}_i^H\boldsymbol{R}_i^{1/2}\boldsymbol{w}_i]+1.$$

It can be easily verified that

$$\min_{\boldsymbol{u}_i\in\mathbb{C}^N}g_i(\boldsymbol{W},\boldsymbol{u}_i)=\frac{1}{1+\frac{\boldsymbol{w}_i^H\boldsymbol{R}_i\boldsymbol{w}_i}{\sum_{j\neq i}\boldsymbol{w}_j^H\boldsymbol{R}_i\boldsymbol{w}_j+\sigma_i^2}},\qquad (10)$$

and the minimum is attained when

$$\boldsymbol{u}_i=\frac{1}{\sum_j\boldsymbol{w}_j^H\boldsymbol{R}_i\boldsymbol{w}_j+\sigma_i^2}\boldsymbol{R}_i^{1/2}\boldsymbol{w}_i.\qquad (11)$$

Also, consider the following standard result: For any $x>0$, we have

$$\log(x)=\min_{\theta\geq0}\ x\theta-\log(\theta)-1,\qquad (12)$$

and the minimum is attained when $\theta=1/x$. Using (10) and (12), one can show that problem (9) is equivalent to

$$\begin{aligned}
\min_{\boldsymbol{W},\boldsymbol{U},\boldsymbol{\theta}}\ &\textstyle\sum_{i=1}^K\theta_ig_i(\boldsymbol{W},\boldsymbol{u}_i)-\log(\theta_i)\\
\text{s.t. }&\boldsymbol{\theta}\geq\boldsymbol{0},\ \textstyle\sum_{i=1}^K\|\boldsymbol{w}_i\|_2^2\leq P,
\end{aligned}\qquad (13)$$

where $\boldsymbol{U}=[\ \boldsymbol{u}_1,\ldots,\boldsymbol{u}_K\ ]\in\mathbb{C}^{N\times K}$. The equivalent formulation in (13) allows us to handle the problem conveniently via alternating optimization. Specifically, we optimize $\boldsymbol{W}$, $\boldsymbol{U}$ and $\boldsymbol{\theta}$ in an alternating or cyclic fashion. The optimization of $\boldsymbol{U}$ given $\boldsymbol{W},\boldsymbol{\theta}$ has a closed form and is given in (11). The optimization of $\boldsymbol{\theta}$ given $\boldsymbol{W},\boldsymbol{U}$ also has a closed form and is given by $\theta_i=1/g_i(\boldsymbol{W},\boldsymbol{u}_i)$ for all i. The optimization of $\boldsymbol{W}$ given $\boldsymbol{U},\boldsymbol{\theta}$ is a quadratic program subject to one quadratic constraint, and it can be efficiently solved using a water-filling algorithm [15]. The pseudo-code of the resulting algorithm is shown in Algorithm 1.

Algorithm 1 : Alternating Optimization Algorithm for Problem (9)

1: **given** a feasible $\boldsymbol{W}$
2: **repeat**
3: $\quad\boldsymbol{u}_i\leftarrow\frac{1}{\sum_j\boldsymbol{w}_j^H\boldsymbol{R}_i\boldsymbol{w}_j+\sigma_i^2}\boldsymbol{R}_i^{1/2}\boldsymbol{w}_i$, $i=1,\ldots,K$;
4: $\quad\theta_i\leftarrow\left(1-\boldsymbol{u}_i^H\boldsymbol{R}_i^{1/2}\boldsymbol{w}_i\right)^{-1}$, $i=1,\ldots,K$;
5: $\quad\boldsymbol{w}_i\leftarrow\left(\sum_{j=1}^K\theta_j\boldsymbol{u}_j^H\boldsymbol{u}_j\boldsymbol{R}_j+\lambda\boldsymbol{I}\right)^{-1}\boldsymbol{R}_i^{1/2}\boldsymbol{u}_i\theta_i$, $i=1,\ldots,K$,
$\quad$ where λ is chosen such that the above $\boldsymbol{w}_i$'s satisfy $\sum_{i=1}^K\|\boldsymbol{w}_i\|_2^2=P$; see [15] for an algorithm of doing so.
6: **until** a stopping rule is satisfied.

Before we finish, we should give a few remarks. First, and as a rather subtle point, our algorithm is a minimum mean square error (MMSE)-free version of the original WMMSE approach [15]. The original WMMSE approach relies on a notion that utilizes an equivalent relationship between MMSE and SINR. Such an MMSE notion is specifically for the perfect CSIT case, and it does not apply to our problem due to the inherent problem structures in (9) (specifically, the estimated SINR structures). However, we are able to derive a replacement of the MMSE notion, which is (10) and uses simple algebraic concepts. Thus, our algorithm has nothing to do with MMSE and only adopts the insight of alternating optimization in the WMMSE approach. Second, the technique shown above can be easily extended to other formulations such as the multicell scenario, proportionally fair utility maximization [15] and user grouping [22]; such extensions are easy and straightforward, and the details are omitted here in view of space limitation.

5. NUMERICAL RESULTS

In this section, simulations are used to show the performance of the proposed algorithm in Algorithm 1. We also benchmark our algorithm against an existing SA-based algorithm, specifically, the stochastic WMMSE algorithm which is the application of the SSUM framework [8].

The simulation settings are as follows. The number of transmit antennas is $N = 8$, the number of users is $K = 4$, and equal noise power $\sigma_1^2 = \ldots = \sigma_K^2 = \sigma^2$ is assumed. We model the channel covariances by a similar way as the specular model in [23]. Specifically, we have

$$\boldsymbol{R}_i = \sum_{j=1}^{L} \alpha_{i,j} \boldsymbol{a}(\varphi_{i,j}) \boldsymbol{a}(\varphi_{i,j})^H,$$

for some L, $\{\varphi_{i,j}\}$, $\{\alpha_{i,j}\}$, where $\alpha_{ij} > 0$;

$$\boldsymbol{a}(\varphi) = \left[1, e^{-j2\pi \frac{d\cos(\varphi)}{\lambda}}, \ldots, e^{-j2\pi \frac{(N-1)d\cos(\varphi)}{\lambda}}\right]^T$$

is a steering vector with angle φ, wavelength λ and antenna spacing d. We set $d = 0.5\lambda$, $L = 3$, $\alpha_{i,j} = 1/L$ for all i, j. In each simulation trial, we first randomly generate a set of mean angles $\bar{\varphi}_i$'s via a uniform distribution on $[0, 2\pi]$. Then, the angles $\varphi_{i,j}$ are randomly generated following a uniform distribution on $\left[\bar{\varphi}_i - \frac{\pi}{3}, \bar{\varphi}_i + \frac{\pi}{3}\right]$.

Fig. 1 compares the convergence rates of the proposed algorithm and the stochastic WMMSE algorithm. It is a one-trial result, and we set $P/\sigma^2 = 20$ dB. Note that the average sum rate $\sum_i \mathbb{E}_{\boldsymbol{h}_i \sim \mathcal{D}_i}[\mathsf{R}_i(\boldsymbol{W}, \boldsymbol{h}_i)]$ in the figure are obtained by Monte-Carlo averaging over 1,000 independent channel realizations. Also, "Surrogate" refers to the approximate average sum rate value $\sum_i \bar{\mathsf{R}}_i(\boldsymbol{W})$ of our algorithm. We see that the convergence speed of our algorithm is much faster than that of stochastic WMMSE. Furthermore, the worst-case gap between the approximate and true average sum rates of our algorithm is about 1.87 bits in this example, which is smaller than $4 \times \gamma \approx 3.3256$ bits and agrees with our analysis in Proposition 1.

Fig. 2 shows the average sum rate performance of the two algorithms over 500 simulation trials. We terminate the algorithm when its number of iterations exceeds a certain value, shown as "iter" in the legend. We observe from Fig. 2 that the stochastic WMMSE algorithm requires $2,000$ iterations to approach the performance of our algorithm with 20 iterations. Also, notice that further increasing the iteration number does not lead to significant performance improvement for stochastic WMMSE by our empirical experience. This suggests that the proposed algorithm is much more efficient than the stochastic WMMSE algorithm.

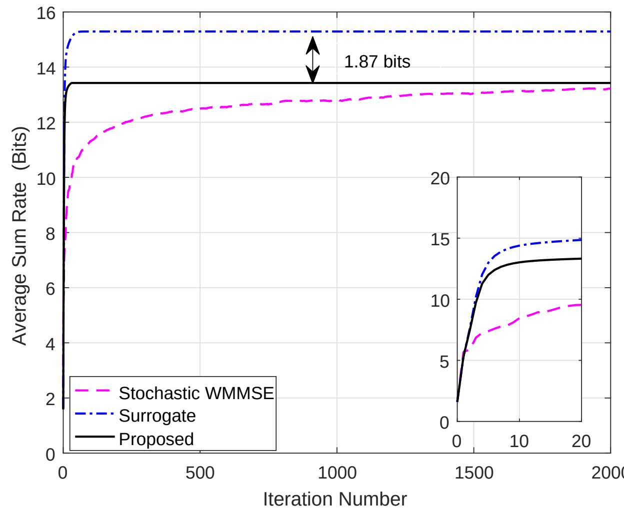

Fig. 1. Average sum rate versus iteration number.

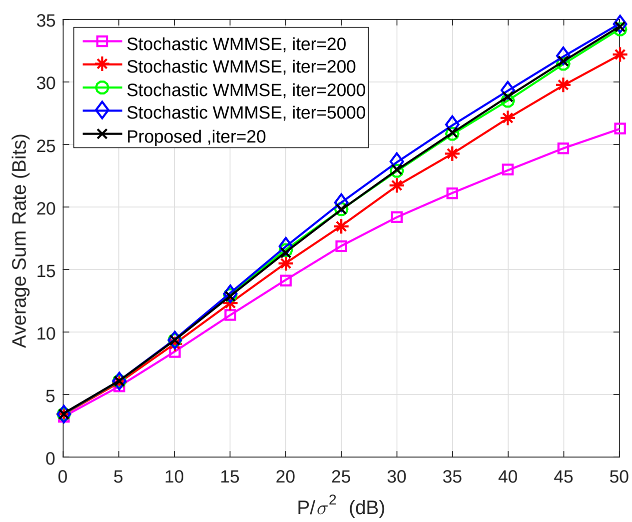

Fig. 2. Average sum rate performance versus P/σ^2.

6. CONCLUSION AND DISCUSSION

In this paper, we have proposed a simple low-complexity approximation scheme for the average robust sum-rate maximization problem for multiuser MISO downlink and under a statistical covariance-based CSIT model. As a future direction, it would be interesting to extend this work to the MIMO case and to the nonzero channel mean case.

7. REFERENCES

[1] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 5, pp. 684–702, 2003.

[2] G. Caire, N. Jindal, M. Kobayashi, and N. Ravindran, "Multiuser MIMO achievable rates with downlink training and channel state feedback," *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2845–2866, Jun. 2010.

[3] N. Jindal, "MIMO broadcast channels with finite-rate feedback," *IEEE Trans. Inf. Theory*, vol. 52, no. 11, pp. 5045–5060, 2006.

[4] M. B. Shenouda and T. N. Davidson, "Convex conic formulations of robust downlink precoder designs with quality of service constraints," *IEEE J. Sel. Top. Signal Process.*, vol. 1, no. 4, pp. 714–724, 2007.

[5] G. Zheng, K.-K. Wong, and B. Ottersten, "Robust cognitive beamforming with bounded channel uncertainties," *IEEE Trans. Signal Process.*, vol. 57, no. 12, pp. 4871–4881, 2009.

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

[7] M. B. Shenouda, T. N. Davidson, and L. Lampe, "Outage-based design of robust Tomlinson–Harashima transceivers for the MISO downlink with QoS requirements," *Signal Process.*, vol. 93, no. 12, pp. 3341–3352, 2013.

[8] M. Razaviyayn, M. Sanjabi, and Z.-Q. Luo, "A stochastic successive minimization method for nonsmooth nonconvex optimization with applications to transceiver design in wireless communication networks," *Mathematical Programming*, vol. 157, no. 2, pp. 515–545, 2016.

[9] 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," *IEEE Trans. Signal Process.*, vol. 62, no. 21, pp. 5690–5705, 2014.

[10] W.-K. Ma, J. Pan, A. M.-C. So, and T.-H. Chang, "Unraveling the rank-one solution mystery of robust miso downlink transmit optimization: A verifiable sufficient condition via a new duality result," to apprear in *IEEE Trans. Signal Process.*, 2017, available online at https://arxiv.org/abs/1602.01569.

[11] 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.

[12] M. Vu and A. Paulraj, "A robust transmit CSI framework with applications in MIMO wireless precoding," in *Conf. Record 39th Asilomar Conf. Signals, Systems, Computers*, 2005, pp. 623–627.

[13] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*. Cambridge University Press, 2003.

[14] Y. Yang, G. Scutari, D. P. Palomar, and M. Pesavento, "A parallel decomposition method for nonconvex stochastic multi-agent optimization problems," *IEEE Trans. Signal Process.*, vol. 64, no. 11, pp. 2949–2964, 2015.

[15] Q. Shi, M. Razaviyayn, Z. Q. Luo, and C. He, "An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4331–4340, Sep. 2011.

[16] A. B. Gershman, N. D. Sidiropoulos, S. Shahbazpanahi, M. Bengtsson, and B. Ottersten, "Convex optimization-based beamforming: From receive to transmit and network designs," *IEEE Signal Process. Mag*, vol. 27, no. 3, pp. 62–75, 2010.

[17] E. Björnson, M. Bengtsson, and B. Ottersten, "Optimal multiuser transmit beamforming: A difficult problem with a simple solution structure," *IEEE Signal. Proc. Mag.*, vol. 31, no. 4, pp. 142–148, 2014.

[18] E. Bjornson, R. Zakhour, D. Gesbert, and B. Ottersten, "Cooperative multicell precoding: Rate region characterization and distributed strategies with instantaneous and statistical CSI," *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 4298–4310, 2010.

[19] E. Jorswieck and H. Boche, *Majorization and Matrix-Monotone Functions in Wireless Communications*. Hanover, MA, USA: Now Publishers Inc., 2007.

[20] S. X. Wu, W.-K. Ma, and A. M.-C. So, "Physical-layer multicasting by stochastic transmit beamforming and Alamouti space-time coding," *IEEE Trans. Signal Process.*, vol. 61, no. 17, pp. 4230–4245, Sep. 2013, also the companion technical report at https://arxiv.org/pdf/1305.1427.pdf.

[21] 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, 2004.

[22] M. Razaviyayn, M. Baligh, A. Callard, and Z.-Q. Luo, "Joint user grouping and transceiver design in a MIMO interfering broadcast channel," *IEEE Trans. Signal Process.*, vol. 62, no. 1, pp. 85–94, 2014.

[23] D. Gesbert, L. Pittman, and M. Kountouris, "Transmit correlation-aided scheduling in multiuser MIMO networks," in *Proc. Int. Conf. Acous., Speech, Signal Process.*, vol. 4. IEEE, 2006, pp. 249–252.