Robust Downlink Beamforming for BDMA Massive MIMO System

Fengchao Zhu, Feifei Gao, *Senior Member, IEEE*, Shi Jin, *Member, IEEE*,
Hai Lin, *Senior Member, IEEE*, and Minli Yao

***Abstract*— In this paper, we design robust downlink beamforming against the imperfect channel state information (CSI) for beam division multiple access (BDMA) massive multiple-input multiple output (MIMO) systems. Following a worst-case deterministic model, the proposed design is formulated as minimizing the power consumption of base station (BS) under different signal-to-interference-plus-noise ratio (SINR) constraints. The S-Procedure and semi-definite relaxation (SDR) are used to convert the initial non-convex optimization to a convex semi-definite programming problem. Then the optimality of SDR is strictly proved by showing the rank-one property of the optimal beamforming thanks to the orthogonal channels under BDMA scheme. More importantly, we make one step further by deriving the optimal beamforming directions and optimal beamforming power allocation of the SDR in closed-form, which greatly reduces the optimization complexity and makes the proposed design practical for a real word massive MIMO system. Simulation results are then provided to verify the efficiency of the proposed robust beamforming algorithm.**

***Index Terms*— Robust beamforming, massive MIMO, semi-definite programming (SDP), closed-form solutions, beam division multiple access (BDMA).**

I. INTRODUCTION

MASSIVE multiple-input multiple-output (MIMO) systems [1] are currently hot topics of 5G wireless communications that demand for high spectral utilization to support the quality of service (QoS) for a wide variety of multimedia applications. Because of its potential to provide low power consumption, high spectral efficiency, security, and reliable linkage, massive MIMO has already been marked as a key technology under EURASIP METIS-2020 project and under China IMT-2020 project [2]. The basic premise behind massive MIMO is that when the number of antennas at the base station (BS) grows to infinity [3], the propagation channel vectors of different users become asymptotically orthogonal, which eliminates the effects of uncorrelated noise and intra-cell interference [4]. In this case, simple linear signal processing approaches, such as matched filter (MF) receiver and beamforming precoder become optimal in massive MIMO systems.

All these potential advantages of massive MIMO system rely heavily on the perfect channel state information (CSI) at the base station (BS) [5]. However, erroneous CSI is unavoidable in practice due to the finite length of training signal [6], [7] and/or limited feedback bandwidth [8]. The key difficulty for massive MIMO channel estimation lies in the requirement that the length of the training signal has to be no less than the number of antennas as well as its accompanied high computational complexity. A lot of prominent researches have been implemented to reduce the computational complexity and make channel estimation feasible in massive MIMO systems. By exploiting the low-rank approximation of channel covariance matrix, Yin *et al.* [9] and Adhikary *et al.* [10] utilized non-overlapping spatial information of different users to reduce the estimated parameters. Based on the assumption that the angular spread (AS) of the incident signals at BS from each user is narrow, the authors in [11]–[13] proposed an angle domain spatial basis expansion model (SBEM) that could represent the corresponding channels with only a few parameters. A technique called beam division multiple access (BDMA) was presented in [14] and [15], where it is shown that the channel after discrete Fourier transform (DFT) is approximately sparse, and then assign different users with orthogonal beams for transmission. Recently, the BDMA technique has been generalized to millimeter wave communication systems [16], [17], where the results say that the beam domain channel fading in time and frequency disappears asymptotically. Note that the basic idea of the above mentioned channel estimation algorithms [9]–[17] is to approximate massive MIMO channels by various orthogonal bases, by which means the effect parameters can be reduced.

On the other side, robust beamforming has a long history in signal processing community [18], [19], whose

Manuscript received April 10, 2017; revised August 4, 2017, November 19, 2017, and December 16, 2017; accepted December 16, 2017. Date of publication December 27, 2017; date of current version April 16, 2018. The work of F. Zhu was supported by the National Natural Science Foundation of China under Grant 61601474; The work of F. Gao was supported in part by the National Natural Science Foundation of China under Grant 61422109, 61771274. The work of S. Jin was supported in part by the National Science Foundation for Distinguished Young Scholars of China with Grant 61625106 and the National Natural Science Foundation of China under Grant 61531011; The work of H. Lin was supported by JSPS KAKENHI Grant Number 17K06435. The associate editor coordinating the review of this paper and approving it for publication was J. Choi. *(Corresponding author: Feifei Gao.)*

F. Zhu and M. Yao are with the High-Tech Institute of Xi'an, Xi'an 710025, China (e-mail: fengchao_zhu@126.com; yaominli@sohu.com).

F. Gao is with the State Key Laboratory of Intelligent Technology and Systems, Tsinghua National Laboratory for Information Science and Technology, Department of Automation, Tsinghua University, Beijing 100084, China (e-mail: feifeigao@ieee.org).

S. Jin is with the National Communications Research Laboratory, Southeast University, Nanjing 210096, China (e-mail: jinshi@seu.edu.cn).

H. Lin is with the Department of Electrical and Information Systems, Osaka Prefecture University, Sakai 599-8531, Japan (e-mail: hai.lin@ieee.org).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCOMM.2017.2787582

philosophy has drawn considerable interest in CSI uncertainty mitigation [20]–[22]. In general, there are two major types of robust design: the stochastic robust beamforming and the worst-case robust beamforming. For the stochastic robust beamforming [23]–[27], the design is usually formulated as minimizing the transmission power while ensuring that all users' signal-to-interference-and-noise ratio (SINR) requirements are satisfied with high probabilities, where the CSI errors are modeled as random variables with known statistical properties. In contrast, the worst-case robust beamforming [28]–[32] is usually designed to satisfy the SINR requirements for all possible channel realizations in the uncertainty regions, where the CSI errors belong to some known bounded uncertainty sets. It has been shown in [28]–[32] that the downlink multiuser robust beamforming for multi-input single-output (MISO) channel needs to solve NP-hard optimizations [33], and only suboptimal solutions could be found through different rank-one approximations. Thus, existing robust beamforming algorithms [28]–[32] are very difficult to use in massive MIMO systems due to its forbiddingly huge computational complexity. To make robust beamforming feasible in massive MIMO systems, Hanif *et al.* [34] proposed a low complexity algorithm, which could greatly reduce the computational complexity than existing works. Due to the fact that the complexity reduction of [34] is based on an approximation of the initial problem, only suboptimal solutions could be derived. To the best of the authors' knowledge, the optimal solutions of robust beamforming for multiuser massive MIMO systems has not been reported yet.

In this paper, we consider robust beamforming for multiuser massive MIMO systems following BMDA scheme [14], [15], where the estimated channels are lied orthogonally to each other and the channel estimation errors are bounded by ball constants. Our target is to minimize the transmit power of BS, and meanwhile provide the intended users with satisfied SINRs for all possible channel realizations. The key result of this paper is that the robust beamforming problem for massive MIMO systems can be globally solved under elliptically bounded CSI errors, where semi-definite relaxation (SDR) [35] is strictly proved to be optimal. More importantly, we demonstrate that global optimal robust beamforming could be derived in closed-form, which will greatly reduce the computational complexity [36] for massive MIMO systems. Interestingly, it is shown that the direction of the optimal robust beamforming must be the same as that of the estimated channel, i.e., the matched filter is still optimal even though there is channel estimation error. Compared to directly applying the conventional SDP method [30], [31] onto massive MIMO, the proposed scheme yield the optimal solution with much less computational complexity and is applicable for real world massive MIMO system.

The rest of this paper is organized as follows: Section II describes the system model and formulates the proposed robust design; Section III converts the initial non-convex optimization into a convex semi-definite programming (SDP) and proves the optimality of SDR by showing the rank-one property of the optimal solutions; Closed-form solutions for the power allocation problem are derived in Section IV; Simulation results are provided in Section V and conclusions are drawn in Section VI.

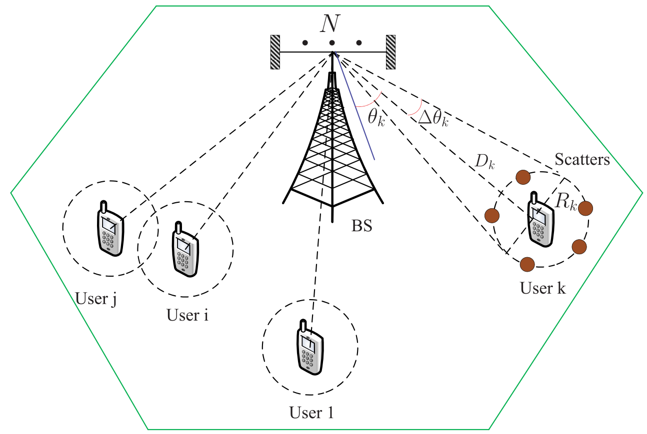


Fig. 1. Multiuser massive MIMO channel model in the form of ULA.

Notation: The Hermitian, inverse and Moore-Penrose inverse of $\boldsymbol{A}$ are denoted by $\boldsymbol{A}^{\mathrm{H}}$, $\boldsymbol{A}^{-1}$ and $\boldsymbol{A}^{\dagger}$ respectively; $\mathrm{Tr}(\boldsymbol{A})$ defines the trace; $\boldsymbol{I}$ and $\boldsymbol{0}$ represent an identity matrix and an all-zero matrix, respectively, with appropriate dimensions; $\boldsymbol{A} \succeq \boldsymbol{0}$ and $\boldsymbol{A} \succ \boldsymbol{0}$ mean that $\boldsymbol{A}$ is positive semi-definite and positive definite, respectively; The distribution of a circularly symmetric complex Gaussian (CSCG) random variable with zero mean and variance σ^2 is defined as $\mathcal{CN}(0, \sigma^2)$, and $\sim$ means "distributed as"; $\mathbb{R}^{a\times b}$ and $\mathbb{C}^{a\times b}$ denote the spaces of $a \times b$ matrices with real- and complex-valued entries, respectively; $\|\boldsymbol{x}\|$ is the Euclidean norm of a vector $\boldsymbol{x}$.

II. System Model and Problem Formulation

A. System Model

Let us consider a multiuser massive MIMO system shown in Fig. 1, where BS is equipped with $N \gg 1$ antennas in the form of ULA with supercritical antenna spacing (i.e., less than or equal to half wavelength). There are K single-antenna users with index set $\mathcal{K} = \{1, \ldots, K\}$ randomly distributed in the coverage area. The kth user is located at D_k meters away from BS and is surrounded by a ring of $P \gg 1$ local scatterers with the radius R_k [9], [13]. The channel from the kth user to BS is composed of P rays and can be expressed as [9], [13]:

$$\boldsymbol{h}_k = \frac{1}{\sqrt{P}} \sum_{p=1}^{P} \alpha_{k,p} \boldsymbol{a}(\theta_{k,p}), \tag{1}$$

where $\alpha_{k,p} \sim \mathcal{CN}(0, \zeta_{k,p})$ represents the complex gain of the pth ray. Moreover, $\boldsymbol{a}(\theta_{k,p}) \in \mathbb{C}^{N\times 1}$ is the steering vector and has the form

$$\boldsymbol{a}(\theta_{k,p}) = \left[1, e^{j\frac{2\pi d}{\lambda}\sin\theta_{k,p}}, \ldots, e^{j\frac{2\pi d}{\lambda}(N-1)\sin\theta_{k,p}}\right]^{\mathrm{H}}, \tag{2}$$

where d is the antenna spacing, λ denotes the signal wavelength, and $\theta_{k,p}$ represents the direction of arrival (DOA) of the pth ray.

In ideal massive MIMO case with $N \to \infty$, there is $\boldsymbol{h}_i^{\mathrm{H}}\boldsymbol{h}_j = 0$ but this is not true for a practical massive MIMO system with finite N. We can then assume that the channel vector

$\boldsymbol{h}_k$ for the kth user lies within a ball of radius ϵ_k around the estimated channel vector $\tilde{\boldsymbol{h}}_k$, i.e.,

$$\boldsymbol{h}_k \in \mathcal{U}_k = \left\{\tilde{\boldsymbol{h}}_k + \boldsymbol{\delta}_k \mid \|\boldsymbol{\delta}_k\| \le \epsilon_k\right\}, \tag{3}$$

where $\boldsymbol{\delta}_k \in \mathbb{C}^{N\times 1}$ is the channel estimation error whose norm is assumed to be bounded by ϵ_k.

For a general training scheme, there is no obvious relationship among $\tilde{\boldsymbol{h}}_k$. Hence, in this paper, we adopt the recently popular BDMA scheme [14], [15] such that pilots of different users are transmitted through orthogonal spatial directions.[1] In this case, the estimated channels $\tilde{\boldsymbol{h}}_i$ and $\tilde{\boldsymbol{h}}_j$ for different users must be orthogonal,[2] i.e.,

$$\tilde{\boldsymbol{h}}_i^{\mathrm{H}}\tilde{\boldsymbol{h}}_j = 0, \quad \forall i \ne j, \tag{4}$$

which is the key point of our proposed design. Please refer to [14] and [15] for detailed discussion of BDMA.

B. Problem Formulation

The downlink signal at the kth user can be expressed as

$$y_k = \boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{s}_k v_k + \sum_{i\ne k, i\in\mathcal{K}} \boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{s}_i v_i + n_k, \tag{5}$$

where $v_k \sim \mathcal{CN}(0,1)$ denotes the data symbol for the kth user; $\boldsymbol{s}_k$ is the corresponding downlink beamforming vector to be designed; $n_k \sim \mathcal{CN}(0,\sigma_k^2)$ represents the receive noise.

To meet the requirement of the energy efficiency of 5G communications, our target here is to minimize the transmit power of BS, at the same time providing intended users with satisfied SINR. The worst-case robust transmit beamforming design can then be formulated as [30], [31]

$$\mathbf{P1}: \quad \min_{\{\boldsymbol{s}_k\}} \sum_{k=1}^{K} \|\boldsymbol{s}_k\|^2 \tag{6}$$

$$\text{s.t.} \ \frac{\|\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{s}_k\|^2}{\sum_{i\ne k}\|\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{s}_i\|^2 + \sigma_k^2} \ge \gamma_k, \quad \forall \boldsymbol{h}_k \in \mathcal{U}_k, \tag{7}$$

$$k = 1,2,\ldots,K, \tag{8}$$

where $\gamma_k > 0$ is the desired SINR for the kth user. In small MIMO systems,[3] **P1** is NP hard, and it is difficult to obtain the optimal beamforming vectors $\{\boldsymbol{s}_k, k = 1,\ldots,K\}$. Thus, different approximation algorithms [30], [31] are designed to extract suboptimal solutions. Nevertheless, we will next show that the optimal solutions could be derived in closed-form thanks to (4) under BDMA massive MIMO scheme.

[1] In particular, the columns of discrete Fourier transform (DFT) matrix with non-overlap indices are assigned to different users as the training directions. Such a scheme is not optimal in terms of training design but could be implemented with much higher efficiency. Moreover, as long as N is large, the performance loss of the channel estimation accuracy is small compared to the optimal one.

[2] Note that the BDMA scheme can also be applied in millimeter wave communication systems [15], [37], where the estimated channels can be presented as orthogonal vectors with BDMA channel estimation algorithms.

[3] For contrast, we use small MIMO system to represent the conventional MIMO system where the number of transmit antennas is small.

III. The Optimal Robust Beamforming

A. Problem Re-formulation

The main difficulty for solving **P1** lies in the constraints (7). Usually, the S-Procedure [38] is adopted to convert the constraints (7) to easier cases [30], [31]. Substituting (3) into (7), we can equivalently rewrite (7) as

$$\begin{cases} \left(\tilde{\boldsymbol{h}}_k + \boldsymbol{\delta}_k\right)^{\mathrm{H}}\left(\dfrac{1}{\gamma_k}\boldsymbol{s}_k\boldsymbol{s}_k^{\mathrm{H}} - \sum\limits_{i\ne k}\boldsymbol{s}_i\boldsymbol{s}_i^{\mathrm{H}}\right)\left(\tilde{\boldsymbol{h}}_k + \boldsymbol{\delta}_k\right) - \sigma_k^2 \ge 0, \\ -\boldsymbol{\delta}_k^{\mathrm{H}}\boldsymbol{I}\boldsymbol{\delta}_k + \epsilon_k^2 \ge 0, \quad k = 1,2,\ldots,K. \end{cases} \tag{9}$$

Lemma 1 (S-Procedure [38]): For any $\boldsymbol{b}_1 \in \mathbb{C}^{n\times 1}$, $\boldsymbol{b}_2 \in \mathbb{C}^{n\times 1}$, $c_1 \in \mathbb{R}$, $c_2 \in \mathbb{R}$ and for any complex Hermitian matrices $\boldsymbol{A}_1 \in \mathbb{C}^{n\times n}$, $\boldsymbol{A}_2 \in \mathbb{C}^{n\times n}$, define $f_1(\boldsymbol{x})$ and $f_2(\boldsymbol{x})$ as

$$\begin{aligned} f_1(\boldsymbol{x}) &= \boldsymbol{x}^{\mathrm{H}}\boldsymbol{A}_1\boldsymbol{x} + \boldsymbol{b}_1^{\mathrm{H}}\boldsymbol{x} + \boldsymbol{x}^{\mathrm{H}}\boldsymbol{b}_1 + c_1, \\ f_2(\boldsymbol{x}) &= \boldsymbol{x}^{\mathrm{H}}\boldsymbol{A}_2\boldsymbol{x} + \boldsymbol{b}_2^{\mathrm{H}}\boldsymbol{x} + \boldsymbol{x}^{\mathrm{H}}\boldsymbol{b}_2 + c_2, \end{aligned} \tag{10}$$

respectively. The condition $f_1(\boldsymbol{x}) \ge 0 \Rightarrow f_2(\boldsymbol{x}) \ge 0$ holds true if and only if there exists a nonnegative μ, such that

$$\begin{bmatrix} \boldsymbol{A}_2 & \boldsymbol{b}_2 \\ \boldsymbol{b}_2^{\mathrm{H}} & c_2 \end{bmatrix} - \mu \begin{bmatrix} \boldsymbol{A}_1 & \boldsymbol{b}_1 \\ \boldsymbol{b}_1^{\mathrm{H}} & c_1 \end{bmatrix} \succeq \mathbf{0}.$$

According to Lemma 1, we know that the constraints in (9) hold true if and only if there exists $\mu_k \ge 0,\ k = 1,2,\ldots,K$ such that

$$\begin{bmatrix} \boldsymbol{X}_k + \mu_k\boldsymbol{I} & \boldsymbol{X}_k\tilde{\boldsymbol{h}}_k \\ \tilde{\boldsymbol{h}}_k^{\mathrm{H}}\boldsymbol{X}_k^{\mathrm{H}} & \tilde{\boldsymbol{h}}_k^{\mathrm{H}}\boldsymbol{X}_k\tilde{\boldsymbol{h}}_k - \sigma_k^2 - \mu_k\epsilon_k^2 \end{bmatrix} \succeq \mathbf{0}, \tag{11}$$

where for simplicity, we define $\boldsymbol{X}_k$ as

$$\boldsymbol{X}_k = \frac{1}{\gamma_k}\boldsymbol{s}_k\boldsymbol{s}_k^{\mathrm{H}} - \sum_{i\ne k}\boldsymbol{s}_i\boldsymbol{s}_i^{\mathrm{H}}. \tag{12}$$

It is then clear that **P1** can be equivalently expressed as

P1 − EQV :

$$\min_{\{\mu_k\},\{\boldsymbol{s}_k\},\{\boldsymbol{S}_k\}} \sum_{k=1}^{K} \mathrm{Tr}(\boldsymbol{S}_k) \tag{13}$$

$$\text{s.t.} \begin{bmatrix} \boldsymbol{X}_k + \mu_k\boldsymbol{I} & \boldsymbol{X}_k\tilde{\boldsymbol{h}}_k \\ \tilde{\boldsymbol{h}}_k^{\mathrm{H}}\boldsymbol{X}_k^{\mathrm{H}} & \tilde{\boldsymbol{h}}_k^{\mathrm{H}}\boldsymbol{X}_k\tilde{\boldsymbol{h}}_k - \sigma_k^2 - \mu_k\epsilon_k^2 \end{bmatrix} \succeq \mathbf{0}, \tag{14}$$

$$\boldsymbol{X}_k = \frac{1}{\gamma_k}\boldsymbol{S}_k - \sum_{i\ne k}\boldsymbol{S}_i, \tag{15}$$

$$\mu_k \ge 0, \quad \boldsymbol{S}_k = \boldsymbol{s}_k\boldsymbol{s}_k^{\mathrm{H}},\ k = 1,2,\ldots,K. \tag{16}$$

Note that the nonlinear constraint $\boldsymbol{S}_k = \boldsymbol{s}_k\boldsymbol{s}_k^{\mathrm{H}}$ in (16) is equivalent to: $\boldsymbol{S}_k \succeq \mathbf{0}$ and $\mathrm{Rank}(\boldsymbol{S}_k) = 1$.

However, **P1 − EQV** is still NP hard due to the fact that the rank constraint (16) is non-convex. Using the SDR technique [35] (i.e., dropping the constraints $\mathrm{Rank}(\boldsymbol{S}_k) = 1$, $k = 1,2,\ldots,K$), we can obtain the following relaxed convex optimization:

P1 − SDR :

$$\min_{\{\mu_k\},\{\boldsymbol{S}_k\}} \sum_{k=1}^{K} \mathrm{Tr}(\boldsymbol{S}_k) \tag{17}$$

$$\text{s.t.} \begin{bmatrix} \boldsymbol{X}_k + \mu_k \boldsymbol{I} & \boldsymbol{X}_k \tilde{\boldsymbol{h}}_k \\ \tilde{\boldsymbol{h}}_k^{\mathrm{H}} \boldsymbol{X}_k^{\mathrm{H}} & \tilde{\boldsymbol{h}}_k^{\mathrm{H}} \boldsymbol{X}_k \tilde{\boldsymbol{h}}_k - \sigma_k^2 - \mu_k \epsilon_k^2 \end{bmatrix} \succeq \boldsymbol{0}, \tag{18}$$

$$\boldsymbol{X}_k = \frac{1}{\gamma_k} \boldsymbol{S}_k - \sum_{i \neq k} \boldsymbol{S}_i, \tag{19}$$

$$\mu_k \geq 0, \quad \boldsymbol{S}_k \succeq \boldsymbol{0}, \ k = 1, 2, \ldots, K. \tag{20}$$

In fact, **P1 − SDR** is a semi-definite programming (SDP) problem, which can be efficiently solved by the standard convex optimization tools [33]. However, SDR in small MIMO system does not guarantee the rank-one solutions for the relaxed optimization, where only suboptimal solutions could be derived.

Nevertheless, we will show that **P1 − SDR** indeed guarantees rank-one optimal solutions for BDMA massive MIMO scheme.

B. Optimal Rank-One Solutions

The following proposition will be useful throughout the rest of the paper.

Proposition 1: Suppose $\{\mu_k^\star\}$ and $\{\boldsymbol{S}_k^\star\}$ are the optimal solutions of **P1 − SDR**. There must be

(a) $\mu_k^\star > 0$, $\forall k \in \{1, \ldots, K\}$;

(b) $\mathrm{Rank}\left(\boldsymbol{X}_k^\star + \mu_k^\star \boldsymbol{I}\right) \geq 1$, where $\boldsymbol{X}_k^\star = \frac{1}{\gamma_k} \boldsymbol{S}_k^\star - \sum_{i \neq k} \boldsymbol{S}_i^\star$, $\forall k \in \{1, \ldots, K\}$.

Proof: See Appendix A. ■

From Proposition 1, we can further reformulate the constraints (18) to gain more insightful solutions. Let the EVD of $\boldsymbol{X}_k$ be $\boldsymbol{X}_k = \boldsymbol{U}_{x,k} \boldsymbol{\Lambda}_{x,k} \boldsymbol{U}_{x,k}^{\mathrm{H}}$ with the eigenvalues $q_{k,1} \geq \ldots \geq q_{k,N}$. Due to the fact that $\mu_k \boldsymbol{I} = \boldsymbol{U}_{x,k} (\mu_k \boldsymbol{I}) \boldsymbol{U}_{x,k}^{\mathrm{H}}$, the EVD of $\boldsymbol{X}_k + \mu_k \boldsymbol{I}$ can be represented as $\boldsymbol{X}_k + \mu_k \boldsymbol{I} = \boldsymbol{U}_{x,k} \left(\boldsymbol{\Lambda}_{x,k} + \mu_k \boldsymbol{I}\right) \boldsymbol{U}_{x,k}^{\mathrm{H}}$ with the eigenvalues $q_{k,1} + \mu_k \geq \ldots \geq q_{k,N} + \mu_k$. Using Proposition 1, since $\mathrm{Rank}\left(\boldsymbol{X}_k^\star + \mu_k^\star \boldsymbol{I}\right) \geq 1$ should be satisfied at the optimal point, there must be $q_{k,1} + \mu_k > 0$. Then assuming $\mathrm{Rank}\,(\boldsymbol{X}_k + \mu_k \boldsymbol{I}) = l \geq 1$, there are

$$q_{k,i} + \mu_k = 0, \quad \forall i \in \{l+1, \ldots, N\}, \ \forall k \in \{1, \ldots, K\}. \tag{21}$$

Thus, the EVD of $\boldsymbol{X}_k + \mu_k \boldsymbol{I}$ can be further expressed as

$$\boldsymbol{X}_k + \mu_k \boldsymbol{I} = \boldsymbol{U}_{x,k} \begin{bmatrix} \boldsymbol{\Sigma} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix} \boldsymbol{U}_{x,k}^{\mathrm{H}}, \tag{22}$$

where $\boldsymbol{\Sigma}$ is positive diagonal with the diagonal eigenvalues $q_{k,1} + \mu_k \geq \ldots \geq q_{k,l} + \mu_k > 0$. Aiming to obtain a compact form of (18) for **P1 − SDR**, we provide the following proposition.

Proposition 2: The Moore-Penrose inverse of $\boldsymbol{X}_k + \mu_k \boldsymbol{I}$ can be derived as[4]

$$(\boldsymbol{X}_k + \mu_k \boldsymbol{I})^\dagger = \boldsymbol{U}_{x,k} \begin{bmatrix} \boldsymbol{\Sigma}^{-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix} \boldsymbol{U}_{x,k}^{\mathrm{H}}. \tag{23}$$

[4]Note that if $\mathrm{Rank}\,(\boldsymbol{X}_k + \mu_k \boldsymbol{I}) = N$, then there is $(\boldsymbol{X}_k + \mu_k \boldsymbol{I})^\dagger = (\boldsymbol{X}_k + \mu_k \boldsymbol{I})^{-1}$.

Proof: It can be shown that Eq. (23) satisfies the Moore-Penrose conditions [39]: (a) $\boldsymbol{A}\boldsymbol{A}^\dagger\boldsymbol{A} = \boldsymbol{A}$; (b) $\boldsymbol{A}^\dagger\boldsymbol{A}\boldsymbol{A}^\dagger = \boldsymbol{A}^\dagger$; (c) $\left(\boldsymbol{A}\boldsymbol{A}^\dagger\right)^{\mathrm{H}} = \boldsymbol{A}\boldsymbol{A}^\dagger$; (d) $\left(\boldsymbol{A}^\dagger\boldsymbol{A}\right)^{\mathrm{H}} = \boldsymbol{A}^\dagger\boldsymbol{A}$. ■

Lemma 2 (Generalized Schur's Complement [40]): Let

$$\boldsymbol{M} = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{B}^{\mathrm{H}} & \boldsymbol{C} \end{bmatrix}$$

be a Hermitian matrix. Then, $\boldsymbol{M} \succeq \boldsymbol{0}$ if and only if $\boldsymbol{C} - \boldsymbol{B}^{\mathrm{H}}\boldsymbol{A}^\dagger\boldsymbol{B} \succeq \boldsymbol{0}$ and $\left(\boldsymbol{I} - \boldsymbol{A}\boldsymbol{A}^\dagger\right)\boldsymbol{B} = \boldsymbol{0}$ (assuming $\boldsymbol{A} \succeq \boldsymbol{0}$), or $\boldsymbol{A} - \boldsymbol{B}\boldsymbol{C}^\dagger\boldsymbol{B}^{\mathrm{H}} \succeq \boldsymbol{0}$ and $\left(\boldsymbol{I} - \boldsymbol{C}\boldsymbol{C}^\dagger\right)\boldsymbol{B}^{\mathrm{H}} = \boldsymbol{0}$ (assuming $\boldsymbol{C} \succeq \boldsymbol{0}$), where $\boldsymbol{A}^\dagger$ and $\boldsymbol{C}^\dagger$ are the generalized inverses of $\boldsymbol{A}$ and $\boldsymbol{C}$, respectively.

From Proposition 2 and Lemma 2, we know that (18) holds if and only if the following equations are satisfied:

$$\begin{cases} \left[\boldsymbol{I} - (\boldsymbol{X}_k + \mu_k \boldsymbol{I})(\boldsymbol{X}_k + \mu_k \boldsymbol{I})^\dagger\right] \boldsymbol{X}_k \tilde{\boldsymbol{h}}_k = \boldsymbol{0}, \\ \tilde{\boldsymbol{h}}_k^{\mathrm{H}} \boldsymbol{X}_k \tilde{\boldsymbol{h}}_k - \sigma_k^2 - \mu_k \epsilon_k^2 - \tilde{\boldsymbol{h}}_k^{\mathrm{H}} \boldsymbol{X}_k^{\mathrm{H}} (\boldsymbol{X}_k + \mu_k \boldsymbol{I})^\dagger \boldsymbol{X}_k \tilde{\boldsymbol{h}}_k \geq 0, \\ \boldsymbol{X}_k + \mu_k \boldsymbol{I} \succeq \boldsymbol{0}. \end{cases} \tag{24}$$

Then using (21), we know from the first equation of (24) that

$$\begin{aligned} &\left[\boldsymbol{I} - (\boldsymbol{X}_k + \mu_k \boldsymbol{I})(\boldsymbol{X}_k + \mu_k \boldsymbol{I})^\dagger\right] \boldsymbol{X}_k \tilde{\boldsymbol{h}}_k \\ &\quad = \boldsymbol{U}_{x,k} \begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & -\mu_k \boldsymbol{I}^{(N-l)\times(N-l)} \end{bmatrix} \boldsymbol{U}_{x,k}^{\mathrm{H}} \tilde{\boldsymbol{h}}_k = \boldsymbol{0}, \end{aligned} \tag{25}$$

where $\boldsymbol{\Lambda}_{x,k}^{l \times l}$ is diagonal with the values $q_{k,1} \geq \ldots \geq q_{k,l}$. It follows from (25) that $\left[\boldsymbol{u}_{x,k,l+1}^{\mathrm{H}}, \ldots, \boldsymbol{u}_{x,k,N}^{\mathrm{H}}\right] \tilde{\boldsymbol{h}}_k = \boldsymbol{0}$ holds true, where $\boldsymbol{u}_{x,k,i}$, $\forall i \in \{l+1, \ldots, N\}$ is the ith column of $\boldsymbol{U}_{x,k}$. Due to the fact that $\boldsymbol{X}_k = \boldsymbol{X}_k^{\mathrm{H}}$, from the second equation of (24) we know

$$\begin{aligned} &\tilde{\boldsymbol{h}}_k^{\mathrm{H}} \boldsymbol{X}_k \tilde{\boldsymbol{h}}_k - \sigma_k^2 - \mu_k \epsilon_k^2 - \tilde{\boldsymbol{h}}_k^{\mathrm{H}} \boldsymbol{X}_k^{\mathrm{H}} (\boldsymbol{X}_k + \mu_k \boldsymbol{I})^\dagger \boldsymbol{X}_k \tilde{\boldsymbol{h}}_k \\ &\quad = \mathrm{Tr}\left(\tilde{\boldsymbol{H}}_k \boldsymbol{X}_k\right) - \sigma_k^2 - \mu_k \epsilon_k^2 - \mathrm{Tr}\left[\tilde{\boldsymbol{H}}_k \boldsymbol{X}_k^{\mathrm{H}} (\boldsymbol{X}_k + \mu_k \boldsymbol{I})^\dagger \boldsymbol{X}_k\right] \\ &\quad = \mathrm{Tr}\left\{\left(\tilde{\boldsymbol{H}}_k \boldsymbol{X}_k\right)\left[\boldsymbol{I} - (\boldsymbol{X}_k + \mu_k \boldsymbol{I})^\dagger \boldsymbol{X}_k\right]\right\} - \sigma_k^2 - \mu_k \epsilon_k^2 \geq 0, \end{aligned} \tag{26}$$

where $\tilde{\boldsymbol{H}}_k = \tilde{\boldsymbol{h}}_k \tilde{\boldsymbol{h}}_k^{\mathrm{H}}$.

Lemma 3 (Theory of Majorization [41]): Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be two $N \times N$ positive semi-definite matrices with eigenvalues $\alpha_1 \geq \ldots \geq \alpha_N$ and $\beta_1 \geq \ldots \geq \beta_N$, respectively. Then there holds

$$\sum_{i=1}^{N} \alpha_i \beta_{N-i+1} \leq \mathrm{Tr}(\boldsymbol{A}\boldsymbol{B}) \leq \sum_{i=1}^{N} \alpha_i \beta_i. \tag{27}$$

Let the EVD of $\tilde{\boldsymbol{H}}_k$ be $\tilde{\boldsymbol{H}}_k = \boldsymbol{U}_{\tilde{h},k} \boldsymbol{\Lambda}_{\tilde{h},k} \boldsymbol{U}_{\tilde{h},k}^{\mathrm{H}}$ with the eigenvalues $\lambda_{k,1} \geq \ldots \geq \lambda_{k,N}$. Since $\tilde{\boldsymbol{H}}_k = \tilde{\boldsymbol{h}}_k \tilde{\boldsymbol{h}}_k^{\mathrm{H}}$ has only one nonzero eigenvalue, there must be $\lambda_{k,1} > 0$ and $\lambda_{k,2} = \ldots = \lambda_{k,N} = 0$. According to (25) and Lemma 3, we know that (26) can be reformulated as

$$\begin{aligned}&\mathrm{Tr}\left\{\left(\tilde{\boldsymbol{H}}_k\boldsymbol{X}_k\right)\left[\boldsymbol{I}-\left(\boldsymbol{X}_k+\mu_k\boldsymbol{I}\right)^{\dagger}\boldsymbol{X}_k\right]\right\}-\sigma_k^2-\mu_k\epsilon_k^2\\&\quad=\mathrm{Tr}\left\{\left(\tilde{\boldsymbol{H}}_k\boldsymbol{X}_k\right)\left[\boldsymbol{I}-\left(\boldsymbol{X}_k+\mu_k\boldsymbol{I}\right)^{\dagger}\left(\boldsymbol{X}_k+\mu_k\boldsymbol{I}-\mu_k\boldsymbol{I}\right)\right]\right\}\\&\qquad-\sigma_k^2-\mu_k\epsilon_k^2\\&\quad=\mu_k\mathrm{Tr}\left[\tilde{\boldsymbol{H}}_k\boldsymbol{X}_k\left(\boldsymbol{X}_k+\mu_k\boldsymbol{I}\right)^{\dagger}\right]-\sigma_k^2-\mu_k\epsilon_k^2\\&\quad\leq\frac{\mu_k q_{k,1}\lambda_{k,1}}{\mu_k+q_{k,1}}-\sigma_k^2-\mu_k\epsilon_k^2,\end{aligned}\tag{28}$$

where the equality holds when $\boldsymbol{u}_{x,k,1}=\boldsymbol{u}_{\tilde{h},k,1}$, and $\boldsymbol{u}_{x,k,1}$, $\boldsymbol{u}_{\tilde{h},k,1}$ are the first column of $\boldsymbol{U}_{x,k}$ and $\boldsymbol{U}_{\tilde{h},k}$, respectively. Due to $\tilde{\boldsymbol{H}}_k=\tilde{\boldsymbol{h}}_k\tilde{\boldsymbol{h}}_k^{\mathrm{H}}$, there must be $\boldsymbol{u}_{x,k,1}=\tilde{\boldsymbol{h}}_k/\|\tilde{\boldsymbol{h}}_k\|$. Combining the discussions of (25) and (28), we know that when $\boldsymbol{u}_{x,k,1}=\tilde{\boldsymbol{h}}_k/\|\tilde{\boldsymbol{h}}_k\|$, the constraints in (18) can be equivalently expressed as $\mu_k q_{k,1}\lambda_{k,1}/(\mu_k+p_{k,1})-\sigma_k^2-\mu_k\epsilon_k^2\geq 0$ and $\boldsymbol{X}_k+\mu_k\boldsymbol{I}\succeq\boldsymbol{0}$. Then we have the following proposition.

Proposition 3: Suppose $\{\boldsymbol{S}_k^\star\}$ are the optimal transmit covariances of $\mathbf{P1-SDR}$, there must be

(a) $\boldsymbol{S}_j^\star\boldsymbol{S}_k^\star=\boldsymbol{0}$, $\forall j\neq k$ and $j,k\in\{1,\ldots,K\}$;

(b) $\mathrm{Rank}\left(\boldsymbol{S}_k^\star\right)\geq 1$, i.e., there exists at least one nonzero eigenvalue for $\boldsymbol{S}_k^\star$.

Proof: See Appendix B. ■

Result (a) says that the optimal transmit covariances must be orthogonal to each other, which implies that zero-forcing beamforming is optimal for $\mathbf{P1-SDR}$, which is consistent with the general advantage of massive MIMO. Result (b) says that the optimal transmit covariance could not be zero, which is mainly due to the fact that the SINR requirements for all users are nonzero.

Theorem 1: Under the equality of (4), the optimal robust beamforming vectors of **P1** can be derived in closed-form

$$\boldsymbol{s}_k^\star=\sqrt{p_k^\star}\tilde{\boldsymbol{h}}_k/\|\tilde{\boldsymbol{h}}_k\|,\quad\forall k\in\{1,\ldots,K\},\tag{29}$$

where $p_k^\star=q_{k,1}\gamma_k$ is the power allocated for the kth beamforming vector that will be designed later. The optimal transmit covariances must satisfy

$$\mathrm{Rank}\left(\boldsymbol{S}_k^\star\right)=1,\quad\forall k\in\{1,\ldots,K\},\tag{30}$$

where $\boldsymbol{S}_k^\star=\boldsymbol{s}_k^\star\boldsymbol{s}_k^{\star\mathrm{H}}=p_k^\star\tilde{\boldsymbol{h}}_k\tilde{\boldsymbol{h}}_k^{\mathrm{H}}/\|\tilde{\boldsymbol{h}}_k\|^2$ holds.

Proof: See Appendix C. ■

Remark 1: Theorem 1 indicates that beamforming is the optimal transmit strategy for **P1** rather than the joint precoding scheme. Meanwhile, the optimal beamforming vector must be $\boldsymbol{s}_k^\star=\sqrt{p_k^\star}\tilde{\boldsymbol{h}}_k/\|\tilde{\boldsymbol{h}}_k\|$ which says that the matched filter is optimal even for the robust transmission design.

However, $\mathbf{P1-SDR}$ will suffer the infeasible case when ϵ_k is large enough, which is inevitable due to the fact that we could not control the channel estimate error. Firstly, since $\boldsymbol{X}_k+\mu_k\boldsymbol{I}\succeq\boldsymbol{0}$ must be satisfied in (18), it follows from (C.3) that $\mu_k\geq p_k$ holds. Secondly, since $\gamma_k>0$ is true, we obtain $\mu_k\geq q_{k,1}\gamma_k$. Thirdly, it is easily shown that Eq. (28) is a monotonically increasing function with respect to $q_{k,1}$. As a result, from Eq. (28) we obtain

$$\begin{aligned}0&\leq\frac{\mu_k q_{k,1}\lambda_{k,1}}{\mu_k+q_{k,1}}-\sigma_k^2-\mu_k\epsilon_k^2\leq\frac{\mu_k q_{k,1}^\star\lambda_{k,1}}{\mu_k+q_{k,1}^\star}-\sigma_k^2-\mu_k\epsilon_k^2\\&\leq\mu_k\left(\frac{\lambda_{k,1}}{1+\gamma_k}-\epsilon_k^2\right)-\sigma_k^2,\end{aligned}\tag{31}$$

where $\lambda_{k,1}=\mathrm{Tr}(\tilde{\boldsymbol{H}}_k)=\|\tilde{\boldsymbol{h}}_k\|^2$. It follows from (31) that $\epsilon_k<\sqrt{\lambda_{k,1}}/\sqrt{1+\gamma_k}=\|\tilde{\boldsymbol{h}}_k\|/\sqrt{1+\gamma_k}$ must be satisfied. If $\epsilon_k\geq\|\tilde{\boldsymbol{h}}_k\|/\sqrt{1+\gamma_k}$, from (31) we know that $\mathbf{P1-SDR}$ is infeasible. Nevertheless, we could assume that

$$\epsilon_k<\|\tilde{\boldsymbol{h}}_k\|/\sqrt{1+\gamma_k},\tag{32}$$

which is reasonable since large channel estimate errors usually lead to unacceptable degradation in performance.

IV. The Optimal Power Allocation

In Section III, it is shown that the optimal robust beamforming directions must be the same as that of the estimated channels. In this section, we will solve the remaining power allocation problem. Since $\mu_k^\star>0$ has been strictly proved in Proposition 1, we know form Theorem 1 that **P1** can be simplified to

P2 :

$$\min_{\{\mu_k\},\{q_{k,1}\}}\sum_{k=1}^{K}q_{k,1}\gamma_k\tag{33}$$

$$\text{s.t.}\ \frac{\mu_k q_{k,1}\|\tilde{\boldsymbol{h}}_k\|^2}{\mu_k+q_{k,1}}-\sigma_k^2-\mu_k\epsilon_k^2\geq 0,\quad k=1,2,\ldots,K,\tag{34}$$

$$\mu_k>0,\quad q_{k,1}>0,\ k=1,2,\ldots,K,\tag{35}$$

where $q_{k,1}=p_k/\gamma_k$. Define

$$f(\mu_k,q_{k,1})=\frac{\mu_k q_{k,1}\|\tilde{\boldsymbol{h}}_k\|^2}{\mu_k+q_{k,1}},\tag{36}$$

whose Hessian matrix is given by

$$\nabla^2 f(\mu_k,q_{k,1})=\frac{-2\|\tilde{\boldsymbol{h}}_k\|^2}{(\mu_k+q_{k,1})^3}\begin{bmatrix}\mu_k^2&-\mu_k q_{k,1}\\-\mu_k q_{k,1}&q_{k,1}^2\end{bmatrix}\prec\boldsymbol{0},\tag{37}$$

which implys that $f(\mu_k,q_{k,1})$ is a concave function. Thus, **P2** is convex.

The Lagrange of **P2** is defined as

$$\begin{aligned}&\mathcal{L}\left(\{\mu_k\},\{q_{k,1}\},\{\xi_k\},\{\varsigma_k\},\{\tau_k\}\right)\\&\quad=\sum_{k=1}^{K}q_{k,1}\gamma_k-\sum_{k=1}^{K}\left[\xi_k\left(\frac{\mu_k q_{k,1}\|\tilde{\boldsymbol{h}}_k\|^2}{\mu_k+q_{k,1}}-\sigma_k^2-\mu_k\epsilon_k^2\right)\right]\\&\qquad-\sum_{k=1}^{K}\varsigma_k\mu_k-\sum_{k=1}^{K}\tau_k q_{k,1},\end{aligned}\tag{38}$$

where $\{\xi_k\geq 0\}$, $\{\varsigma_k\geq 0\}$ and $\{\tau_k\geq 0\}$ are the dual variables associated with the constraints in (34) and (35). The Karush-Kuhn-Tucker (KKT) conditions related to $\{\mu_k\}$ and $\{q_{k,1}\}$ can be formulated as

$$\frac{\partial \mathcal{L}}{\partial q_{k,1}^\star} = \gamma_k - \frac{\xi_k^\star \mu_k^{\star 2} \|\tilde{\boldsymbol{h}}_k\|^2}{(\mu_k^\star + q_{k,1}^\star)^2} - \tau_k = 0, \tag{39}$$

$$\frac{\partial \mathcal{L}}{\partial \mu_k^\star} = \xi_k^\star \epsilon_k^2 - \frac{\xi_k^\star q_{k,1}^{\star 2} \|\tilde{\boldsymbol{h}}_k\|^2}{(\mu_k^\star + q_{k,1}^\star)^2} - \varsigma_k = 0, \tag{40}$$

$$\xi_k^\star \left(\frac{\mu_k^\star q_{k,1}^\star \|\tilde{\boldsymbol{h}}_k\|^2}{\mu_k^\star + q_{k,1}^\star} - \sigma_k^2 - \mu_k^\star \epsilon_k^2 \right) = 0, \tag{41}$$

$$\varsigma_k^\star \mu_k^\star = 0, \quad \tau_k^\star q_{k,1}^\star = 0, \ \forall k \in \{1, \dots, K\}, \tag{42}$$

where $\{\mu_k^\star > 0\}$ and $\{q_{k,1}^\star > 0\}$ are the optimal primal variables, and $\{\xi_k^\star \geq 0\}$, $\{\varsigma_k^\star \geq 0\}$ and $\{\tau_k^\star \geq 0\}$ are the optimal dual variables.

Due to the fact that $\{\mu_k^\star > 0\}$ and $\{q_{k,1}^\star > 0\}$, it follows from (42) that $\{\varsigma_k^\star = 0\}$ and $\{\tau_k^\star = 0\}$. Since $\{\mu_k^\star > 0\}$, $\{q_k^\star > 0\}$, $\{\gamma_k > 0\}$ and $\{\xi_k^\star \geq 0\}$, it follows from (39) that $\{\xi_k^\star > 0\}$ must be satisfied. Then from (41) we know

$$\frac{\mu_k^\star q_{k,1}^\star \|\tilde{\boldsymbol{h}}_k\|^2}{\mu_k^\star + q_{k,1}^\star} - \sigma_k^2 - \mu_k^\star \epsilon_k^2 = 0, \quad \forall k \in \{1, \dots, K\}, \tag{43}$$

which says that the constraints in (34) of **P2** must be satisfied with equality at the optimal point. Moreover, it is easy to obtain from (39) that

$$q_{k,1}^\star = \mu_k^\star (\|\tilde{\boldsymbol{h}}_k\| \sqrt{\xi_k^\star / \gamma_k} - 1), \quad \forall k \in \{1, \dots, K\}, \tag{44}$$

and from (39) and (40) that

$$\xi_k^\star = \frac{\gamma_k q_{k,1}^{\star 2}}{\mu_k^{\star 2} \epsilon_k^2}. \tag{45}$$

Substituting (45) into (44), there is

$$q_{k,1}^\star = \frac{\mu_k^\star \epsilon_k}{\|\tilde{\boldsymbol{h}}_k\| - \epsilon_k}, \quad \forall k \in \{1, \dots, K\}. \tag{46}$$

Replacing $q_k^\star$ in (43) with (46), we have

$$\mu_k^\star = \frac{\sigma_k^2}{\|\tilde{\boldsymbol{h}}_k\| \epsilon_k - \epsilon_k^2}, \quad \forall k \in \{1, \dots, K\}. \tag{47}$$

Then the optimal power allocation can be derived as

$$p_k^\star = q_{k,1}^\star \gamma_k = \frac{\mu_k^\star \gamma_k \epsilon_k}{\|\tilde{\boldsymbol{h}}_k\| - \epsilon_k}, \quad \forall k \in \{1, \dots, K\}. \tag{48}$$

Note that if $\epsilon_k \geq \|\tilde{\boldsymbol{h}}_k\|$, then **P2** is infeasible. Nevertheless, since $\gamma_k > 0$, the nonexistence of the trivial solution for **P2** can be guaranteed when (32) is satisfied.

V. Simulation Results

In this section, computer simulations are presented to evaluate the performance of the proposed robust beamforming algorithm. The BS is equipped with $N \geq 64$ transmit antennas and $K = 30$ users are all equipped with single antenna. The received noises per antenna for all users are generated as independent CSCG random variables distributed with $\mathcal{CN}(0, 1)$. For simplicity, the receive SINR for all users are assumed to be the same, i.e., $\gamma = \gamma_1, \dots, = \gamma_K$. The simulation results are averaged over 10000 Monte Carlo runs.

TABLE I
Running Time Comparison

Methods \ N	32	64	128	256
SDP [31]	2.1 s	4.5 s	7.3 s	14.5 s
SOCP [34]	1.8 s	4.2 s	7.0 s	14.2 s
Proposed	0.0002 s	0.0003 s	0.0004 s	0.0005 s

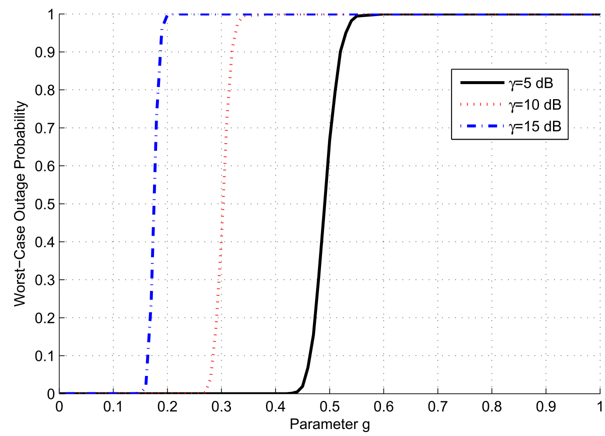


Fig. 2. Worst-case outage probability of beamforming versus the parameter g with $K = 30$ and $N = 128$.

The philosophy of the robustness in this paper is to guarantee a performance level for any channel realization in the uncertainty region. It needs to be mentioned that if $\epsilon_k \geq \|\tilde{\boldsymbol{h}}_k\| / \sqrt{1 + \gamma_k}$ (c.f. Eq. (32)), then there is no guarantee of performance. Even if $\epsilon_k < \|\tilde{\boldsymbol{h}}_k\| / \sqrt{1 + \gamma_k}$, there is still a possibility that the designed beamforming could not guarantee any performance because of $\mu_k \left(\frac{\|\tilde{\boldsymbol{h}}_k\|^2}{1+\gamma_k} - \epsilon_k^2 \right) - \sigma_k^2 < 0$ (c.f. Eq. (31) for details). Define $\epsilon_k = g \|\tilde{\boldsymbol{h}}_k\|$ with $g \in [0, 1)$. Then the worst-case outage probability of beamforming is $P_{out} = \text{Prob} \left\{ \mu_k \left(\frac{\|\tilde{\boldsymbol{h}}_k\|^2}{1+\gamma_k} - \epsilon_k^2 \right) - \sigma_k^2 < 0 \right\}$.

In the first example, we compare the average CPU running times for solving $\mathbf{P1-SDR}$ through the SDP method [31], the SOCP method [34] and the proposed closed-form method in Table I, respectively. We see the CPU running times of existing works are much larger than that of the proposed closed-form method, say about 4 order of magnitude. Though the running times do not strictly present the computational complexity, such a huge difference well demonstrates the superiority of the proposed method in terms of the complexity. Moreover, we do not compare the performance of [31] and [34] and the proposed one because they would both achieve the optimal solutions under the condition (4).

In the second example, we plot p_{out} versus g with different γ, N and K for the proposed method and the non-robust method in Fig. 2, Fig. 3 and Fig. 4, respectively. It is shown in Fig. 2 that the worst-case outage probability is decreased with the decrease of required SINR γ. This is mainly due to the fact that when γ is small, the channel estimate error are more flexible to guarantee the feasibility of the proposed

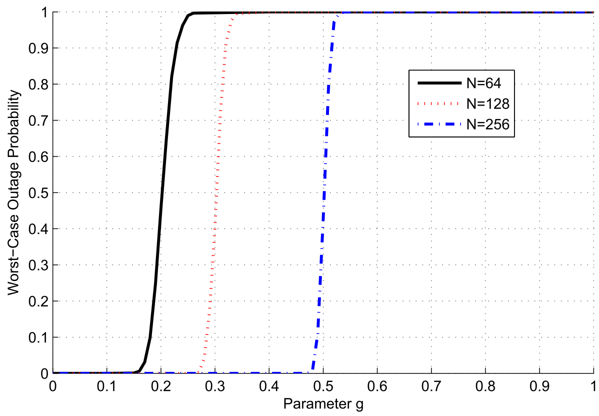


Fig. 3. Worst-case outage probability of beamforming versus the parameter g with $K = 30$ and $\gamma = 10$ dB.

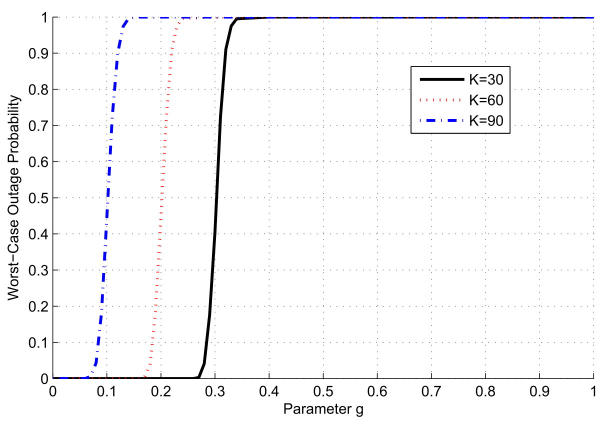


Fig. 4. Worst-case outage probability of beamforming versus the parameter g with $N = 128$ and $\gamma = 10$ dB.

design (please c.f. Eq. (31) for details). Moreover, it is clear in Fig. 2 that the worst-case outage probability is increased with the increase of g, i.e., the worst-case outage probability is zero when $g \leq 0.4$ and $\gamma = 5$ dB. While the worst-case outage probability will always be one when $g \geq 0.6$ and $\gamma = 5$ dB. In fact, the proposed design **P2** will be infeasible almost surely when g becomes large. It is shown in Fig. 3 that the worst-case outage probability will decrease when the number of transmit antennas becomes large. Similarly, it is seen from Fig. 4 that the worst-case outage probability will decrease when the number of users becomes small. From Fig. 2, Fig. 3 and Fig. 4, we know that decreasing the SINR thresholds for all users, increasing the number of transmit antennas at BS, or decreasing the number of users will help to improve the feasibility of the proposed robust beamforming algorithm.

In the third example, we plot the average user SINR versus parameter g with different γ, N and K for the proposed method and the non-robust method in Fig. 5, Fig. 6 and Fig. 7, respectively.[5] It is clear in Fig. 5 that for the non-robust method, the average user SINR is a moronically decreasing

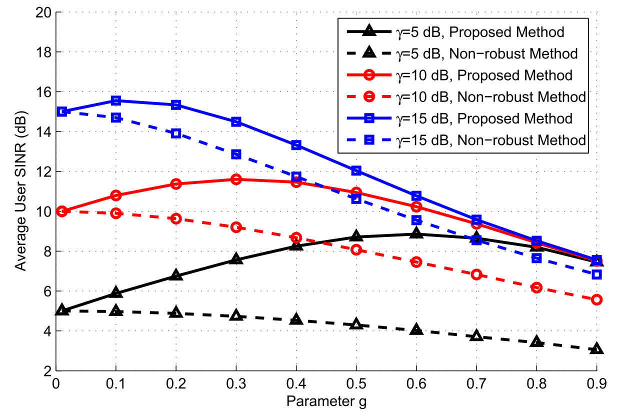


Fig. 5. Average user SINR versus g with different γ for $K = 30$ and $N = 128$.

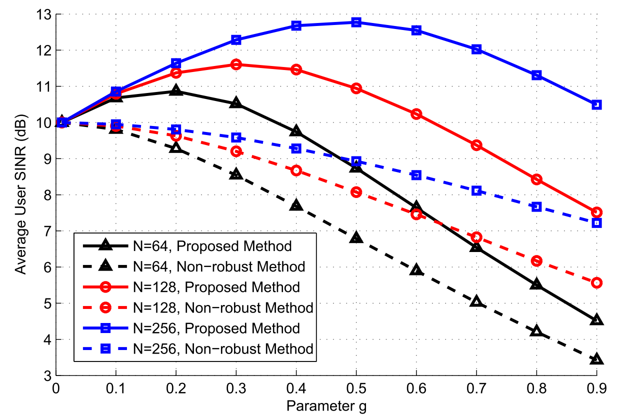


Fig. 6. Average user SINR versus g with different γ for $K = 30$ and $\gamma = 10$ dB.

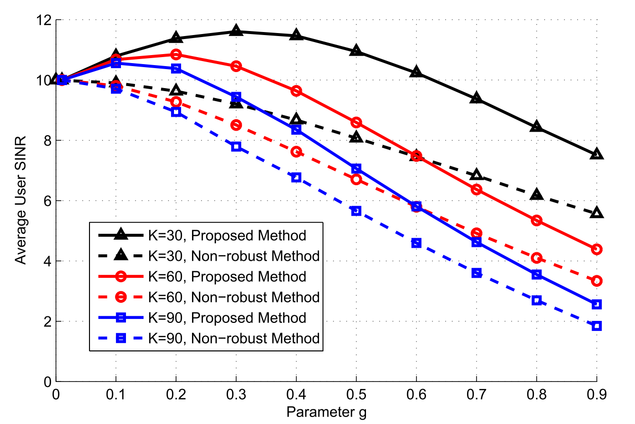


Fig. 7. Average user SINR versus g with different γ for $N = 128$ and $\gamma = 10$ dB.

function respect to parameter g. Moreover, the maximum average user SINR is equal to γ when $g = 0$, which says that the non-robust method will always obtain infeasible solutions of **P2** when $g > 0$. Interestingly, the average user SINR obtained by the proposed method will first increase with g increasing, and then decrease with g increasing. This phenomenon can be explained using Fig. 2, where it is clear that **P2** is infeasible when $g \geq 0.2$ (respect to $\gamma = 15$ dB), $g \geq 0.3$ (respect to $\gamma = 10$ dB) and $g \geq 0.5$ (respect to

[5] Note that when **P2** is infeasible, the solutions derived from Theorem 1, Theorem 2 and (48) for the proposed method are still existed but nontrivial. Nevertheless, the nontrivial solutions can be served as a suboptimal transmit strategy when **P2** is infeasible.

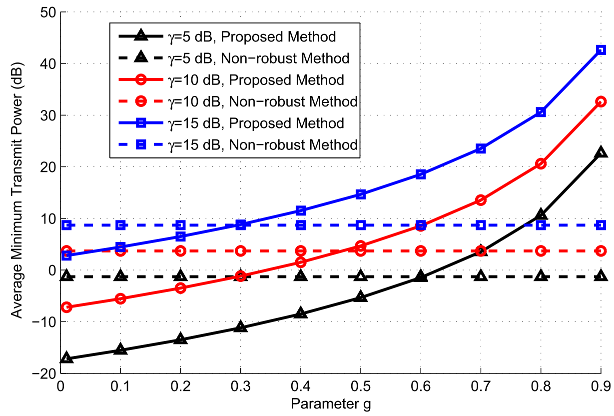


Fig. 8. Average minimum transmit power versus g with different N for $K = 30$ and $N = 128$.

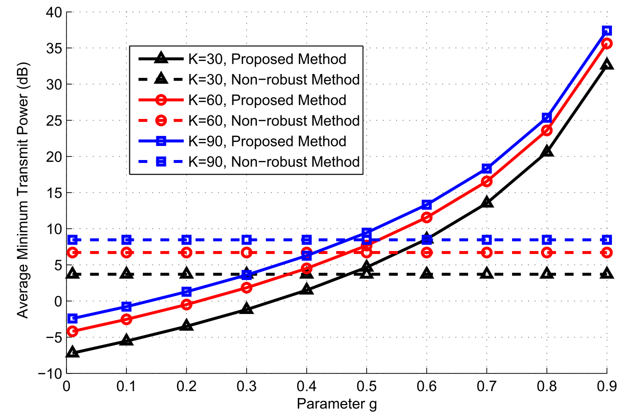


Fig. 10. Average minimum transmit power versus g with different N for $N = 128$ and $\gamma = 10$ dB.

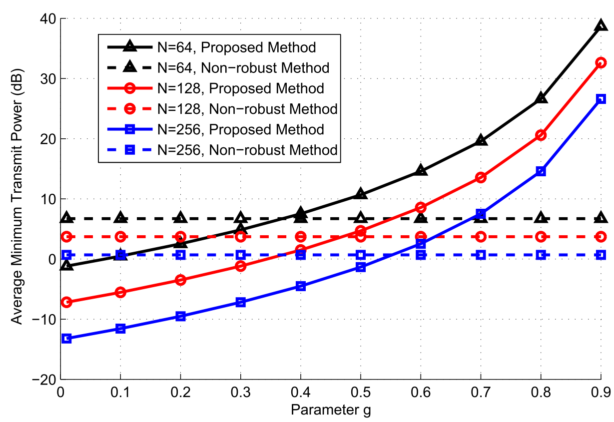


Fig. 9. Average minimum transmit power versus g with different N for $K = 30$ and $\gamma = 10$ dB.

$\gamma = 5$ dB). We know from Fig. 2 and Fig. 5 that the average user SINR obtained by the proposed method will increase with g when **P2** is feasible, and will decrease with g when **P2** is infeasible. Similarly, it is obvious in Fig. 6 and Fig. 7 that the average user SINR obtained by the non-robust method is a moronically decreasing function respect to N and K, where the derived solutions for **P2** is always infeasible when $g > 0$. Comparing Fig. 3 and Fig. 6, Fig. 4 and Fig. 7 respectively, we know that when **P2** is feasible, increasing N or decreasing K will help to improve the average user SINR obtained by the proposed method.

In the last example, we plot the average minimum transmit power versus parameter g with different γ, N and K for the proposed method and the non-robust method in Fig. 8, Fig. 9 and Fig. 10, respectively. It is obvious that the transmit power derived by the non-robust method will not change with parameter g. This is mainly due to the fact that the channel estimated errors will not affect the optimal solutions of the non-robust method. This phenomenon also explains the trends of the SINR performance for the non-robust method in Fig. 5, Fig. 6 and Fig. 7. Moreover, it is clear that the transmit power derived by the proposed method will moronically increase with g. The reason lies in that when g increases, the proposed method needs more power to eliminate the infeasible case. However, comparing the solutions derived in Fig. 5, Fig. 6 and Fig. 7 with that in Fig. 8, Fig. 9 and Fig. 10, respectively, it is implied that increasing the transmit power will not further improve the SINR performance when **P2** is infeasible (i.e., when g is big enough). Interestingly, comparing the solutions derived in Fig. 2, Fig. 3 and Fig. 4 with that in Fig. 8, Fig. 9 and Fig. 10, respectively, it is obvious that the average minimum transmit power of the proposed method is always smaller than that of the non-robust method when **P2** is feasible.

The above numerical results imply that we could choose proper γ and K to satisfy the feasibility conditions of **P2**. More importantly, it is always better to exploit more antennas at BS, which will bring smaller worst-case outage probability, better SINR performance and lower power consumption.

VI. Conclusions

In this paper, we designed a worst-case robust beamforming algorithm for multiuser massive MIMO systems with BDMA scheme. The target is to minimize the transmit power of BS provided that the intended users are satisfied with bounded SINR for all possible channel realizations. The conventional method for a regular MIMO system is NP-hard whose optimal solutions are generally unavailable. Nevertheless, the orthogonal property for BDMA massive MIMO channels can be exploited to equivalently transfer the initial non-convex problem as an SDP, where rank-one solutions are guaranteed. More importantly, we further derive the global optimal solutions in closed-form, which will greatly reduces the computational complexity and makes the proposed algorithm feasible for practical applications. Simulation results are provided to corroborate the proposed studies.

Appendix A
Proof of Proposition 1

Let us first show that $\mu_k^\star > 0$, $\forall k \in \{1, \ldots, K\}$ must hold from contradiction. Assuming $\mu_k^\star = 0$ for some k that $k \in \{1, \ldots, K\}$, it follows from (18) that

$$\boldsymbol{\Gamma}_k = \begin{bmatrix} \boldsymbol{X}_k & \boldsymbol{X}_k \tilde{\boldsymbol{h}}_k \\ \tilde{\boldsymbol{h}}_k^{\mathrm{H}} \boldsymbol{X}_k^{\mathrm{H}} & \tilde{\boldsymbol{h}}_k^{\mathrm{H}} \boldsymbol{X}_k \tilde{\boldsymbol{h}}_k - \sigma_k^2 \end{bmatrix} \succeq \boldsymbol{0}, \qquad \text{(A.1)}$$

must be satisfied. Left and right multiplying both sides of $\mathbf{\Gamma}_k$ by $[-\tilde{\boldsymbol{h}}_k^{\mathrm{H}}\ 1]$ and $[-\tilde{\boldsymbol{h}}_k^{\mathrm{H}}\ 1]^{\mathrm{H}}$, respectively, yields

$$[-\tilde{\boldsymbol{h}}_k^{\mathrm{H}} \quad 1]\mathbf{\Gamma}_k[-\tilde{\boldsymbol{h}}_k^{\mathrm{H}} \quad 1]^{\mathrm{H}} = -\sigma_k^2 \geq 0, \tag{A.2}$$

which cannot be true due to $\sigma_k^2 > 0$. Thus, at the optimal point, $\mu_k^\star > 0$, $\forall k \in \{1, \ldots, K\}$ must be satisfied in $\mathbf{P1-SDR}$.

Secondly, assume $\mathrm{Rank}\left(\boldsymbol{X}_k^\star + \mu_k^\star \boldsymbol{I}\right) = 0$, which implies that $\boldsymbol{X}_k^\star + \mu_k^\star \boldsymbol{I} = \mathbf{0}$. Since $\mu_k^\star > 0$, $\forall k \in \{1, \ldots, K\}$ hold, there is $\boldsymbol{X}_k^\star = -\mu_k^\star \boldsymbol{I} \prec \mathbf{0}$. Then it follows from (18) that

$$\tilde{\boldsymbol{h}}_k^{\mathrm{H}} \boldsymbol{X}_k^\star \tilde{\boldsymbol{h}}_k - \sigma_k^2 - \mu_k^\star \epsilon_k^2 < -\sigma_k^2 - \mu_k^\star \epsilon_k^2 < 0 \tag{A.3}$$

is true, which contradicts with the constraint (18) that

$$\begin{bmatrix} \boldsymbol{X}_k^\star + \mu_k^\star \boldsymbol{I} & \boldsymbol{X}_k^\star \tilde{\boldsymbol{h}}_k \\ \tilde{\boldsymbol{h}}_k^{\mathrm{H}} \boldsymbol{X}_k^{\star \mathrm{H}} & \tilde{\boldsymbol{h}}_k^{\mathrm{H}} \boldsymbol{X}_k^\star \tilde{\boldsymbol{h}}_k - \sigma_k^2 - \mu_k^\star \epsilon_k^2 \end{bmatrix} \succeq \mathbf{0}. \tag{A.4}$$

Thus, at the optimal point, there must be $\boldsymbol{X}_k^\star + \mu_k^\star \boldsymbol{I} \neq \mathbf{0}$. Moreover, since $\boldsymbol{X}_k^\star + \mu_k^\star \boldsymbol{I} \succeq \mathbf{0}$ should be satisfied, we obtain $\mathrm{Rank}\left(\boldsymbol{X}_k^\star + \mu_k^\star \boldsymbol{I}\right) \geq 1$.

Based on the above discussions, the proof is completed.

Appendix B
Proof of Proposition 3

A. Proof of Part (a)

The Lagrange of $\mathbf{P1-SDR}$ is defined as

$$\begin{aligned} &\mathcal{L}\left(\{\mu_k\}, \{\boldsymbol{S}_k\}, \left\{\begin{bmatrix} \boldsymbol{Z}_k & \boldsymbol{z}_k \\ \boldsymbol{z}_k^{\mathrm{H}} & \eta_k \end{bmatrix}\right\}, \{\Phi_{\mu,k}\}, \{\mathbf{\Phi}_{s,k}\}\right) \\ &= \sum_{k=1}^{K} \Big\{ \mathrm{Tr}(\boldsymbol{S}_k) - \mathrm{Tr}\left[\boldsymbol{Z}_k\left(\boldsymbol{X}_k + \mu_k \boldsymbol{I}\right) + \boldsymbol{z}_k \tilde{\boldsymbol{h}}_k^{\mathrm{H}} \boldsymbol{X}_k^{\mathrm{H}}\right] \\ &\quad - \left[\boldsymbol{z}_k^{\mathrm{H}} \boldsymbol{X}_k \tilde{\boldsymbol{h}}_k + \eta_k\left(\tilde{\boldsymbol{h}}_k^{\mathrm{H}} \boldsymbol{X}_k \tilde{\boldsymbol{h}}_k - \sigma_k^2 - \mu_k \epsilon_k^2\right)\right]\Big\} \\ &\quad - \sum_{k=1}^{K} \Phi_{\mu,k}\mu_k - \sum_{k=1}^{K} \mathrm{Tr}(\mathbf{\Phi}_{s,k}\boldsymbol{S}_k), \end{aligned} \tag{B.1}$$

where

$$\mathbf{\Gamma}_k = \begin{bmatrix} \boldsymbol{Z}_k & \boldsymbol{z}_k \\ \boldsymbol{z}_k^{\mathrm{H}} & \eta_k \end{bmatrix} \succeq \mathbf{0} \tag{B.2}$$

is the dual variable associated with the kth constraint in (18), while $\Phi_{\mu,k} \geq 0$ and $\mathbf{\Phi}_{s,k} \succeq \mathbf{0}$ are the dual variables associated with the kth constraint in (20). Using the definition in (B.2) and substituting (19) into (B.1), we further obtain

$$\begin{aligned} &\mathcal{L}\left(\{\mu_k\}, \{\boldsymbol{S}_k\}, \{\mathbf{\Gamma}_k\}, \{\Phi_{\mu,k}\}, \{\mathbf{\Phi}_{s,k}\}\right) \\ &= \sum_{k=1}^{K} \mathrm{Tr}\left(\boldsymbol{A}_{s,k}\boldsymbol{S}_k\right) + \sum_{k=1}^{K} B_{\mu,k}\mu_k + \sum_{k=1}^{K} \eta_k \sigma_k^2, \end{aligned} \tag{B.3}$$

where for simplicity, we define $\boldsymbol{A}_{s,k}$ and $B_{\mu,k}$ as

$$\begin{aligned} \boldsymbol{A}_{s,k} &= \boldsymbol{I} - \frac{1}{\gamma_k}\left(\boldsymbol{Z}_k + \boldsymbol{z}_k \tilde{\boldsymbol{h}}_k^{\mathrm{H}} + \tilde{\boldsymbol{h}}_k \boldsymbol{z}_k^{\mathrm{H}} + \eta_k \tilde{\boldsymbol{h}}_k \tilde{\boldsymbol{h}}_k^{\mathrm{H}}\right) \\ &\quad + \sum_{i \neq k}\left(\boldsymbol{Z}_i + \boldsymbol{z}_i \tilde{\boldsymbol{h}}_i^{\mathrm{H}} + \tilde{\boldsymbol{h}}_i \boldsymbol{z}_i^{\mathrm{H}} + \eta_i \tilde{\boldsymbol{h}}_i \tilde{\boldsymbol{h}}_i^{\mathrm{H}}\right) - \mathbf{\Phi}_{s,k}, \end{aligned} \tag{B.4}$$

$$B_{\mu,k} = \eta_k \epsilon_k^2 - \mathrm{Tr}(\boldsymbol{Z}_k) - \Phi_{\mu,k}, \tag{B.5}$$

respectively. Then the Karush-Kuhn-Tucker (KKT) conditions related to $\{\boldsymbol{S}_k\}$ and $\{\mu_k\}$ can be formulated as

$$\begin{aligned} &\boldsymbol{I} - \frac{1}{\gamma_k}\left(\boldsymbol{Z}_k^\star + \boldsymbol{z}_k^\star \tilde{\boldsymbol{h}}_k^{\mathrm{H}} + \tilde{\boldsymbol{h}}_k \boldsymbol{z}_k^{\star \mathrm{H}} + \eta_k \tilde{\boldsymbol{h}}_k \tilde{\boldsymbol{h}}_k^{\mathrm{H}}\right) \\ &\quad + \sum_{i \neq k}\left(\boldsymbol{Z}_i^\star + \boldsymbol{z}_i^\star \tilde{\boldsymbol{h}}_i^{\mathrm{H}} + \tilde{\boldsymbol{h}}_i \boldsymbol{z}_i^{\star \mathrm{H}} + \eta_i \tilde{\boldsymbol{h}}_i \tilde{\boldsymbol{h}}_i^{\mathrm{H}}\right) = \mathbf{\Phi}_{s,k}^\star \succeq \mathbf{0}, \end{aligned} \tag{B.6}$$

$$\eta_k \epsilon_k^2 - \mathrm{Tr}(\boldsymbol{Z}_k^\star) = \Phi_{\mu,k}^\star \geq 0, \tag{B.7}$$

$$\mathbf{\Phi}_{s,k}^\star \boldsymbol{S}_k^\star = \mathbf{0}, \quad \Phi_{\mu,k}^\star \mu_k^\star = 0, \quad \forall k \in \{1, \ldots, K\}, \tag{B.8}$$

where $\{\mu_k^\star\}$ and $\{\boldsymbol{S}_k^\star\}$ are the optimal primal variables, and $\mathbf{\Gamma}_k^\star$ is the optimal dual variable with

$$\mathbf{\Gamma}_k^\star = \begin{bmatrix} \boldsymbol{Z}_k^\star & \boldsymbol{z}_k^\star \\ \boldsymbol{z}_k^{\star \mathrm{H}} & \eta_k^\star \end{bmatrix} \succeq \mathbf{0}. \tag{B.9}$$

It follows from (B.6) that $\forall j \neq k$ and $j, k \in \{1, \ldots, K\}$, there must be

$$\begin{aligned} &\mathbf{\Phi}_{s,j}^\star + \mathbf{\Phi}_{s,k}^\star \\ &= 2\boldsymbol{I} + \left(1 - \frac{1}{\gamma_j}\right)\left(\boldsymbol{Z}_j^\star + \boldsymbol{z}_j^\star \tilde{\boldsymbol{h}}_j^{\mathrm{H}} + \tilde{\boldsymbol{h}}_j \boldsymbol{z}_j^{\star \mathrm{H}} + \eta_j \tilde{\boldsymbol{h}}_j \tilde{\boldsymbol{h}}_j^{\mathrm{H}}\right) \\ &\quad + \left(1 - \frac{1}{\gamma_k}\right)\left(\boldsymbol{Z}_k^\star + \boldsymbol{z}_k^\star \tilde{\boldsymbol{h}}_k^{\mathrm{H}} + \tilde{\boldsymbol{h}}_k \boldsymbol{z}_k^{\star \mathrm{H}} + \eta_k \tilde{\boldsymbol{h}}_k \tilde{\boldsymbol{h}}_k^{\mathrm{H}}\right) \\ &\quad + 2\sum_{i \neq j,k}\left(\boldsymbol{Z}_i^\star + \boldsymbol{z}_i^\star \tilde{\boldsymbol{h}}_i^{\mathrm{H}} + \tilde{\boldsymbol{h}}_i \boldsymbol{z}_i^{\star \mathrm{H}} + \eta_i \tilde{\boldsymbol{h}}_i \tilde{\boldsymbol{h}}_i^{\mathrm{H}}\right). \end{aligned} \tag{B.10}$$

Moreover, left and right multiplying both sides of (B.9) by $[\boldsymbol{I}\ \tilde{\boldsymbol{h}}_k]$ and $[\boldsymbol{I}\ \tilde{\boldsymbol{h}}_k]^{\mathrm{H}}$, respectively, yields

$$\begin{aligned} [\boldsymbol{I} \quad \tilde{\boldsymbol{h}}_k]\mathbf{\Gamma}_k^\star[\boldsymbol{I} \quad \tilde{\boldsymbol{h}}_k]^{\mathrm{H}} &= [\boldsymbol{I} \quad \tilde{\boldsymbol{h}}_k]\begin{bmatrix} \boldsymbol{Z}_k^\star & \boldsymbol{z}_k^\star \\ \boldsymbol{z}_k^{\star \mathrm{H}} & \eta_k^\star \end{bmatrix}[\boldsymbol{I} \quad \tilde{\boldsymbol{h}}_k]^{\mathrm{H}} \\ &= \boldsymbol{Z}_k^\star + \boldsymbol{z}_k^\star \tilde{\boldsymbol{h}}_k^{\mathrm{H}} + \tilde{\boldsymbol{h}}_k \boldsymbol{z}_k^{\star \mathrm{H}} + \eta_k \tilde{\boldsymbol{h}}_k \tilde{\boldsymbol{h}}_k^{\mathrm{H}} \succeq \mathbf{0}. \end{aligned} \tag{B.11}$$

From (B.10) and (B.11), we obtain

$$\mathbf{\Phi}_{s,j}^\star + \mathbf{\Phi}_{s,k}^\star \succeq 2\boldsymbol{I}, \tag{B.12}$$

which says that $\mathrm{Rank}\left(\mathbf{\Phi}_{s,j}^\star + \mathbf{\Phi}_{s,k}^\star\right) = N$, $\forall j \neq k$ and $j, k \in \{1, \ldots, K\}$.

Secondly, it follows from (B.3)∼(B.9) that the dual problem of $\mathbf{P1-SDR}$ can be formulated as

$\mathbf{P1-DUAL}$:

$$\max_{\{\boldsymbol{Z}_k\},\{\boldsymbol{z}_k\},\{\eta_k\}} \quad \sum_{k=1}^{K} \eta_k \sigma_k^2 \tag{B.13}$$

$$\begin{aligned} \text{s.t. } \mathbf{\Phi}_{s,k} &= \boldsymbol{I} - \frac{1}{\gamma_k}\left(\boldsymbol{Z}_k + \boldsymbol{z}_k \tilde{\boldsymbol{h}}_k^{\mathrm{H}} + \tilde{\boldsymbol{h}}_k \boldsymbol{z}_k^{\mathrm{H}} + \eta_k \tilde{\boldsymbol{h}}_k \tilde{\boldsymbol{h}}_k^{\mathrm{H}}\right) \\ &\quad + \sum_{i \neq k}\left(\boldsymbol{Z}_i + \boldsymbol{z}_i \tilde{\boldsymbol{h}}_i^{\mathrm{H}} + \tilde{\boldsymbol{h}}_i \boldsymbol{z}_i^{\mathrm{H}} + \eta_i \tilde{\boldsymbol{h}}_i \tilde{\boldsymbol{h}}_i^{\mathrm{H}}\right) \succeq \mathbf{0}, \end{aligned} \tag{B.14}$$

$$\Phi_{\mu,k} = \eta_k \epsilon_k^2 - \mathrm{Tr}(\boldsymbol{Z}_k) \geq 0, \tag{B.15}$$

$$\mathbf{\Gamma}_k = \begin{bmatrix} \boldsymbol{Z}_k & \boldsymbol{z}_k \\ \boldsymbol{z}_k^{\mathrm{H}} & \eta_k \end{bmatrix} \succeq \mathbf{0}, \quad \forall k \in \{1, \ldots, K\}. \tag{B.16}$$

Denote $\{\boldsymbol{Z}_k^\star\}$, $\{\boldsymbol{z}_k^\star\}$ and $\{\eta_k^\star\}$ as the optimal solutions of $\mathbf{P1-DUAL}$.

Let us show that $\text{Rank}\left(\mathbf{\Phi}^{\star}_{s,k}\right) < N$ from contradiction. Assuming $\text{Rank}\left(\mathbf{\Phi}^{\star}_{s,k}\right) = N$ or $\mathbf{\Phi}^{\star}_{s,k} \succ \mathbf{0}$, we can always find an alternative η^{*}_{k} satisfying $\eta^{*}_{k} > \eta^{\star}_{k}$ and $\mathbf{\Phi}^{*}_{s,k}\left(\{\mathbf{Z}^{\star}_{k}\}, \{\mathbf{z}^{\star}_{k}\}, \{\eta^{*}_{k}\}\right) \succeq \mathbf{0}$ simultaneously. Then substituting $\{\mathbf{Z}^{\star}_{k}\}$, $\{\mathbf{z}^{\star}_{k}\}$ and $\{\eta^{*}_{k}\}$ into **P1 − DUAL**, we obtain

$$\sum_{i=1}^{K} \eta^{*}_{i}\sigma^{2}_{i} > \sum_{k=1}^{K} \eta^{\star}_{k}\sigma^{2}_{k} \tag{B.17}$$

$$\begin{aligned}\mathbf{\Phi}^{*}_{s,k} = \mathbf{I} - \frac{1}{\gamma_k}\left(\mathbf{Z}^{\star}_{k} + \mathbf{z}^{\star}_{k}\tilde{\mathbf{h}}^{\mathrm{H}}_{k} + \tilde{\mathbf{h}}_{k}\mathbf{z}^{\star\mathrm{H}}_{k} + \eta^{*}_{k}\tilde{\mathbf{h}}_{k}\tilde{\mathbf{h}}^{\mathrm{H}}_{k}\right) \\ + \sum_{i\neq k}\left(\mathbf{Z}^{\star}_{i} + \mathbf{z}^{\star}_{i}\tilde{\mathbf{h}}^{\mathrm{H}}_{i} + \tilde{\mathbf{h}}_{i}\mathbf{z}^{\star\mathrm{H}}_{i} + \eta^{*}_{i}\tilde{\mathbf{h}}_{i}\tilde{\mathbf{h}}^{\mathrm{H}}_{i}\right) \succeq \mathbf{0},\end{aligned} \tag{B.18}$$

$$\Phi^{*}_{\mu,k} = \eta^{*}_{k}\epsilon^{2}_{k} - \text{Tr}(\mathbf{Z}^{\star}_{k}) > \eta^{\star}_{k}\epsilon^{2}_{k} - \text{Tr}(\mathbf{Z}^{\star}_{k}) = \Phi^{\star}_{\mu,k} \geq 0, \tag{B.19}$$

$$\mathbf{\Gamma}^{*}_{k} = \begin{bmatrix} \mathbf{Z}^{\star}_{k} & \mathbf{z}^{\star}_{k} \\ \mathbf{z}^{\star\mathrm{H}}_{k} & \eta^{*}_{k}\end{bmatrix} \succeq \begin{bmatrix} \mathbf{Z}^{\star}_{k} & \mathbf{z}^{\star}_{k} \\ \mathbf{z}^{\star\mathrm{H}}_{k} & \eta^{\star}_{k}\end{bmatrix} = \mathbf{\Gamma}^{\star}_{k} \succeq \mathbf{0}, \tag{B.20}$$

$$\forall k \in \{1, \ldots, K\}. \tag{B.21}$$

It can be easily checked that $\{\eta^{*}_{k}\}$ will not violate any constraints, but will always increase the objective value. Thus η^{*}_{k} is a better solution than $\eta^{\star}_{k}$, which contradicts the assumption that $\eta^{\star}_{k}$ is optimal. Consequently, we know that at the optimal point,

$$\text{Rank}\left(\mathbf{\Phi}^{\star}_{s,k}\right) < N, \quad \forall k \in \{1, \ldots, K\}, \tag{B.22}$$

must hold.

Let us then assume $\text{Rank}\left(\mathbf{\Phi}^{\star}_{s,k}\right) = D_k < N$ and $\text{Rank}\left(\mathbf{\Phi}^{\star}_{s,j}\right) = D_j < N$. Denote $\{\mathbf{u}_p\}$, $p \in \{1, \ldots, D_k\}$ and $\{\mathbf{u}'_q\}$, $q \in \{1, \ldots, D_j\}$ as the orthogonal bases of $\mathbf{\Phi}^{\star}_{s,k}$ and $\mathbf{\Phi}^{\star}_{s,j}$, respectively. Let $\{\bar{\mathbf{u}}_{\bar{p}}\}$ where $\bar{p} \in \{1, \ldots, N - D_k\}$ and $\{\bar{\mathbf{u}}'_{\bar{q}}\}$ where $\bar{q} \in \{1, \ldots, N - D_j\}$ are the orthogonal bases of the null spaces of $\mathbf{\Phi}^{\star}_{s,k}$ and $\mathbf{\Phi}^{\star}_{s,j}$, respectively. Since $\text{Rank}\left(\mathbf{\Phi}^{\star}_{s,j} + \mathbf{\Phi}^{\star}_{s,k}\right) = N$ (c.f. Eq.(B.12) for details), there must be

$$\text{Span}\{\bar{\mathbf{u}}_1, \ldots, \bar{\mathbf{u}}_{N-D_k}\} \subseteq \text{Span}\{\mathbf{u}'_1, \ldots, \mathbf{u}'_{D_j}\}, \tag{B.23}$$

which says that the null space of $\mathbf{\Phi}^{\star}_{s,k}$ lies in the range space of $\mathbf{\Phi}^{\star}_{s,j}$, or equivalently,

$$\text{Span}\{\bar{\mathbf{u}}_1, \ldots, \bar{\mathbf{u}}_{N-D_k}\} \perp \text{Span}\{\bar{\mathbf{u}}'_1, \ldots, \bar{\mathbf{u}}'_{N-D_j}\}. \tag{B.24}$$

Moreover, it follows from (B.8) that $\mathbf{\Phi}^{\star}_{s,k}\mathbf{S}^{\star}_{k} = \mathbf{0}$, $\forall k \in \{1, \ldots, K\}$ must hold. Thus $\mathbf{S}^{\star}_{k}$ lies in the null space of $\mathbf{\Phi}^{\star}_{s,k}$ and $\mathbf{S}^{\star}_{j}$ lies in the null space of $\mathbf{\Phi}^{\star}_{s,j}$. Using (B.24) there is

$$\mathbf{S}^{\star}_{j}\mathbf{S}^{\star}_{k} = \mathbf{0}, \quad \forall j \neq k \in \{1, \ldots, K\}. \tag{B.25}$$

B. Proof of Part (b)

Let us show that there exists at least one nonzero eigenvalue for $\mathbf{S}^{\star}_{k}$ from contradiction. Assuming all the eigenvalues of $\mathbf{S}^{\star}_{k}$ are zero, i.e., $\mathbf{S}^{\star}_{k} = \mathbf{0}$ for some k, $k \in \{1, \ldots, K\}$, it follows from (19) that

$$\mathbf{X}^{\star}_{k} = -\sum_{i\neq k}\mathbf{S}^{\star}_{i} \preceq \mathbf{0}. \tag{B.26}$$

Then using part (a) of Proposition 1 and from (18) we know

$$\tilde{\mathbf{h}}^{\mathrm{H}}_{k}\mathbf{X}^{\star}_{k}\tilde{\mathbf{h}}_{k} - \sigma^{2}_{k} - \mu^{\star}_{k}\epsilon^{2}_{k} < -\sigma^{2}_{k} - \mu^{\star}_{k}\epsilon^{2}_{k} < 0, \tag{B.27}$$

which cannot be true due to $\tilde{\mathbf{h}}^{\mathrm{H}}_{k}\mathbf{X}^{\star}_{k}\tilde{\mathbf{h}}_{k} - \sigma^{2}_{k} - \mu^{\star}_{k}\epsilon^{2}_{k} \geq 0$. Thus there must be at least one nonzero eigenvalue for $\mathbf{S}^{\star}_{k}$, which implies that $\text{Rank}\left(\mathbf{S}^{\star}_{k}\right) \geq 1$.

The proof of Proposition 3 is completed.

APPENDIX C
PROOF OF THEOREM 1

Let us show that $\text{Rank}\left(\mathbf{S}^{\star}_{k}\right) = 1$, $\forall k \in \{1, \ldots, K\}$ from contradiction. Assume $\{\mu^{\star}_{k}\}$ and $\{\mathbf{S}^{\star}_{k}\}$ are the optimal solutions of **P1 − SDR**, where $\text{Rank}\left(\mathbf{S}^{\star}_{k}\right) = C_k > 1$. Since the optimal solutions satisfy $\mathbf{S}^{\star}_{j}\mathbf{S}^{\star}_{k} = \mathbf{0}$, $\forall j \neq k$ and $\text{Rank}\left(\mathbf{S}^{\star}_{k}\right) \geq 1$, $\forall k \in \{1, \ldots, K\}$ (c.f. Proposition 3 for details), the EVD of $\mathbf{X}^{\star}_{k}$ can be further expressed as

$$\mathbf{X}^{\star}_{k} = \frac{1}{\gamma_k}\mathbf{S}^{\star}_{k} - \sum_{i\neq k}\mathbf{S}^{\star}_{i} = \mathbf{U}^{\star}_{x,k}\begin{bmatrix}\mathbf{\Lambda}^{\star}_{k} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & -\mathbf{\Lambda}^{\star}_{i\neq k}\end{bmatrix}\mathbf{U}^{\star\mathrm{H}}_{x,k}, \tag{C.1}$$

where $\mathbf{\Lambda}^{\star}_{k}$ is a positive diagonal matrix with the diagonal eigenvalues $p^{\star}_{k,1}/\gamma_k \geq \ldots \geq p^{\star}_{k,C_k}/\gamma_k > 0$ related to $\mathbf{S}^{\star}_{k}/\gamma_k$, and $-\mathbf{\Lambda}^{\star}_{i\neq k}$ is a negative diagonal matrix with the diagonal eigenvalues $0 > -p^{\star}_{i\neq k,1} \geq \ldots \geq -p^{\star}_{i\neq k,\sum_{i\neq k}C_k}$ related to $-\sum_{i\neq k}\mathbf{S}^{\star}_{i}$.

Next, let $p^{\star}_{k,2} = \ldots = p^{\star}_{k,c_k} = 0$, $\forall k \in \{1, \ldots, K\}$. Then we can construct the following new solutions

$$\mathbf{S}^{*}_{k} = p^{\star}_{k,1}\frac{\tilde{\mathbf{h}}_{k}\tilde{\mathbf{h}}^{\mathrm{H}}_{k}}{\|\tilde{\mathbf{h}}_{k}\|^2}, \quad \forall k \in \{1, \ldots, K\}, \tag{C.2}$$

which satisfies $\text{Rank}\left(\mathbf{S}^{*}_{k}\right) = 1$. With BDMA massive MIMO regime, the estimated channel vectors satisfy $\tilde{\mathbf{h}}_{i} \perp \tilde{\mathbf{h}}_{k}$, $\forall i \neq k$. The EVD of $\mathbf{X}^{*}_{k}$ can be expressed as

$$\begin{aligned}\mathbf{X}^{*}_{k} &= \frac{1}{\gamma_k}\mathbf{S}^{*}_{k} - \sum_{i\neq k}\mathbf{S}^{*}_{i} = \frac{p^{\star}_{k,1}}{\gamma_k}\tilde{\mathbf{h}}_{k}\tilde{\mathbf{h}}^{\mathrm{H}}_{k} - \sum_{i\neq k}p^{\star}_{i,1}\tilde{\mathbf{h}}_{i}\tilde{\mathbf{h}}^{\mathrm{H}}_{i} \\ &= \mathbf{U}^{*}_{x,k}\begin{bmatrix} p^{\star}_{k,1}/\gamma_k & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \text{diag}(\{-p^{\star}_{i,1}, i \neq k\})\end{bmatrix}\mathbf{U}^{*\mathrm{H}}_{x,k},\end{aligned} \tag{C.3}$$

where $\{-p^{\star}_{i,1}, i \neq k\}$ in decreasing order are the $K - 1$ eigenvalues related to $\{\mathbf{S}^{*}_{i}, i \neq k\}$. Denote $\mathbf{u}^{*}_{x,k,i}$ as the ith column of $\mathbf{U}^{*}_{x,k}$. It follows from (C.3) that there must be

$$\mathbf{u}^{*}_{x,k,1} = \frac{\tilde{\mathbf{h}}_{k}}{\|\tilde{\mathbf{h}}_{k}\|}, \tag{C.4}$$

which says that (28) holds with equality. Moreover, since $\mathbf{X}^{\star}_{k} + \mu_k\mathbf{I} \succeq \mathbf{0}$ should be satisfied, it is easily known from (C.1) and (C.3) that $\mathbf{X}^{*}_{k} + \mu_k\mathbf{I} \succeq \mathbf{0}$ is true.

Noting that in (22), $q_{k,1}$ is the maximum eigenvalue of $\mathbf{X}_k$, and in (C.1), $p^{\star}_{k,1}$ is the maximum eigenvalue of $\mathbf{S}^{\star}_{k}$. Thus, there must be $p^\star_{k,1} = q_{k,1}\gamma_k$. Then substituting $\{\mu^\star_k\}$ and $\{\boldsymbol{S}^*_k\}$ into $\mathbf{P1-SDR}$, we obtain

$$\sum_{k=1}^{K}\mathrm{Tr}(\boldsymbol{S}^*_k) = \sum_{k=1}^{K} p^\star_{k,1} < \sum_{k=1}^{K}\sum_{j=1}^{C_k} p^\star_{k,j} = \sum_{k=1}^{K}\mathrm{Tr}(\boldsymbol{S}^\star_k) \tag{C.5}$$

$$\begin{aligned}\mathrm{Tr}&\left\{\left(\tilde{\boldsymbol{H}}_k\boldsymbol{X}^*_k\right)\left[\boldsymbol{I}-\left(\boldsymbol{X}^*_k+\mu_k\boldsymbol{I}\right)^\dagger\boldsymbol{X}^*_k\right]\right\}-\sigma_k^2-\mu_k\epsilon_k^2\\ &=\frac{\mu^\star_k q_{k,1}\lambda_{k,1}}{\mu_k+p_{k,1}}-\sigma_k^2-\mu_k\epsilon_k^2\geq-\sigma_k^2-\mu^\star_k\epsilon_k^2\\ &\quad+\mathrm{Tr}\left\{\left(\tilde{\boldsymbol{H}}_k\boldsymbol{X}^\star_k\right)\left[\boldsymbol{I}-\left(\boldsymbol{X}^\star_k+\mu_k\boldsymbol{I}\right)^\dagger\boldsymbol{X}^\star_k\right]\right\}>0,\end{aligned} \tag{C.6}$$

$$\boldsymbol{X}^*_k+\mu_k\boldsymbol{I}\succeq\boldsymbol{0},\quad \boldsymbol{X}^*_k=\frac{1}{\gamma_k}\boldsymbol{S}^*_k-\sum_{i\neq k}\boldsymbol{S}^*_i, \tag{C.7}$$

$$\mu^\star_k\geq 0,\quad \boldsymbol{S}^*_k\succeq\boldsymbol{0},\ k=1,2,\ldots,K. \tag{C.8}$$

From (C.6)$\sim$(C.8) we know that $\{\boldsymbol{S}^*_k\}$ satisfy all constraints of $\mathbf{P1-SDR}$. From (C.5) we know that $\{\boldsymbol{S}^*_k\}$ will always provide smaller objective value. Consequently, $\{\boldsymbol{S}^*_k\}$ are better solutions than $\{\boldsymbol{S}^\star_k\}$, which contradicts with our first place assumption.

Based on the above discussions, we know that at the optimal point there must be $\boldsymbol{S}^\star_k = p^\star_{k,1}\tilde{\boldsymbol{h}}_k\tilde{\boldsymbol{h}}_k^{\mathrm{H}}/\|\tilde{\boldsymbol{h}}_k\|^2$, i.e., $\mathrm{Rank}\left(\boldsymbol{S}^\star_k\right) = 1, \forall k\in\{1,\ldots,K\}$. The proof of Theorem 1 is thus completed.

References

[1] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[2] J. G. Andrews *et al.*, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.

[3] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

[4] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 742–758, Oct. 2014.

[5] E. Björnson, E. G. Larsson, and T. L. Marzetta, "Massive MIMO: Ten myths and one critical question," *IEEE Commun. Mag.*, vol. 54, no. 2, pp. 114–123, Feb. 2016.

[6] F. Gao, T. Cui, and A. Nallanathan, "On channel estimation and optimal training design for amplify and forward relay networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 5, pp. 1907–1916, May 2008.

[7] T. Cui, F. Gao, T. Ho, and A. Nallanathan, "Distributed space–time coding for two-way wireless relay networks," *IEEE Trans. Signal Process.*, vol. 57, no. 2, pp. 658–671, Feb. 2009.

[8] F. Gao, R. Zhang, and Y.-C. Liang, "Optimal channel estimation and training design for two-way relay networks," *IEEE Trans. Commun.*, vol. 57, no. 10, pp. 3024–3033, Oct. 2009.

[9] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 264–273, Feb. 2013.

[10] A. Adhikary, J. Nam, J.-Y. Ahn, and G. Caire, "Joint spatial division and multiplexing—The large-scale array regime," *IEEE Trans. Inf. Theory*, vol. 59, no. 10, pp. 6441–6463, Oct. 2013.

[11] H. Xie, B. Wang, F. Gao, and S. Jin, "A full-space spectrum-sharing strategy for massive MIMO cognitive radio systems," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 10, pp. 2537–2549, Oct. 2016.

[12] H. Xie, F. Gao, and S. Jin, "An overview of low-rank channel estimation for massive MIMO systems," *IEEE Access*, vol. 4, pp. 7313–7321, Nov. 2016.

[13] H. Xie, F. Gao, S. Zhang, and S. Jin, "A unified transmission strategy for TDD/FDD massive MIMO systems with spatial basis expansion model," *IEEE Trans. Veh. Technol.*, vol. 66, no. 4, pp. 3170–3184, Apr. 2016.

[14] W. Choi, A. Forenza, J. G. Andrews, and J. R. W. Heath, Jr., "Opportunistic space-division multiple access with beam selection," *IEEE Trans. Commun.*, vol. 55, no. 12, pp. 2371–2380, Dec. 2007.

[15] X. Gao, L. Dai, Z. Chen, Z. Wang, and Z. Zhang, "Near-optimal beam selection for beamspace mmwave massive MIMO systems," *IEEE Commun. Lett.*, vol. 20, no. 5, pp. 1054–1057, May 2016.

[16] L. You, X. Gao, G. Y. Li, X.-G. Xia, and N. Ma, "BDMA for millimeter-wave/terahertz massive MIMO transmission with per-beam synchronization," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 7, pp. 1550–1563, Jul. 2017.

[17] Z. Xiao, P. Xia, and X.-G. Xia, "Enabling UAV cellular with millimeter-wave communication: Potentials and approaches," *IEEE Commun. Mag.*, vol. 54, no. 5, pp. 66–73, May 2016.

[18] S. Verdú and H. V. Poor, "On minimax robustness: A general approach and applications," *IEEE Trans. Inf. Theory*, vol. 30, no. 2, pp. 328–340, Mar. 1984.

[19] S. A. Kassam and H. V. Poor, "Robust techniques for signal processing: A survey," *Proc. IEEE*, vol. 73, no. 3, pp. 433–481, Mar. 1985.

[20] J. Li, P. Stoica, and Z. Wang, "On robust Capon beamforming and diagonal loading," *IEEE Trans. Signal Process.*, vol. 51, no. 7, pp. 1702–1715, Jul. 2003.

[21] 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 Process.*, vol. 51, no. 2, pp. 313–324, Feb. 2003.

[22] R. G. Lorenz and S. P. Boyd, "Robust minimum variance beamforming," *IEEE Trans. Signal Process.*, vol. 53, no. 5, pp. 1684–1696, May 2005.

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

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

[25] X. Zhang, D. P. Palomar, and B. Ottersten, "Statistically robust design of linear MIMO transceivers," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3678–3689, Aug. 2008.

[26] 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, Nov. 2014.

[27] Z. Zhu, Z. Chu, Z. Wang, and I. Lee, "Outage constrained robust beamforming for secure broadcasting systems with energy harvesting," *IEEE Trans. Wireless Commun.*, vol. 15, no. 11, pp. 7610–7620, Nov. 2016.

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

[29] 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, Feb. 2009.

[30] 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.

[31] I. Wajid, M. Pesavento, Y. C. Eldar, and D. Ciochina, "Robust downlink beamforming with partial channel state information for conventional and cognitive radio networks," *IEEE Trans. Signal Process.*, vol. 61, no. 14, pp. 3656–3670, Jul. 2013.

[32] A. Mostafa and L. Lampe, "Optimal and robust beamforming for secure transmission in MISO visible-light communication links," *IEEE Trans. Signal Process.*, vol. 64, no. 24, pp. 6501–6516, Dec. 2016.

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

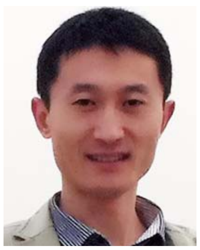
[34] M. F. Hanif, L.-N. Tran, A. Tolli, and M. Juntti, "Computationally efficient robust beamforming for SINR balancing in multicell downlink with applications to large antenna array systems," *IEEE Trans. Commun.*, vol. 62, no. 6, pp. 1908–1920, Jun. 2014.

[35] 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.

[36] X. Li, T. Jiang, S. Cui, J. An, and Q. Zhang, "Cooperative communications based on rateless network coding in distributed MIMO systems [coordinated and distributed MIMO]," *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 60–67, Jun. 2010.

[37] J. Brady, N. Behdad, and A. M. Sayeed, "Beamspace MIMO for millimeter-wave communications: System architecture, modeling, analysis, and measurements," *IEEE Trans. Antennas Propag.*, vol. 61, no. 7, pp. 3814–3827, Jul. 2013.

[38] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory* (Studies in Applied Mathematics). Philadelphia, PA, USA: SIAM, 1994.
[39] R. Penrose, "A generalized inverse for matrices," *Math. Proc. Cambridge Philos. Soc.*, vol. 51, no. 3, pp. 406–413, 1955.
[40] R. A. Horn and C. R. Johnson, *Matrix Analysis*. New York, NY, USA: Cambridge Univ. Press, 1985.
[41] A. W. Marshall, I. Olkin, and B. C. Arnold, *Inequalities: Theory of Majorization and Its Applications*. New York, NY, USA: Academic, 1979.
[42] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block coding for wireless communications: Performance results," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 3, pp. 451–460, Mar. 1999.

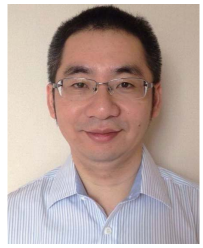


Fengchao Zhu received the B.E. degree in information engineering, the M.E. degree in signal and information processing, and the Ph.D. degree in signal and information processing from the Xian Research Institute of Hi-Tech, Xian, China, in 2008, 2011, and 2015, respectively. From 2011 to 2015, he was with the Department of Automation, Tsinghua University, Beijing, China, for the Ph.D. degree. His current research interests include convex optimization and deep learning, together with their applications in beamforming for massive MIMO systems, array signal processing, and physical layer security for wireless communications.

Feifei Gao (M'09–SM'14) received the B.Eng. degree from Xian Jiaotong University, Xi'an, China, in 2002, the M.Sc. degree from McMaster University, Hamilton, ON, Canada, in 2004, and the Ph.D. degree from the National University of Singapore, Singapore, in 2007. He was a Research Fellow with the Institute for Infocomm Research, A*STAR, Singapore, in 2008, and an Assistant Professor with the School of Engineering and Science, Jacobs University, Bremen, Germany, from 2009 to 2010. In 2011, he joined the Department of Automation, Tsinghua University, Beijing, China, where he is currently an Associate Professor.

His research areas include communication theory, signal processing for communications, array signal processing, and convex optimizations, with particular interests in MIMO techniques, multi-carrier communications, cooperative communication, and cognitive radio networks. He has authored/co-authored over 120 refereed IEEE journal papers and over 120 IEEE conference proceeding papers.

Prof. Gao has serves as the Symposium Co-Chair for the 2018 IEEE Vehicular Technology Conference Spring, the 2015 IEEE Conference on Communications, the 2014 IEEE Global Communications Conference, the 2014 IEEE Vehicular Technology Conference Fall, and a technical committee member for many other IEEE conferences. He has served as an Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, the IEEE SIGNAL PROCESSING LETTERS, the IEEE COMMUNICATIONS LETTERS, the IEEE WIRELESS COMMUNICATIONS LETTERS, the *International Journal on Antennas and Propagations*, and the *China Communications*.



Shi Jin (S'06–M'07) received the B.S. degree in communications engineering from the Guilin University of Electronic Technology, Guilin, China, in 1996, the M.S. degree from the Nanjing University of Posts and Telecommunications, Nanjing, China, in 2003, and the Ph.D. degree in communications and information systems from the Southeast University, Nanjing, in 2007. From 2007 to 2009, he was a Research Fellow with Adastral Park Research Campus, University College London, London, U.K. He is currently with the Faculty of the National Mobile Communications Research Laboratory, Southeast University. His research interests include space time wireless communications, random matrix theory, and information theory. He and his co-authors have been received the 2011 IEEE Communications Society Stephen O. Rice Prize Paper Award in the field of communication theory and the 2010 Young Author Best Paper Award by the IEEE Signal Processing Society. He serves as an Associate Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, and the IEEE COMMUNICATIONS LETTERS, and the *IET Communications*.

Hai Lin (M'05–SM'12) received the B.E. degree from Shanghai JiaoTong University, China, in 1993, the M.E. degree from the University of the Ryukyus, Japan, in 2000, and the Dr. Eng. degree from Osaka Prefecture University, Japan, in 2005. Since 2000, he has been a Research Associate with the Graduate School of Engineering, Osaka Prefecture University, where he is currently an Associate Professor.

His research interests are in signal processing for communications, wireless communications, and statistical signal processing. He has acted many times as technical program committee member for the IEEE ICC, GLOBECOM, WCNC, and VTC. He is a member of the IEICE. He served as a technical program co-chair for the Wireless Communications Symposium of the IEEE ICC 2011, the Signal Processing for Communications Symposium of the IEEE ICC 2013, and the Wireless Communications Symposium of the IEEE GLOBECOM 2013. He was an Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, and currently an Editor of the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY.

Minli Yao received the M.Sc. degree from the Xi'an Research Institute of High Technology, Xi'an, China, in 1992, and the Ph.D. degree from Xi'an Jiao Tong University, Xi'an China, in 1999, respectively. He is currently a Professor with the Department of Communication Engineering, Xi'an Research Institute of High Technology.

His current interests include satellite navigation signal structure design and signal processing, Sat-COM, and spreading communication.