This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2019.2916405, IEEE Transactions on Wireless Communications

Robust Simultaneous Wireless Information and Power Transfer in Beamspace Massive MIMO

Fengchao Zhu, Feifei Gao, Yonina C. Eldar, and Gongbin Qian

***Abstract*—We investigate worst-case robust beamforming for simultaneous wireless information and power transfer (SWIPT) in a multiuser beamspace massive multiple-input multiple-output (mMIMO) system. The objective is to minimize the transmit power of the base station (BS) subject to the individual signal-to-interference-plus-noise ratio (SINR) and the energy-harvesting constraints under imperfect channel state information (CSI). Instead of directly resorting to semi-definite relaxation (SDR), we convert the initial non-convex optimization to a power allocation problem, which greatly reduces the computational complexity. The beamforming vectors are proved as scaled versions of the estimated channels, and then the optimal scaling factors are derived in closed-form. Simulations demonstrate that the proposed robust beamforming method achieves the globally optimal point for the initial design when the channel estimation errors are small, and leads to satisfactory performance when the channel estimation errors are large.**

***Index Terms*—Simultaneous wireless information and power transfer (SWIPT), robust beamforming, massive MIMO, beamspace, non-convex optimization.**

I. Introduction

Simultaneous wireless information and power transfer (SWIPT) [1] has received considerable interest recently, since it can offer unlimited supplies to energy-constrained wireless networks. Many prominent works have studied the fundamental performance of SWIPT systems. For example, the rate-energy regions of multiple-input multiple-output (MIMO) channels with separate and co-located SWIPT receivers were characterized in [2]. An orthogonal frequency division multiplexing (OFDM)-based wireless powered communication system was investigated in [3]. The energy beamforming vector and time split parameter were designed for a power beacon assisted two-way relaying network in [4]. An artificial noise (AN) assisted interference alignment (IA) scheme with wireless power transfer is proposed in [5]. Moreover, SWIPT has been studied under different channel setups, i.e., multiple fading channels [6], relay channels [7], [8], and multiple-input single-output (MISO) channels [9].

Most existing SWIPT works assume perfect channel state information (CSI) is available at the base station (BS). However, it is often difficult to obtain perfect CSI in practice because of channel estimation and quantization errors, which greatly degrades system performance. The authors in [10] proposed a robust secure beamformer for multiuser MISO SWIPT systems, where imperfect CSI of potential eavesdroppers is involved. In [11], a probabilistic robust SWIPT algorithm was designed, where rank-one beamforming solutions were derived with convex relaxations. Two robust joint beamforming and power splitting algorithms for MISO SWIPT systems were investigated in [12]. Moreover, the authors in [13], [14] presented various semidefinite relaxation (SDR) methods to solve the robust beamforming problem. Unfortunately, only suboptimal solutions can be derived while globally optimal solutions are currently unknown for multiuser SWIPT systems.

On the other hand, massive MIMO [15]–[17] has been considered as an additional attractive technology for SWIPT since it can significantly improve spectrum efficiency, energy efficiency and reliability [18], [19]. The authors in [20] and [21] investigated the wireless-energy-transfer problem in massive MIMO systems, where asymptotically optimal solutions and interesting insights into the optimal design were derived. SWIPT techniques for multi-way massive MIMO relay networks were developed in [22], where the fundamental tradeoff between harvested energy and sum rate was quantified. However, the works [20]–[22] again assume perfect CSI, which is even difficult to obtain in massive MIMO systems [23], [24]. In practice, CSI in massive MIMO systems can only be obtained from some low complexity channel estimation approaches [25]–[31]. For instance, [25], [26] applied low-rank approximations of the channel covariance matrices to reduce the number of estimated parameters. The authors in [27]–[29] applied an angle division multiple access (ADMA) model to represent massive MIMO channels with a few channel gains and angular parameters. A beamspace channel estimation scheme was designed in [30], [31], where the channel vectors are approximated by a few orthogonal basis

Manuscript received November 12, 2018; revised January 14, 2019; accepted April 29, 2019. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Guoqiang Mao. 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 {61831013, 61771274, 61531011}, the Beijing Municipal Natural Science Foundation under Grant {4182030, L182042}, and the R&D Project of the Science, Technology and Innovation Commission of Shenzhen Municipality No. JCYJ20180306170617062. The work of Y. C. Eldar was supported by the funding from the European Union's Horizon 2020 research and innovation program under grant agreement No. 646804-ERCCOG-BNYQ, and from the Israel Science Foundation under Grant No. 0100101.

F. Zhu is with Rocket Force University of Engineering, Xi'an, Shaanxi 710025, P. R. China (e-mail: fengchao_zhu@126.com).

F. Gao (corresponding author) is with Beijing National Research Center for Information Science and Technology (BNRist), Department of Automation, Tsinghua University, Beijing, Beijing 100084, P. R. China, and is also with Key Laboratory of Digital TV System of Guangdong Province and Shenzhen City, Research Institute of Tsinghua University in Shenzhen, Shenzhen 518057, P. R. China (e-mail: feifeigao@ieee.org).

Y. C. Eldar is with the Department of Mathematics and Computer Science, Weizmann Institute of Science, Rehovot, 32000, Israel (e-mail: yonina@ee.technion.ac.il).

G. Qian is with the Department of Communication Engineering, College of Information Engineering, Shenzhen University, Shenzhen 518000, P. R. China (e-mail: qiangb@szu.edu.cn).

from the discrete Fourier transform (DFT). However, all these works are based on approximately-sparse models such that the channel estimation errors are inevitable [34].

In this paper, we consider worst-case robust beamforming for SWIPT under beamspace massive MIMO scheme [30], [31], where the estimated channels are orthogonal to each other and have bounded estimation errors. Our objective is to minimize the transmit power of the BS, while providing the information user and the energy user with different signal-to-interference-and-noise ratio (SINR) and power, respectively, for all possible channel realizations. The resulting problem belongs to a well known non-convex [35] optimization formulation, for which only suboptimal solutions are available from existing works. In addition, the conventional robust designs [10]–[14] will suffer from high computational complexity with a large number of transmit antennas. By utilizing the orthogonality property within channel estimates [30], [31], we demonstrate that the problem can be globally solved when the channel estimation errors are smaller than a certain threshold. In particular, the optimal beamforming vectors are scaled versions of the estimated channels, and optimal scaling factors can be analytically obtained from a power allocation problem. Hence, the proposed approach greatly reduces the computational complexity compared to the conventional ones and is suitable for practical applications. Interestingly, the simulations further demonstrate that the proposed solution performs well even when the channel errors are large.

The rest of this paper is organized as follows: Section II describes the channel model of beamspace massive MIMO and formulates the proposed robust design. Section III derives the solutions of the non-convex problem and proves their optimality. Simulation results are provided in Section IV and conclusions are drawn in Section V.

Throughout the paper we use the following notations: Vectors are denoted by boldface small and matrices by capital letters. The Hermitian, inverse and Moore-Penrose inverse of $\boldsymbol{A}$ are written as $\boldsymbol{A}^{\mathrm{H}}$, $\boldsymbol{A}^{-1}$ and $\boldsymbol{A}^{\dagger}$ respectively. The inequalities $\boldsymbol{A} \succeq \mathbf{0}$ and $\boldsymbol{A} \succ \mathbf{0}$ mean that $\boldsymbol{A}$ is positive semi-definite and positive definite, respectively. We use $\mathrm{Tr}(\boldsymbol{A})$ to denote the trace, $\|\boldsymbol{x}\|$ is the Euclidean norm of a vector $\boldsymbol{x}$, $\mathbb{E}[\cdot]$ is the statistical expectation, and $\mathbb{R}^{a\times b}$ and $\mathbb{C}^{a\times b}$ are the spaces of $a\times b$ matrices with real- and complex-valued entries, respectively. Define $\mathrm{diag}(\cdot)$ as the operation of selecting diagonal elements of any $N\times N$ matrix. The distribution of a circularly symmetric complex Gaussian (CSCG) random variable with zero mean and variance σ^2 is written as $\mathcal{CN}(0,\sigma^2)$, and $\sim$ means "distributed as".

II. System Model and Problem Formulation

A. System Model

We consider SWIPT for a multiuser massive MIMO system shown in Fig. 1, where the BS is equipped with $N \gg 1$ antennas in the form of uniform linear array (ULA) with supercritical antenna spacing (i.e., less than or equal to half wavelength). There are $K+1$ single-antenna users randomly distributed in the coverage area, which contains K single-antenna information decoding users (or information users)

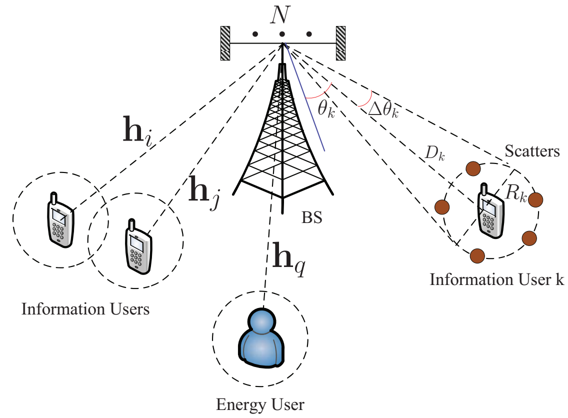


Fig. 1. Multiuser massive MIMO channel model with a ULA at the BS.

with index set $\mathcal{K} = \{1,\ldots,K\}$ and one energy-harvesting-user (or energy user). The kth user is located at D_k meters away from BS and is surrounded by a ring of $G_k \gg 1$ local scatterers with the radius R_k [25], [29]. The channel from the kth information user to BS is composed of G_k rays and can be expressed as [25], [29]:

$$\boldsymbol{h}_k = \frac{1}{\sqrt{G_k}}\sum_{g=1}^{G_k}\alpha_{k,g}\boldsymbol{a}(\theta_{k,g}), \quad 1\le k\le K, \tag{1}$$

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

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

where d is the antenna spacing, λ denotes the signal wavelength, and $\theta_{k,g}$ represents the direction of arrival (DOA) of the gth ray. We can similarly define the channel from the energy information user to BS as

$$\boldsymbol{h}_q = \frac{1}{\sqrt{G_q}}\sum_{g=1}^{G_q}\alpha_{q,g}\boldsymbol{a}(\theta_{q,g}). \tag{3}$$

In ideal massive MIMO case with $N \to \infty$, there is $\boldsymbol{h}_i^{\mathrm{H}}\boldsymbol{h}_j = 0$, $\boldsymbol{h}_i^{\mathrm{H}}\boldsymbol{h}_q = 0$, $\forall i \ne j$. However, in practice N cannot approach infinity and users with nearly orthogonal channels are allowed to transmit simultaneously with tolerable interference. Particularly, a popular low complexity beamspace channel scheme [30], [31] assigns non-overlapped columns of a discrete Fourier transform (DFT) matrix to different users and dictionary, such that the estimated channels for different users exactly satisfy

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

Such channel estimation is not exact, resulting in an error between $\boldsymbol{h}_i$ and $\tilde{\boldsymbol{h}}_i$ but can be implemented efficiently. Moreover, when N is large, the performance loss in channel estimation accuracy is small compared to optimal estimation.

We therefore assume that the real channel vectors $\boldsymbol{h}_k$ and $\boldsymbol{h}_q$ lie around the estimated channel vectors $\tilde{\boldsymbol{h}}_k$ and $\tilde{\boldsymbol{h}}_q$, respectively, so that

$$\begin{aligned} \boldsymbol{h}_k \in \mathcal{U}_k &= \left\{ \tilde{\boldsymbol{h}}_k + \boldsymbol{\delta}_k \mid \|\boldsymbol{\delta}_k\| \le \epsilon_k \right\}, \\ \boldsymbol{h}_q \in \mathcal{U}_q &= \left\{ \tilde{\boldsymbol{h}}_q + \boldsymbol{\delta}_q \mid \|\boldsymbol{\delta}_q\| \le \epsilon_q \right\}, \end{aligned} \tag{5}$$

where $\boldsymbol{\delta}_k \in \mathbb{C}^{N\times 1}$ and $\boldsymbol{\delta}_q \in \mathbb{C}^{N\times 1}$ are the channel estimation errors [32], [33] with norms bounded by ϵ_k and ϵ_q, respectively.

B. Problem Formulation

Our goal is to design simultaneous information beamforming vectors $\{\boldsymbol{s}_k \in \mathbb{C}^{N\times 1}\}$ and an energy beamforming vector[1] $\boldsymbol{q} \in \mathbb{C}^{N\times 1}$ for the information users and the energy user, respectively, which meet certain target requirements for all possible channel realizations. The baseband signal from BS can be expressed as

$$\boldsymbol{x}_b = \sum_{k=1}^{K} \boldsymbol{s}_k v_k + \boldsymbol{q} v_q, \tag{6}$$

where $v_k \sim \mathcal{CN}(0,1)$ denotes the data symbol for the kth information user, and $v_q \sim \mathcal{CN}(0,1)$ is the energy signal for the energy user. The downlink signal at the kth information user can then be expressed as

$$y_k = \boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{s}_k v_k + \sum_{i\neq k, i\in\mathcal{K}} \boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{s}_i v_i + \boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{q} v_q + n_k, \tag{7}$$

where $n_k \sim \mathcal{CN}(0, \sigma_k^2)$ represents the antenna noise of the kth information user. Similarly, the downlink signal at the energy user is given by

$$y_q = \sum_{k=1}^{K} \boldsymbol{h}_q^{\mathrm{H}} \boldsymbol{s}_k v_k + \boldsymbol{h}_q^{\mathrm{H}} \boldsymbol{q} v_q + n_q, \tag{8}$$

with $n_q \sim \mathcal{CN}(0, \sigma_q^2)$ representing the antenna noise of the energy user.

The harvested energy by the energy user is given by $\zeta\mathbb{E}\|y_q\|^2$, where $\zeta \in (0,1]$ denotes the energy conversion efficiency that depends on the rectification process and the energy harvesting circuit [20]–[22]. Using (8), we have

$$\mathbb{E}\|y_q\|^2 = \sum_{k=1}^{K} \|\boldsymbol{h}_q^{\mathrm{H}} \boldsymbol{s}_k\|^2 + \|\boldsymbol{h}_q^{\mathrm{H}} \boldsymbol{q}\|^2 + \sigma_q^2. \tag{9}$$

Our problem is to minimize the transmit power at the BS subject to the constraints that the harvested energy and SINR are above certain thresholds for all possible values of $\boldsymbol{h}_k$ and $\boldsymbol{h}_q$. This results in the optimization problem:

$$\begin{aligned} \mathbf{P1}: \quad & \min_{\{\boldsymbol{s}_k\},\boldsymbol{q}} \quad \|\boldsymbol{q}\|^2 + \sum_{k=1}^{K} \|\boldsymbol{s}_k\|^2 \\ & \text{s.t.} \quad \zeta\left(\sum_{k=1}^{K} \|\boldsymbol{h}_q^{\mathrm{H}} \boldsymbol{s}_k\|^2 + \|\boldsymbol{h}_q^{\mathrm{H}} \boldsymbol{q}\|^2 + \sigma_q^2 \right) \ge Q, \qquad \text{(10a)} \\ & \qquad \frac{\|\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{s}_k\|^2}{\sum_{i\neq k} \|\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{s}_i\|^2 + \|\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{q}\|^2 + \sigma_k^2} \ge \gamma_k, \qquad \text{(10b)} \\ & \qquad \forall \boldsymbol{h}_q \in \mathcal{U}_q, \quad \forall \boldsymbol{h}_k \in \mathcal{U}_k, \quad k = 1, 2, \ldots, K, \end{aligned}$$

where $Q > 0$ is the desired harvested energy for the energy user, and $\gamma_k > 0$ is the target SINR for the kth information user. It is observed from (10a) that when $\sigma_q^2 - Q/\zeta \ge 0$ (the case that Q is small enough), the energy user can be satisfied by the antenna noise. Nevertheless, such case is trivial, and we consider $\sigma_q^2 - Q/\zeta < 0$ in the rest of the paper.

Remark 1: When the channel estimation errors are equal to zero, i.e., $\mathcal{U}_k = \left\{\tilde{\boldsymbol{h}}_k\right\}$ and $\mathcal{U}_q = \left\{\tilde{\boldsymbol{h}}_q\right\}$, $\mathbf{P1}$ can be simplified as a non-robust optimization for massive MIMO, where zero-forcing (ZF) beamformer [20]–[22] has been proven to be the optimal transmit strategy. In this case [20]–[22], the optimal information and energy beamforming directions for $\mathbf{P1}$ can be chosen as $\boldsymbol{h}_k/\|\boldsymbol{h}_k\|$ and $\boldsymbol{h}_q/\|\boldsymbol{h}_q\|$, respectively. Then, $\mathbf{P1}$ can be further simplified as a power allocation problem, where the zero-forcing (ZF) beamforming solutions are given as

$$\begin{aligned} \boldsymbol{q} &= \left(\sqrt{Q/\zeta - \sigma_q^2}\right) \tilde{\boldsymbol{h}}_q/\|\tilde{\boldsymbol{h}}_q\|^2, \\ \boldsymbol{s}_k &= \left(\sqrt{\gamma_k \sigma_k^2}\right) \tilde{\boldsymbol{h}}_k/\|\tilde{\boldsymbol{h}}_k\|^2. \end{aligned} \tag{11}$$

However, when the channel estimation errors are not zero, i.e., $\epsilon_k > 0$ and $\epsilon_q > 0$, the optimal solutions of $\mathbf{P1}$ are in general hard to obtain [10]–[14]. Nevertheless, we next show that $\mathbf{P1}$ can be globally solved when $\{\epsilon_k\}$ and ϵ_q are sufficiently small, and the solutions reduce to (11) when there are no channel errors.

III. Optimal Robust Beamforming

A. Semidefinite Relaxation (SDR)

We will solve $\mathbf{P1}$ by first obtaining an equivalent problem which has a natural SDR representation. Then, we will show that the SDR has a closed-form solution which is optimal for the original problem under small channel errors.

The main difficulty in solving $\mathbf{P1}$ lies in the constraints (10a) and (10b). Substituting (5) into (10a), we equivalently rewrite (10a) as

$$\begin{aligned} &\left(\tilde{\boldsymbol{h}}_q + \boldsymbol{\delta}_q\right)^{\mathrm{H}} \left(\boldsymbol{q}\boldsymbol{q}^{\mathrm{H}} + \sum_{k=1}^{K} \boldsymbol{s}_k \boldsymbol{s}_k^{\mathrm{H}} \right) \left(\tilde{\boldsymbol{h}}_q + \boldsymbol{\delta}_q\right) + \sigma_q^2 - Q/\zeta \ge 0, \\ &\forall \boldsymbol{\delta}_q, \quad -\boldsymbol{\delta}_q^{\mathrm{H}} \boldsymbol{\delta}_q + \epsilon_q^2 \ge 0. \end{aligned} \tag{12}$$

Similarly, (10b) is equivalent to

$$\begin{aligned} &\left(\tilde{\boldsymbol{h}}_k + \boldsymbol{\delta}_k\right)^{\mathrm{H}} \left(\frac{1}{\gamma_k} \boldsymbol{s}_k \boldsymbol{s}_k^{\mathrm{H}} - \sum_{i\neq k} \boldsymbol{s}_i \boldsymbol{s}_i^{\mathrm{H}} - \boldsymbol{q}\boldsymbol{q}^{\mathrm{H}} \right) \left(\tilde{\boldsymbol{h}}_k + \boldsymbol{\delta}_k\right) - \sigma_k^2 \ge 0, \\ &\forall \boldsymbol{\delta}_k, \quad -\boldsymbol{\delta}_k^{\mathrm{H}} \boldsymbol{\delta}_k + \epsilon_k^2 \ge 0, \quad k = 1, 2, \ldots, K. \end{aligned} \tag{13}$$

[1] Note that the energy user could also absorb the energy from the information signals. However, with high directional beamforming in massive MIMO [36], [37], the energy leaked from information users may not be sufficient for the energy user. We then propose to use a specific energy signal for energy users. This is why we require the energy users to have relatively orthogonal channels from information users to avoid interference, which is a key difference from the conventional SWIPT design.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2019.2916405, IEEE Transactions on Wireless Communications

We next use the following lemma to reformulate the constraints (12) and (13).

Lemma 1 (S-Procedure [38]): Let $f_1(\boldsymbol{x}) = \boldsymbol{x}^{\rm H}\boldsymbol{A}_1\boldsymbol{x}+\boldsymbol{b}_1^{\rm H}\boldsymbol{x}+\boldsymbol{x}^{\rm H}\boldsymbol{b}_1+c_1$ and $f_2(\boldsymbol{x}) = \boldsymbol{x}^{\rm H}\boldsymbol{A}_2\boldsymbol{x}+\boldsymbol{b}_2^{\rm H}\boldsymbol{x}+\boldsymbol{x}^{\rm H}\boldsymbol{b}_2+c_2$, for some $\boldsymbol{A}_1,\boldsymbol{A}_2 \in \mathbb{C}^{n\times n}$, $\boldsymbol{b}_1,\boldsymbol{b}_2 \in \mathbb{C}^{n\times 1}$, $c_1, c_1 \in \mathbb{R}$. The condition $f_1(\boldsymbol{x}) \geq 0 \Rightarrow f_2(\boldsymbol{x}) \geq 0$ holds true if and only if there exists a nonnegative μ, such that

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

From Lemma 1, we know (12) holds true if and only if there exists $\mu_q \geq 0$ such that

$$\begin{bmatrix} \boldsymbol{X}_q + \mu_q \boldsymbol{I} & \boldsymbol{X}_q\tilde{\boldsymbol{h}}_q \\ \tilde{\boldsymbol{h}}_q^{\rm H}\boldsymbol{X}_q^{\rm H} & \tilde{\boldsymbol{h}}_q^{\rm H}\boldsymbol{X}_q\tilde{\boldsymbol{h}}_q + \sigma_q^2 - Q/\zeta - \mu_q\epsilon_q^2 \end{bmatrix} \succeq \boldsymbol{0}, \quad (14)$$

where for simplicity, we define

$$\boldsymbol{X}_q \triangleq \boldsymbol{q}\boldsymbol{q}^{\rm H} + \sum_{k=1}^{K} \boldsymbol{s}_k\boldsymbol{s}_k^{\rm H}. \quad (15)$$

Similarly, (13) holds true if and only if there exist $\mu_{s_k} \geq 0,\ k = 1, 2, \ldots, K$ such that

$$\begin{bmatrix} \boldsymbol{X}_{s_k} + \mu_{s_k} \boldsymbol{I} & \boldsymbol{X}_{s_k}\tilde{\boldsymbol{h}}_k \\ \tilde{\boldsymbol{h}}_k^{\rm H}\boldsymbol{X}_{s_k}^{\rm H} & \tilde{\boldsymbol{h}}_k^{\rm H}\boldsymbol{X}_{s_k}\tilde{\boldsymbol{h}}_k - \sigma_k^2 - \mu_{s_k}\epsilon_k^2 \end{bmatrix} \succeq \boldsymbol{0}, \quad (16)$$

with

$$\boldsymbol{X}_{s_k} \triangleq \frac{1}{\gamma_k}\boldsymbol{s}_k\boldsymbol{s}_k^{\rm H} - \sum_{i\neq k} \boldsymbol{s}_i\boldsymbol{s}_i^{\rm H} - \boldsymbol{q}\boldsymbol{q}^{\rm H}. \quad (17)$$

Using (14)–(17), $\mathbf{P1}$ can be equivalently expressed as

$$\mathbf{P1}-\mathbf{EQV}: \min_{\mu_q,\{\mu_{s_k}\},\boldsymbol{q},\{\boldsymbol{s}_k\}} \ \mathrm{Tr}(\boldsymbol{X}_q) \quad (18a)$$

$$\text{s.t.} \begin{bmatrix} \boldsymbol{X}_q+\mu_q \boldsymbol{I} & \boldsymbol{X}_q\tilde{\boldsymbol{h}}_q \\ \tilde{\boldsymbol{h}}_q^{\rm H}\boldsymbol{X}_q^{\rm H} & \tilde{\boldsymbol{h}}_q^{\rm H}\boldsymbol{X}_q\tilde{\boldsymbol{h}}_q+\sigma_q^2-Q/\zeta-\mu_q\epsilon_q^2 \end{bmatrix} \succeq \boldsymbol{0}, \quad (18b)$$

$$\begin{bmatrix} \boldsymbol{X}_{s_k}+\mu_{s_k} \boldsymbol{I} & \boldsymbol{X}_{s_k}\tilde{\boldsymbol{h}}_k \\ \tilde{\boldsymbol{h}}_k^{\rm H}\boldsymbol{X}_{s_k}^{\rm H} & \tilde{\boldsymbol{h}}_k^{\rm H}\boldsymbol{X}_{s_k}\tilde{\boldsymbol{h}}_k-\sigma_k^2-\mu_{s_k}\epsilon_k^2 \end{bmatrix} \succeq \boldsymbol{0}, \quad (18c)$$

$$\boldsymbol{X}_q = \sum_{k}^{K} \boldsymbol{S}_k + \boldsymbol{Q}, \quad (18d)$$

$$\boldsymbol{X}_{s_k} = \frac{1}{\gamma_k}\boldsymbol{S}_k - \sum_{i\neq k}\boldsymbol{S}_i - \boldsymbol{Q}, \quad (18e)$$

$$\mu_q \geq 0, \quad \mu_{s_k} \geq 0, \quad (18f)$$

$$\boldsymbol{Q} = \boldsymbol{q}\boldsymbol{q}^{\rm H}, \quad \boldsymbol{S}_k = \boldsymbol{s}_k\boldsymbol{s}_k^{\rm H}, \quad k = 1, 2, \ldots, K, \quad (18g)$$

where μ_q and $\{\mu_{s_k}\}$ are the auxiliary variables generated by the S-Procedure. The nonlinear constraints in (18g) are equivalent to:

$$\boldsymbol{Q} \succeq \boldsymbol{0}, \quad \boldsymbol{S}_k \succeq \boldsymbol{0}, \quad \mathrm{Rank}(\boldsymbol{Q}) = 1, \quad \mathrm{Rank}(\boldsymbol{S}_k) = 1. \quad (19)$$

Clearly, $\mathbf{P1}-\mathbf{EQV}$ is non-convex with rank constraints $\mathrm{Rank}(\boldsymbol{Q}) = 1$ and $\mathrm{Rank}(\boldsymbol{S}_k) = 1$. Dropping the rank constraints, we obtain the following relaxed convex optimization:

$$\mathbf{P1}-\mathbf{SDR}: \min_{\mu_q,\{\mu_{s_k}\},\boldsymbol{Q},\{\boldsymbol{S}_k\}} \ \mathrm{Tr}(\boldsymbol{X}_q)$$

$$\text{s.t.} \ (18b) - (18f), \ \boldsymbol{Q} \succeq \boldsymbol{0}, \ \boldsymbol{S}_k \succeq \boldsymbol{0}, \ k = 1, 2, \ldots, K. \quad (20)$$

Note that $\mathbf{P1}-\mathbf{SDR}$ will be directly solved as an SDP problem for the conventional randomized SDR method [41]. If the optimal solutions of $\mathbf{P1}-\mathbf{SDR}$ are rank-1, then the randomized SDR technique is exactly the same as the SDP method [39], [40]. Nevertheless, directly solving $\mathbf{P1}-\mathbf{SDR}$ as an SDP problem is extremely complex due to the large number of antennas. In addition, $\mathbf{P1}-\mathbf{SDR}$ does not guarantee rank-one solutions and hence the derived solutions may not be optimal for the initial problem $\mathbf{P1}$ [13], [14]. Nevertheless, we next show that $\mathbf{P1}-\mathbf{SDR}$ guarantees rank-one solutions in the case of beamspace massive MIMO systems under some reasonable conditions.

B. Optimal Rank-one Solutions

In this subsection, we further investigate $\mathbf{P1}-\mathbf{SDR}$ to provide more insight into the form of optimal solutions.

Bearing in mind that under the beamspace massive MIMO setting (4), the estimated channels are orthogonal to each other, we define $\{\boldsymbol{u}_i = \tilde{\boldsymbol{h}}_i/\|\tilde{\boldsymbol{h}}_i\|,\ 1 \leq i \leq K\}$ and $\boldsymbol{u}_{K+1} = \tilde{\boldsymbol{h}}_q/\|\tilde{\boldsymbol{h}}_q\|$.

Proposition 1: Assume $\boldsymbol{Q}^\star$ and $\{\boldsymbol{S}_k^\star\}$ are the optimal solutions for $\mathbf{P1}-\mathbf{SDR}$. Then $\boldsymbol{Q}^\star$ and $\{\boldsymbol{S}_k^\star\}$ must have the forms:

$$\boldsymbol{Q}^\star = \sum_{i=1}^{K+1} P_{q,i}^\star \boldsymbol{u}_i\boldsymbol{u}_i^{\rm H}, \quad \boldsymbol{S}_k^\star = \sum_{i=1}^{K+1} P_{s_k,i}^\star \boldsymbol{u}_i\boldsymbol{u}_i^{\rm H}, \quad (21)$$

where $\{P_{q,i}^\star \geq 0,\ 1 \leq i \leq K+1\}$ and $\{P_{s_k,i}^\star \geq 0,\ 1 \leq i \leq K+1\}$.

Proof: See Appendix A. ■

Proposition 1 states that the optimal energy transmit covariance $\boldsymbol{Q}^\star$ and the information transmit covariances $\{\boldsymbol{S}_k^\star\}$ must lie in the space spanned by the estimated channels.

Proposition 2: At the optimal point, we have $\mu_q > 0$ in $\mathbf{P1}-\mathbf{SDR}$. Moreover, (18b) can be equivalently rewritten as

$$\frac{\mu_q\|\tilde{\boldsymbol{h}}_q\|^2\left(P_{q,K+1} + \sum_{k=1}^{K} P_{s_k,K+1}\right)}{\mu_q + P_{q,K+1} + \sum_{k=1}^{K} P_{s_k,K+1}} + \sigma_q^2 - Q/\zeta - \mu_q\epsilon_q^2 \geq 0.$$

Proof: See Appendix B. ■

Proposition 3: At the optimal point, we have $\mu_{s_k} > 0$ in $\mathbf{P1}-\mathbf{SDR}$. Moreover, (18c) can be equivalently rewritten as

$$\frac{\mu_{s_k}\|\tilde{\boldsymbol{h}}_k\|^2\left(\frac{P_{s_k,k}}{\gamma_k} - \sum_{i\neq k} P_{s_i,k} - P_{q,k}\right)}{\mu_{s_k} + \frac{P_{s_k,k}}{\gamma_k} - \sum_{i\neq k} P_{s_i,k} - P_{q,k}} - \sigma_k^2 - \mu_{s_k}\epsilon_k^2 \geq 0,$$

$$\frac{P_{s_k,i}}{\gamma_k} - \sum_{j\neq k} P_{s_j,i} - P_{q,i} + \mu_{s_k} \geq 0, \ 1 \leq i \leq K+1.$$

Proof: See Appendix C. ■

Combining the above propositions leads to the following theorem.

Theorem 1: Denote $\mu_q^\star, \{\mu_{s_k}^\star\}$ as the optimal solutions of $\mathbf{P1}-\mathbf{SDR}$. With (21), if

$$\min\{\mu_{s_k}^\star\} \geq \sum_{i=1}^{K} P_{s_i,K+1}^\star + P_{q,K+1}^\star, \tag{22}$$

then:

(a) $\boldsymbol{Q}^\star = \boldsymbol{q}^\star \boldsymbol{q}^{\star\mathrm{H}} = P_q^\star \tilde{\boldsymbol{h}}_q \tilde{\boldsymbol{h}}_q^{\mathrm{H}} / \|\tilde{\boldsymbol{h}}_q\|^2$ is the optimal energy transmit covariance, where $P_q^\star = \sum_{i=1}^{K} P_{s_i,K+1}^\star + P_{q,K+1}^\star$ is the power allocated for energy beamforming;

(b) $\boldsymbol{S}_k^\star = \boldsymbol{s}_k^\star \boldsymbol{s}_k^{\star\mathrm{H}} = P_k^\star \tilde{\boldsymbol{h}}_k \tilde{\boldsymbol{h}}_k^{\mathrm{H}} / \|\tilde{\boldsymbol{h}}_k\|^2$ is the optimal information transmit covariance, where $P_k^\star = P_{s_k,k}^\star - \sum_{i \neq k}^{K} P_{s_i,k}^\star - P_{q,k}^\star$ is the power allocated to the kth information beamforming.

Proof: See Appendix D. ■

Theorem 1 says that when (22) holds, the optimal solutions of $\mathbf{P1}-\mathbf{SDR}$ and $\mathbf{P1}$ are exactly the same, and the optimal beamforming directions are equal to $\tilde{\boldsymbol{h}}_q \tilde{\boldsymbol{h}}_q^{\mathrm{H}} / \|\tilde{\boldsymbol{h}}_q\|^2$ and $\tilde{\boldsymbol{h}}_k \tilde{\boldsymbol{h}}_k^{\mathrm{H}} / \|\tilde{\boldsymbol{h}}_k\|^2$, respectively. The remaining optimal variables $\mu_q^\star$, $\mu_{s_k}^\star$, $P_q^\star$ and $P_k^\star$ are still to be obtained, which will be done in the next subsection.

C. Optimal Power Allocation

When (22) holds, the optimal energy beamforming and information beamforming vectors are $\boldsymbol{q}^\star = \sqrt{P_q^\star} \tilde{\boldsymbol{h}}_q / \|\tilde{\boldsymbol{h}}_q\|$ and $\boldsymbol{s}_k^\star = \sqrt{P_k^\star} \tilde{\boldsymbol{h}}_k / \|\tilde{\boldsymbol{h}}_k\|$, respectively. Using Proposition 2, Proposition 3 and Theorem 1, we know that $\mathbf{P1}-\mathbf{SDR}$ can be simplified to the following power allocation problem:

$$\mathbf{P2}: \min_{\mu_q, \{\mu_{s_k}\}, P_q, \{P_k\}} \quad P_q + \sum_{k=1}^{K} P_k \tag{23a}$$

$$\text{s.t. } \frac{\mu_q P_q \|\tilde{\boldsymbol{h}}_q\|^2}{\mu_q + P_q} + \sigma_q^2 - Q/\zeta - \mu_q \epsilon_q^2 \geq 0, \tag{23b}$$

$$\frac{\mu_{s_k} P_k \|\tilde{\boldsymbol{h}}_k\|^2}{\mu_{s_k} \gamma_k + P_k} - \sigma_k^2 - \mu_{s_k} \epsilon_k^2 \geq 0, \tag{23c}$$

$$\mu_q > 0, \mu_{s_k} > 0, P_q \geq 0, P_k \geq 0, k = 1, 2, \ldots, K. \tag{23d}$$

Note that when $P_q = 0$, the constraint (23b) is always unable to satisfy since $\sigma_q^2 - Q/\zeta < 0$ and $\mu_q \epsilon_q^2 > 0$. Thus, there must be $P_q > 0$. Similarly, we know $P_k > 0$ in $\mathbf{P2}$. Moreover, we know from Theorem 1 that at the optimal point, $\min\{\mu_{s_k}^\star\} \geq \sum_{i=1}^{K} P_{s_i,K+1}^\star + P_{q,K+1}^\star$ can be further expressed as

$$\min\{\mu_{s_k}^\star\} \geq P_q^\star. \tag{24}$$

Consequently, the optimal solutions of $\mathbf{P2}$ are exactly the same as $\mathbf{P1}$ if $\min\{\mu_{s_k}^\star\} \geq P_q^\star$.

Define

$$f_q(\mu_q, P_q) = \frac{\mu_q P_q \|\tilde{\boldsymbol{h}}_q\|^2}{\mu_q + P_q}, f_{s_k}(\mu_{s_k}, P_k) = \frac{\mu_{s_k} P_k \|\tilde{\boldsymbol{h}}_k\|^2}{\mu_{s_k} \gamma_k + P_k}.$$

The Hessian matrices of $f_q(\mu_q, P_q)$ and $f_{s_k}(\mu_{s_k}, P_k)$ are given by

$$\nabla^2 f_q(\mu_q, P_q) = \frac{-2\|\tilde{\boldsymbol{h}}_q\|^2}{(\mu_q + P_q)^3} \begin{bmatrix} P_q^2 & -\mu_q P_q \\ -\mu_q P_q & \mu_q^2 \end{bmatrix},$$

$$\nabla^2 f_{s_k}(\mu_{s_k}, P_k) = \frac{-2\gamma_k \|\tilde{\boldsymbol{h}}_k\|^2}{(\mu_{s_k} \gamma_k + P_k)^3} \begin{bmatrix} P_k^2 & -\mu_{s_k} P_k \\ -\mu_{s_k} P_k & \mu_{s_k}^2 \end{bmatrix},$$

respectively. Due to $\mu_q > 0$, $\mu_{s_k} > 0$, $P_q > 0$, $P_k > 0$, $\|\tilde{\boldsymbol{h}}_k\|^2 > 0$ and $\|\tilde{\boldsymbol{h}}_q\|^2 > 0$, it can be easily derived from the Hessian matrices that $\nabla^2 f_q(\mu_q, P_q) \prec \mathbf{0}$ and $\nabla^2 f_{s_k}(\mu_{s_k}, P_k) \prec \mathbf{0}$, which implies that $f_q(\mu_q, P_q)$ and $f_{s,k}(\mu_{s_k}, P_k)$ are concave functions. Thus, $\mathbf{P2}$ is a convex optimization formulation.

Using the Karush-Kuhn-Tucker (KKT) conditions for $\mathbf{P2}$, we prove the following results.

Proposition 4: When

$$\min\left\{\frac{\sigma_k^2}{(\|\tilde{\boldsymbol{h}}_k\| - \epsilon_k)\epsilon_k}\right\} \geq \frac{Q/\zeta - \sigma_q^2}{(\|\tilde{\boldsymbol{h}}_q\| - \epsilon_q)^2}, \tag{25}$$

the optimal solutions of $\mathbf{P2}$ are also optimal for $\mathbf{P1}$, and are given by

$$\mu_q^\star = \frac{P_q^\star \|\tilde{\boldsymbol{h}}_q\| - P_q^\star \epsilon_q}{\epsilon_q}, \ \mu_{s_k}^\star = \frac{P_k^\star \|\tilde{\boldsymbol{h}}_k\| - P_k^\star \epsilon_k}{\gamma_k \epsilon_k}, \tag{26}$$

$$P_q^\star = \frac{Q/\zeta - \sigma_q^2}{(\|\tilde{\boldsymbol{h}}_q\| - \epsilon_q)^2}, \ P_k^\star = \frac{\gamma_k \sigma_k^2}{(\|\tilde{\boldsymbol{h}}_k\| - \epsilon_k)^2}, \tag{27}$$

where $k \in \{1, \ldots, K\}$.

Proof: See Appendix E. ■

The variables in (25) are all known at the BS. Thus it is easy to determine whether the solutions in (26) and (27) are optimal for $\mathbf{P1}$. Note that when (25) is not satisfied, the optimal solutions of $\mathbf{P2}$ are in general suboptimal for $\mathbf{P1}$.

We summarize the main results of this section in the following theorem.

Theorem 2: If (25) is satisfied, then the optimal beamforming vectors for $\mathbf{P1}$ are given by

$$\boldsymbol{q}^\star = \sqrt{\frac{Q/\zeta - \sigma_q^2}{(\|\tilde{\boldsymbol{h}}_q\| - \epsilon_q)^2}} \tilde{\boldsymbol{h}}_q / \|\tilde{\boldsymbol{h}}_q\|,$$

$$\boldsymbol{s}_k^\star = \sqrt{\frac{\gamma_k \sigma_k^2}{(\|\tilde{\boldsymbol{h}}_k\| - \epsilon_k)^2}} \tilde{\boldsymbol{h}}_k / \|\tilde{\boldsymbol{h}}_k\|. \tag{28}$$

When (25) is not satisfied, it is shown in simulations that (28) can serve as an efficient alternative to the conventional SDR method, even when the estimation errors are large. The number of variables to be solved for $\mathbf{P1}-\mathbf{SDR}$ and $\mathbf{P2}$ are compared in Table I. For $\mathbf{P1}-\mathbf{SDR}$, there are totally $(N+1)(K+1)$ unknown variables. While for $\mathbf{P2}$, there are totally $2(K+1)$ unknown variables. Consequently, $\mathbf{P2}$ possesses a much lower computational complexity than $\mathbf{P1}-\mathbf{SDR}$. Moreover, the optimal solutions of $\mathbf{P2}$ are in closed-form and can be easily calculated.

IV. Simulation Results

In this section, we present simulations to evaluate the performance of the proposed closed-form robust beamforming.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2019.2916405, IEEE Transactions on Wireless Communications

TABLE I
NUMBER OF VARIABLES COMPARISON

Variables \ Problems	$\mathbf{P1-SDR}$	$\mathbf{P2}$
Number of Real	$K+1$	$2\times(K+1)$
Number of Complex	$N\times(K+1)$	0
Total Variables	$(N+1)\times(K+1)$	$2\times(K+1)$

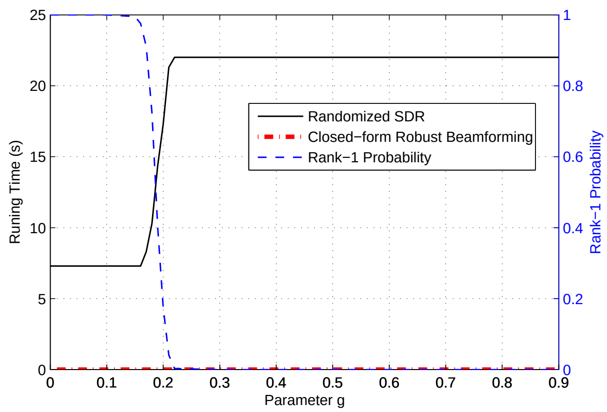


Fig. 2. Rank-1 probability versus the parameter g with $\gamma = 10$ dB, $Q/\zeta = 10$ dBm, $K = 30$ and $N = 128$.

For simplicity, the target SINR of all users are assumed to be the same, i.e., $\gamma = \gamma_1, \ldots, = \gamma_K$. The channel estimation errors are generated as independent CSCG random variables distributed with $\mathcal{CN}(0, g^2)$, where we define $g = \epsilon_k/\|\tilde{\boldsymbol{h}}_k\| = \epsilon_q/\|\tilde{\boldsymbol{h}}_q\|$ with $g \in [0, 1)$. The simulation results are averaged over 10000 Monte Carlo runs.

In the first example, we examine the average CPU running times and the rank-1 probability versus the parameter g in Fig. 2 and the harvest power Q/ζ in Fig. 3, respectively. The simulation results of the conventional randomized SDR method [41] are also displayed for comparison. It is observed from Fig. 2 and Fig. 3 that the rank-1 probability of $\mathbf{P1-SDR}$ is a decreasing function with respect to g and Q/ζ. The optimal solutions of $\mathbf{P1-SDR}$ are always rank-1 when $g \leq 0.15$ and $Q/\zeta \leq 10$ dBm, which means that we can always derive the optimal solutions using the closed-form robust beamforming method when g and Q/ζ are small. In addition, we see that the average CPU running time of the closed-form robust beamforming method is a small constant (say about 0.004s) which does not change with g and Q/ζ. This is mainly due to the fact the optimal solutions have closed-forms, where the change of g or Q/ζ will not affect the computational complexity. On the other hand, the average CPU running time of the conventional randomized SDR is an increasing function with respect to g and Q/ζ, which is much greater than that of the proposed method. In fact, the rank-1 probability of $\mathbf{P1-SDR}$ approaches zero when g and Q/ζ are large enough, where the conventional randomized SDR method needs more time to find the suboptimal rank-1 solutions [41].

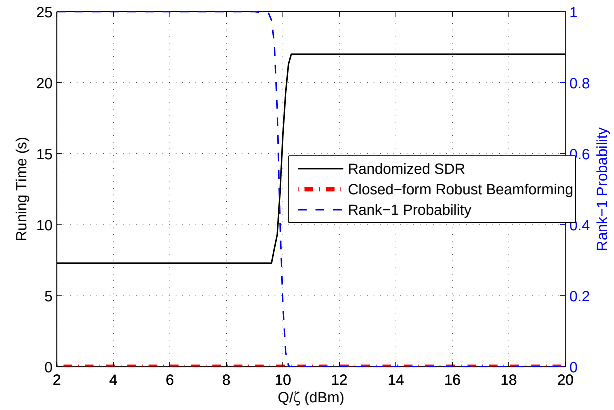


Fig. 3. Rank-1 probability versus harvested power Q/ζ with $\gamma = 10$ dB, $g = 0.1$, $K = 30$ and $N = 128$.

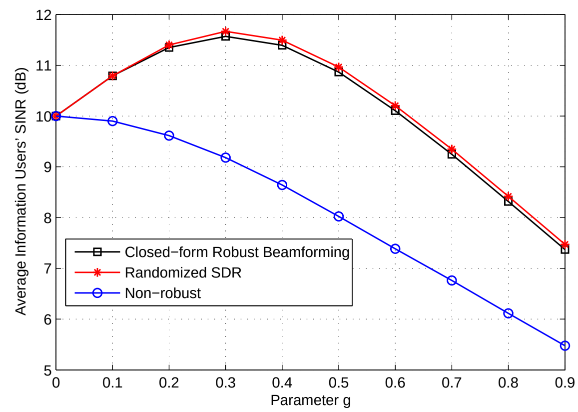


Fig. 4. Average information users' SINR versus parameter g with $K = 30$, $N = 128$, $Q/\zeta = 10$ dBm and $\gamma = 10$ dB.

In the second example, we plot the average information users' SINR versus g and transmit antennas N for the closed-form robust beamforming method, the randomized SDR method, and the non-robust method[2] in Fig. 4 and Fig. 5. The optimal beamforming solutions of $\mathbf{P1}$ for the non-robust method are given by (11): $\boldsymbol{q} = \left(\sqrt{Q/\zeta - \sigma_q^2}\right)\tilde{\boldsymbol{h}}_q/\|\tilde{\boldsymbol{h}}_q\|^2$ and $\boldsymbol{s}_k = \left(\sqrt{\gamma_k\sigma_k^2}\right)\tilde{\boldsymbol{h}}_k/\|\tilde{\boldsymbol{h}}_k\|^2$. It is clearly seen from Fig. 4 that the average SINR of information users for the non-robust method is a decreasing function with respect to g. The reason is that the non-robust solution does not take the channel estimation errors into consideration. Interestingly, the average SINR of information users for the closed-form robust beamforming algorithm increases with g when $g \leq 0.3$. This is mainly due to the fact that when g increases, the BS will allocate more power for information beamforming (see (E.11) for details). While the average information users' SINR for the closed-form robust beamforming algorithm will decrease with the increase of g when $g > 0.3$. This phenomenon can be

[2]Note that for the non-robust method, the estimated channels will be directly assumed as perfect and are then used for beamforming. Thus, the optimal beamforming solutions of $\mathbf{P1}$ for the non-robust method can be easily derived the same as (11).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2019.2916405, IEEE Transactions on Wireless Communications

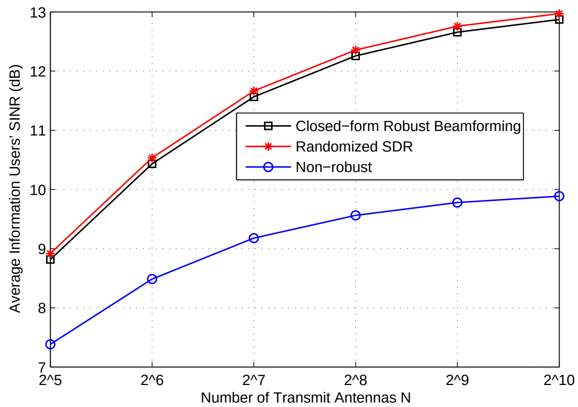


Fig. 5. Average information users' SINR versus number of transmit antennas N with $K = 30$, $g = 0.3$, $Q/\zeta = 10$ dBm and $\gamma = 10$ dB.

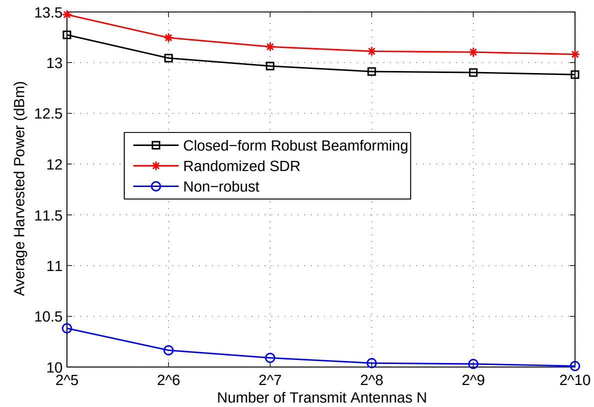


Fig. 7. Average harvested power versus number of transmit antennas N with $K = 30$, $g = 0.3$, $Q/\zeta = 10$ dBm and $\gamma = 10$ dB.

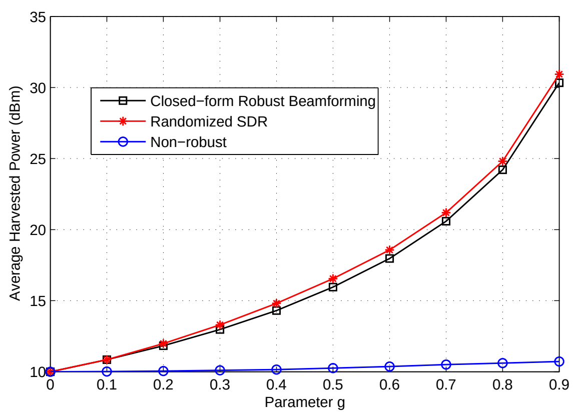


Fig. 6. Average harvested power versus parameter g with $K = 30$, $N = 128$, $g = 0.3$, $Q/\zeta = 10$ dBm and $\gamma = 10$ dB.

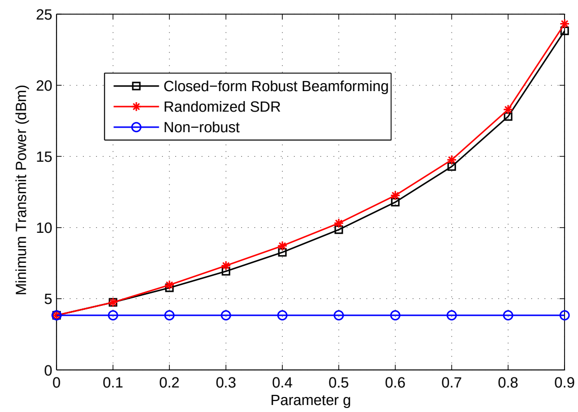


Fig. 8. Minimum transmit power versus parameter g with $K = 30$, $N = 128$, $g = 0.3$, $Q/\zeta = 10$ dBm and $\gamma = 10$ dB.

explained by Fig. 2 which shows that the rank-1 probability will always be zero when $g > 0.3$. In this case, all the optimal solutions of $\mathbf{P1-SDR}$ are not rank-1. Thus, the performance of the closed-form robust beamforming algorithm deteriorates with increasing g. Moreover, it is clear that the performance of the closed-form robust beamforming algorithm is exactly the same as that of the randomized SDR method when g is small (i.e., the rank-1 probability is equal to 1). The gap between the closed-form robust beamforming algorithm and randomized SDR is small when g is large, which means that we can obtain good performance using our approach. Similarly, it is seen from Fig. 5 that the average harvested power for the closed-form robust beamforming algorithm, randomized SDR and the non-robust method increase with N, which implies that we could use more antennas to improve the performance of beamspace massive MIMO systems.

In the third example, we plot the average harvested power versus g and N in Fig. 6 and Fig. 7, respectively. We see from Fig. 6 that the average harvested power is an increasing function with respect to g. This is mainly due to the fact that when g increases, more power is received at the energy user. However, it is clear in Fig. 7 that the average harvested power is a decreasing function of N. Nevertheless, it is seen from Fig. 6 and Fig. 7 that the average harvested power derived by the three methods is always bigger than 10 dBm, which says that the constraint in (10a) will be strictly satisfied.

In the last example, we plot the average minimum transmit power versus g and N in Fig. 8 and Fig. 9, respectively. It is seen from Fig. 8 that the transmit power for the non-robust method will not change with g. While the transmit power for the closed-form robust beamforming algorithm and the randomized SDR approach increases when g becomes large. This is because when the channel estimation errors increase, the BS will allocate more power to eliminate the CSI uncertainty. As a result, the proposed algorithm and the randomized SDR approach can obtain a much higher average SINR of information users and average harvested power than the non-robust method. Moreover, it is clear from Fig. 9 that the transmit power for the three techniques decreases with increasing N, which means that we may save more power with large number of antennas at the BS.

V. Conclusions

In this paper, we design simultaneous robust information and energy beamforming for SWIPT in a multiuser beamspace massive MIMO system. Our target is to minimize the transmit

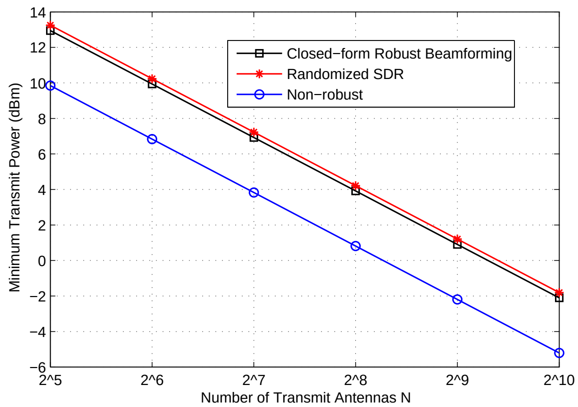


Fig. 9. Minimum transmit power versus number of transmit antennas N with $K = 30$, $g = 0.3$, $Q/\zeta = 10$ dBm and $\gamma = 10$ dB.

power of the BS while providing the information users and the energy user with desired SINRs and harvested power, respectively. Instead of solving the optimization with SDP techniques, we solve a relaxed power allocation problem, where beamforming directions and power allocations are derived in closed-form. More importantly, we strictly prove that the relaxed power allocation problem is equivalent to the initial robust design under beamspace massive MIMO schemes when the channel estimation errors are small enough. Simulation results are provided to corroborate the proposed studies, and show that the developed approach still achieves good performance even when the channel estimation errors are large.

Appendix A
Proof of Proposition 1

Assume $\boldsymbol{Q}^\star$ and $\boldsymbol{S}_k^\star, 1 \le k \le K$ are the optimal solutions of $\mathbf{P1}-\mathbf{SDR}$. Define the $N \times N$ unitary matrix $\boldsymbol{U}$ as $\boldsymbol{U} = [\boldsymbol{u}_1, \ldots, \boldsymbol{u}_N]$ where $\boldsymbol{u}_i, K+2 \le i \le N$ are an arbitrary set of orthonormal vectors that are orthogonal to $\boldsymbol{u}_i, 1 \le i \le K+1$.

A. Part (a): Proving that $\boldsymbol{Q}^\star$ and $\boldsymbol{S}_k^\star$ can be simultaneously diagonalized by $\boldsymbol{U}$ and $\boldsymbol{U}^{\mathrm{H}}$

Since $\boldsymbol{U}$ is unitary, $\boldsymbol{Q}^\star \succeq \mathbf{0}$, and $\boldsymbol{S}_k^\star \succeq \mathbf{0}$, we can always write

$$\boldsymbol{Q}^\star = \boldsymbol{U}\boldsymbol{D}_q\boldsymbol{U}^{\mathrm{H}}, \quad \boldsymbol{S}_k^\star = \boldsymbol{U}\boldsymbol{D}_{s_k}\boldsymbol{U}^{\mathrm{H}}, \tag{A.1}$$

for some $\boldsymbol{D}_q \succeq \mathbf{0}$ and $\boldsymbol{D}_{s_k} \succeq \mathbf{0}$, which are the corresponding matrices. Substituting (A.1) into (18d) and (18e), we have

$$\boldsymbol{X}_q^\star = \boldsymbol{U}\left(\sum_{k=1}^{K}\boldsymbol{D}_{s_k} + \boldsymbol{D}_q\right)\boldsymbol{U}^{\mathrm{H}},$$
$$\boldsymbol{X}_{s_k}^\star = \boldsymbol{U}\left(\frac{1}{\gamma_k}\boldsymbol{D}_{s_k} - \sum_{i\neq k}\boldsymbol{D}_{s_i} - \boldsymbol{D}_q\right)\boldsymbol{U}^{\mathrm{H}}. \tag{A.2}$$

We will show below that we can choose the matrices $\boldsymbol{D}_q$ and $\boldsymbol{D}_{s_k}$ to be diagonal.

Lemma 2 (Schur's Complement [43]): Let $\boldsymbol{M} = [\boldsymbol{A}, \boldsymbol{B}; \boldsymbol{B}^{\mathrm{H}}, \boldsymbol{C}]$ be a Hermitian matrix. Then, $\boldsymbol{M} \succeq \mathbf{0}$ if and only if $\boldsymbol{C} - \boldsymbol{B}^{\mathrm{H}}\boldsymbol{A}^{-1}\boldsymbol{B} \succeq \mathbf{0}$ (assuming $\boldsymbol{A} \succ \mathbf{0}$), or $\boldsymbol{A} - \boldsymbol{B}\boldsymbol{C}^{-1}\boldsymbol{B}^{\mathrm{H}} \succeq \mathbf{0}$ (assuming $\boldsymbol{C} \succ \mathbf{0}$).

Lemma 3 (Generalized Schur's Complement [43]): Let $\boldsymbol{M} = [\boldsymbol{A}, \boldsymbol{B}; \boldsymbol{B}^{\mathrm{H}}, \boldsymbol{C}]$ be a Hermitian matrix. Then, $\boldsymbol{M} \succeq \mathbf{0}$ if and only if $\boldsymbol{C} - \boldsymbol{B}^{\mathrm{H}}\boldsymbol{A}^{\dagger}\boldsymbol{B} \succeq \mathbf{0}$ and $\left(\boldsymbol{I} - \boldsymbol{A}\boldsymbol{A}^{\dagger}\right)\boldsymbol{B} = \mathbf{0}$ (assuming $\boldsymbol{A} \succeq \mathbf{0}$), or $\boldsymbol{A} - \boldsymbol{B}\boldsymbol{C}^{\dagger}\boldsymbol{B}^{\mathrm{H}} \succeq \mathbf{0}$ and $\left(\boldsymbol{I} - \boldsymbol{C}\boldsymbol{C}^{\dagger}\right)\boldsymbol{B}^{\mathrm{H}} = \mathbf{0}$ (assuming $\boldsymbol{C} \succeq \mathbf{0}$).

Denote $\boldsymbol{D}_q(i,j)$ and $\boldsymbol{D}_{s_k}(i,j)$ as the (i,j)th elements of $\boldsymbol{D}_q$ and $\boldsymbol{D}_{s_k}$, respectively. Let $\boldsymbol{A}_q = \boldsymbol{X}_q^\star + \mu_q^\star\boldsymbol{I}$, $\boldsymbol{B}_q = \boldsymbol{X}_q^\star\tilde{\boldsymbol{h}}_q$, and $C_q = \tilde{\boldsymbol{h}}_q^{\mathrm{H}}\boldsymbol{X}_q^\star\tilde{\boldsymbol{h}}_q + \sigma_q^2 - Q/\zeta - \mu_q^\star\epsilon_q^2$ for elements in (18b). Then $\boldsymbol{A}_q \succeq \mathbf{0}$ and $C_q \ge \mathbf{0}$. Substituting (A.2) into (18b) and using the definitions of $\boldsymbol{U}$ and $\{\boldsymbol{u}_k\}$, we obtain

$$\boldsymbol{A}_q = \boldsymbol{U}\left(\sum_{k=1}^{K}\boldsymbol{D}_{s_k} + \boldsymbol{D}_q + \mu_q^\star\boldsymbol{I}\right)\boldsymbol{U}^{\mathrm{H}}, \tag{A.3}$$

$$\begin{aligned}\boldsymbol{B}_q &= \boldsymbol{U}\left(\sum_{k=1}^{K}\boldsymbol{D}_{s_k} + \boldsymbol{D}_q\right)\boldsymbol{U}^{\mathrm{H}}\boldsymbol{u}_{K+1}\|\tilde{\boldsymbol{h}}_q\| \\ &= \|\tilde{\boldsymbol{h}}_q\|\boldsymbol{U}\left(\sum_{k=1}^{K}\boldsymbol{D}_{s_k}(:,K+1) + \boldsymbol{D}_q(:,K+1)\right),\end{aligned} \tag{A.4}$$

$$\begin{aligned}C_q &= \boldsymbol{u}_{K+1}^{\mathrm{H}}\|\tilde{\boldsymbol{h}}_q\|\boldsymbol{B}_q + \sigma_q^2 - Q/\zeta - \mu_q^\star\epsilon_q^2 \\ &= \|\tilde{\boldsymbol{h}}_q\|^2\left(\sum_{k=1}^{K}\boldsymbol{D}_{s_k}(K+1,K+1) + \boldsymbol{D}_q(K+1,K+1)\right) \\ &\quad + \sigma_q^2 - Q/\zeta - \mu_q^\star\epsilon_q^2.\end{aligned} \tag{A.5}$$

We then discuss the following two cases for the constraint (18b).

Case 1: $C_q > 0$: From Lemma 2, we know (18b) is equivalent to $\boldsymbol{A}_q - \boldsymbol{B}_q C_q^{-1}\boldsymbol{B}_q^{\mathrm{H}} \succeq \mathbf{0}$. Using (A.3)-(A.5), we obtain

$$\boldsymbol{A}_q - \boldsymbol{B}_q C_q^{-1}\boldsymbol{B}_q^{\mathrm{H}} = \boldsymbol{U}\boldsymbol{D}_{q,\Pi}\boldsymbol{U}^{\mathrm{H}}, \tag{A.6}$$

where

$$\boldsymbol{D}_{q,\Pi} = \left(\sum_{k=1}^{K}\boldsymbol{D}_{s_k} + \boldsymbol{D}_q + \mu_q^\star\boldsymbol{I}\right) - \frac{\|\tilde{\boldsymbol{h}}_q\|^2}{C_q}\left(\sum_{k=1}^{K}\boldsymbol{D}_{s_k}(:,K+1) + \boldsymbol{D}_q(:,K+1)\right)\left(\sum_{k=1}^{K}\boldsymbol{D}_{s_k}(:,K+1) + \boldsymbol{D}_q(:,K+1)\right)^{\mathrm{H}}.$$

Thus, (18b) is equivalent to $\boldsymbol{D}_{q,\Pi} \succeq \mathbf{0}$. In particular, all the diagonal elements of $\boldsymbol{D}_{q,\Pi}$ must be nonnegative, i.e., $\mathrm{diag}(\boldsymbol{D}_{q,\Pi}) \succeq \mathbf{0}$, which results in the requirement

$$\left(\sum_{k=1}^{K}\boldsymbol{D}_{s_k}(i,i) + \boldsymbol{D}_q(i,i) + \mu_q^\star\right) - \frac{\|\tilde{\boldsymbol{h}}_q\|^2}{C_q}\left|\sum_{k=1}^{K}\boldsymbol{D}_{s_k}(i,K+1) + \boldsymbol{D}_q(i,K+1)\right|^2 \ge 0, \tag{A.7}$$

for all $1 \le i \le N$.

Construct the following new solutions

$$\boldsymbol{Q}^* = \boldsymbol{U}\boldsymbol{\Lambda}_q\boldsymbol{U}^{\mathrm{H}}, \quad \boldsymbol{S}_k^* = \boldsymbol{U}\boldsymbol{\Lambda}_{s_k}\boldsymbol{U}^{\mathrm{H}}, \tag{A.8}$$

where $\boldsymbol{\Lambda}_q = \mathrm{diag}(\boldsymbol{D}_q) \succeq \mathbf{0}$ and $\boldsymbol{\Lambda}_{s_k} = \mathrm{diag}(\boldsymbol{D}_{s_k}) \succeq \mathbf{0}$. We next prove that (A.8) also satisfies (18b).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2019.2916405, IEEE Transactions on Wireless Communications

Using (A.2)-(A.5), we have

$$\boldsymbol{A}_q^* = \boldsymbol{U}\mathrm{diag}\left(\sum_{k=1}^{K}\boldsymbol{D}_{s_k}+\boldsymbol{D}_q+\mu_q^\star\boldsymbol{I}\right)\boldsymbol{U}^{\mathrm{H}}$$

$$\boldsymbol{B}_q^* = \|\tilde{\boldsymbol{h}}_q\|\boldsymbol{u}_{K+1}\left(\sum_{k=1}^{K}\boldsymbol{D}_{s_k}(K+1,K+1)+\boldsymbol{D}_q(K+1,K+1)\right),$$

$$C_q^* = C_q.$$

From $\boldsymbol{A}_q^*$, $\boldsymbol{B}_q^*$ and C_q^*, we obtain

$$\boldsymbol{A}_q^* - \boldsymbol{B}_q^* C_q^{*-1}\boldsymbol{B}_q^{*\mathrm{H}} = \boldsymbol{U}\boldsymbol{\Lambda}_{q,\Sigma}\boldsymbol{U}^{\mathrm{H}}, \tag{A.9}$$

where $\boldsymbol{\Lambda}_{q,\Sigma}(i,i)$ can be expressed

$$\sum_{k=1}^{K}\boldsymbol{D}_{s_k}(i,i)+\boldsymbol{D}_q(i,i)+\mu_q^\star, \quad i\neq K+1$$

$$\sum_{k=1}^{K}\boldsymbol{D}_{s_k}(i,i)+\boldsymbol{D}_q(i,i)+\mu_q^\star-\Omega_{K+1,K+1}, \quad i=K+1.$$

From (A.7), $\boldsymbol{\Lambda}_{q,\Sigma}(K+1,K+1)=\mathrm{diag}(\boldsymbol{D}_{q,\Pi})(K+1,K+1)\geq 0$. In addition, $\boldsymbol{\Lambda}_{q,\Sigma}(i,i)\geq\mathrm{diag}(\boldsymbol{D}_{q,\Pi})(i,i)\geq 0,\ i\neq K+1$. Thus, $\boldsymbol{\Lambda}_{q,\Sigma}(i,i)\geq 0$ holds for all $1\leq i\leq N$, and $\boldsymbol{A}_q^*-\boldsymbol{B}_q^*C_q^{*-1}\boldsymbol{B}_q^{*\mathrm{H}}\succeq\boldsymbol{0}$ also holds.

Case 2: $C_q = 0$: From (18b) and Lemma 3, there must be $\boldsymbol{A}_q-\boldsymbol{B}_qC_q^\dagger\boldsymbol{B}_q^{\mathrm{H}}\succeq\boldsymbol{0}$ and $\left(\boldsymbol{I}-C_qC_q^\dagger\right)\boldsymbol{B}_q^{\mathrm{H}}=\boldsymbol{0}$. Since $C_q=0$, $\boldsymbol{B}_q=\boldsymbol{0}$ must holds. Thus, the constraint (18b) can be equivalently expressed as $\boldsymbol{A}_q\succeq\boldsymbol{0}$, such that $\mathrm{diag}\left(\sum_{k=1}^{K}\boldsymbol{D}_{s_k}+\boldsymbol{D}_q+\mu_q^\star\boldsymbol{I}\right)\succeq\boldsymbol{0}$. Thus we have $\boldsymbol{A}_q^*=\boldsymbol{U}\mathrm{diag}\left(\sum_{k=1}^{K}\boldsymbol{D}_{s_k}+\boldsymbol{D}_q+\mu_q^\star\boldsymbol{I}\right)\boldsymbol{U}^{\mathrm{H}}\succeq\boldsymbol{0}$.

Similarly, we can show that the jointly diagonalized $\boldsymbol{S}_k$ and $\boldsymbol{Q}$ in (A.8) also satisfy the constraint (18c). The details are omitted here for brief.

Moreover, it can be readily checked that (A.8) does not change the objective value and at the same time satisfies the constraints (18d)-(18f), (20). Consequently, the solutions (A.8) are also optimal for $\mathbf{P1}-\mathbf{SDR}$, and $\boldsymbol{Q}^\star$ and $\boldsymbol{S}_k^\star$ can be simultaneously diagonalized by $\boldsymbol{U}$ and $\boldsymbol{U}^{\mathrm{H}}$.

B. Part (b): the proof of (21)

Let us show that (21) must hold from contradiction. Using (A.8) and the definitions $\{\boldsymbol{u}_i=\tilde{\boldsymbol{h}}_i/\|\tilde{\boldsymbol{h}}_i\|,\ 1\leq i\leq K\}$ and $\boldsymbol{u}_{K+1}=\tilde{\boldsymbol{h}}_q/\|\tilde{\boldsymbol{h}}_q\|$, we know that $\boldsymbol{Q}^\star$ and $\boldsymbol{S}_k^\star,\ 1\leq j\leq K$ can be further expressed as

$$\boldsymbol{Q}^\star=\sum_{i=1}^{K}P_{q,i}^\star\frac{\tilde{\boldsymbol{h}}_i\tilde{\boldsymbol{h}}_i^{\mathrm{H}}}{\|\tilde{\boldsymbol{h}}_i\|^2}+P_{q,K+1}^\star\frac{\tilde{\boldsymbol{h}}_q\tilde{\boldsymbol{h}}_q^{\mathrm{H}}}{\|\tilde{\boldsymbol{h}}_q\|^2}+\sum_{j=K+2}^{N}P_{q,j}^\star\boldsymbol{u}_j\boldsymbol{u}_j^{\mathrm{H}},$$

$$\boldsymbol{S}_k^\star=\sum_{i=1}^{K}P_{s_k,i}^\star\frac{\tilde{\boldsymbol{h}}_i\tilde{\boldsymbol{h}}_i^{\mathrm{H}}}{\|\tilde{\boldsymbol{h}}_i\|^2}+P_{s_k,K+1}^\star\frac{\tilde{\boldsymbol{h}}_q\tilde{\boldsymbol{h}}_q^{\mathrm{H}}}{\|\tilde{\boldsymbol{h}}_q\|^2}+\sum_{j=K+2}^{N}P_{s_k,j}^\star\boldsymbol{u}_j\boldsymbol{u}_j^{\mathrm{H}},$$

where $P_{q,j}^\star\geq 0$ and $P_{s_k,j}^\star\geq 0$, for same $j\in\{K+2,\ldots,N\}$. We can then set $\{P_{q,j}^\star=0\}$ and $\{P_{s_k,j}^\star=0\}$ to construct the following new solutions

$$\boldsymbol{Q}^*=\sum_{i=1}^{K+1}P_{q,i}^\star\boldsymbol{u}_i\boldsymbol{u}_i^{\mathrm{H}},\quad \boldsymbol{S}_k^*=\sum_{i=1}^{K+1}P_{s_k,i}^\star\boldsymbol{u}_i\boldsymbol{u}_i^{\mathrm{H}}. \tag{A.10}$$

Similarly as the proof of Part (a), it can be derived from $\mathbf{P1}-\mathbf{SDR}$ that $\boldsymbol{Q}^*$ and $\{\boldsymbol{S}_k^*\}$ always satisfy all constraints of $\mathbf{P1}-\mathbf{SDR}$, but provide smaller objective value. As a result, $\boldsymbol{Q}^*$ and $\{\boldsymbol{S}_k^*\}$ are better solutions than $\boldsymbol{Q}^\star$ and $\{\boldsymbol{S}_k^\star\}$, which contradicts with our first place assumption. Thus, (21) must hold, which completes the proof of Proposition 1.

APPENDIX B
PROOF OF PROPOSITION 2

First, we prove that $\{\mu_q>0\}$ must hold via contradiction. Assuming $\mu_q^\star=0$, it follows from (18b) that

$$\boldsymbol{\Upsilon}_q=\begin{bmatrix}\boldsymbol{X}_q & \boldsymbol{X}_q\tilde{\boldsymbol{h}}_q\\ \tilde{\boldsymbol{h}}_q^{\mathrm{H}}\boldsymbol{X}_q^{\mathrm{H}} & \tilde{\boldsymbol{h}}_q^{\mathrm{H}}\boldsymbol{X}_q\tilde{\boldsymbol{h}}_q+\sigma_q^2-Q/\zeta\end{bmatrix}\succeq\boldsymbol{0}, \tag{B.1}$$

must be satisfied. Left and right multiplying both sides of $\boldsymbol{\Upsilon}_q$ by $[-\tilde{\boldsymbol{h}}_q^{\mathrm{H}}\ 1]$ and $[-\tilde{\boldsymbol{h}}_q^{\mathrm{H}}\ 1]^{\mathrm{H}}$ yields

$$[-\tilde{\boldsymbol{h}}_q^{\mathrm{H}}\ 1]\boldsymbol{\Upsilon}_q[-\tilde{\boldsymbol{h}}_q^{\mathrm{H}}\ 1]^{\mathrm{H}}=\sigma_q^2-Q/\zeta\geq 0, \tag{B.2}$$

which, however, cannot be true due to $\sigma_q^2-Q/\zeta<0$ (see $\mathbf{P1}$ for details). Thus, there must be $\mu_q>0$ in $\mathbf{P1}-\mathbf{SDR}$.

Second, due to the fact that $\{\boldsymbol{S}_k\succeq\boldsymbol{0}\}$ and $\boldsymbol{Q}\succeq\boldsymbol{0}$, there holds

$$\boldsymbol{X}_q+\mu_q\boldsymbol{I}=\sum_{k=1}^{K}\boldsymbol{S}_k+\boldsymbol{Q}+\mu_q\boldsymbol{I}\succ\boldsymbol{0}, \tag{B.3}$$

and hence $(\boldsymbol{X}_q+\mu_q\boldsymbol{I})^{-1}$ exists. Define $\tilde{\boldsymbol{H}}_q=\tilde{\boldsymbol{h}}_q\tilde{\boldsymbol{h}}_q^{\mathrm{H}}$, the constraint in (18b) of $\mathbf{P1}-\mathbf{SDR}$ can be equivalently expressed as

$$\begin{aligned}&\tilde{\boldsymbol{h}}_q^{\mathrm{H}}\boldsymbol{X}_q\tilde{\boldsymbol{h}}_q+\sigma_q^2-Q/\zeta-\mu_q\epsilon_q^2-\tilde{\boldsymbol{h}}_q^{\mathrm{H}}\boldsymbol{X}_q^{\mathrm{H}}\left(\boldsymbol{X}_q+\mu_q\boldsymbol{I}\right)^{-1}\boldsymbol{X}_q\tilde{\boldsymbol{h}}_q\\ &=\mathrm{Tr}\left\{\left(\tilde{\boldsymbol{H}}_q\boldsymbol{X}_q\right)\left[\boldsymbol{I}-\left(\boldsymbol{X}_q+\mu_q\boldsymbol{I}\right)^{-1}\boldsymbol{X}_q\right]\right\}+\sigma_q^2-Q/\zeta-\mu_q\epsilon_q^2\\ &=\mu_q\mathrm{Tr}\left[\tilde{\boldsymbol{H}}_q\boldsymbol{X}_q\left(\boldsymbol{X}_q+\mu_q\boldsymbol{I}\right)^{-1}\right]+\sigma_q^2-Q/\zeta-\mu_q\epsilon_q^2\geq 0.\end{aligned} \tag{B.4}$$

where $\boldsymbol{X}_q+\mu_q\boldsymbol{I}$ is equivalent to

$$\sum_{i=1}^{K+1}\left(P_{q,i}+\sum_{k=1}^{K}P_{s_k,i}+\mu_q\right)\boldsymbol{u}_i\boldsymbol{u}_i^{\mathrm{H}}+\sum_{i=K+2}^{N}\mu_q\boldsymbol{u}_i\boldsymbol{u}_i^{\mathrm{H}}. \tag{B.5}$$

Noting that $\tilde{\boldsymbol{H}}_q=\|\tilde{\boldsymbol{h}}_q\|^2\tilde{\boldsymbol{u}}_{K+1}\tilde{\boldsymbol{u}}_{K+1}^{\mathrm{H}}$ and substituting (B.5) into (B.4), we obtain

$$\frac{\mu_q\|\tilde{\boldsymbol{h}}_q\|^2\left(P_{q,K+1}+\sum_{k=1}^{K}P_{s_k,K+1}\right)}{\mu_q+P_{q,K+1}+\sum_{k=1}^{K}P_{s_k,K+1}}+\sigma_q^2-Q/\zeta-\mu_q\epsilon_q^2\geq 0,$$

completing the proof of Proposition 2.

APPENDIX C
PROOF OF PROPOSITION 3

First, we show that $\{\mu_{s_k}>0\}$ must hold via contradiction. Assuming $\mu_{s_k}=0$ for some $1\leq k\leq K$, it follows from (18c) that

$$\boldsymbol{\Upsilon}_{s_k}=\begin{bmatrix}\boldsymbol{X}_{s_k} & \boldsymbol{X}_{s_k}\tilde{\boldsymbol{h}}_k\\ \tilde{\boldsymbol{h}}_k^{\mathrm{H}}\boldsymbol{X}_{s_k}^{\mathrm{H}} & \tilde{\boldsymbol{h}}_k^{\mathrm{H}}\boldsymbol{X}_{s_k}\tilde{\boldsymbol{h}}_k-\sigma_k^2\end{bmatrix}\succeq\boldsymbol{0}, \tag{C.1}$$

must be satisfied. Left and right multiplying both sides of $\boldsymbol{\Upsilon}_{s_k}$ by $[-\tilde{\boldsymbol{h}}_k^{\mathrm{H}}\ 1]$ and $[-\tilde{\boldsymbol{h}}_k^{\mathrm{H}}\ 1]^{\mathrm{H}}$, respectively, yields $[-\tilde{\boldsymbol{h}}_k^{\mathrm{H}}\ 1]\boldsymbol{\Upsilon}_{s_k}[-\tilde{\boldsymbol{h}}_k^{\mathrm{H}}\ 1]^{\mathrm{H}} = -\sigma_k^2 \geq 0$, which cannot be true due to $\sigma_k^2 > 0$. Thus, there must be $\mu_{s_k} > 0$ for all $1 \leq k \leq K$ in $\mathbf{P1-SDR}$.

Second, let us show that $\mathrm{Rank}\left(\boldsymbol{X}_{s_k}^\star + \mu_{s_k}^\star \boldsymbol{I}\right) \geq 1$ by contradiction. Assume $\mathrm{Rank}\left(\boldsymbol{X}_{s_k}^\star + \mu_{s_k}^\star \boldsymbol{I}\right) = 0$ or $\boldsymbol{X}_{s_k}^\star + \mu_{s_k}^\star \boldsymbol{I} = \boldsymbol{0}$ at the optimal point. We know from (18c) that

$$\begin{bmatrix} \boldsymbol{0} & \boldsymbol{X}_{s_k}^\star \tilde{\boldsymbol{h}}_k \\ \tilde{\boldsymbol{h}}_k^{\mathrm{H}} \boldsymbol{X}_{s_k}^{\star\mathrm{H}} & \tilde{\boldsymbol{h}}_k^{\mathrm{H}}(-\mu_{s_k}^\star \boldsymbol{I})\tilde{\boldsymbol{h}}_k - \sigma_k^2 - \mu_{s_k}^\star \epsilon_k^2 \end{bmatrix} \succeq \boldsymbol{0}. \quad \text{(C.2)}$$

However, (C.2) cannot be true due to $\mu_{s_k}^\star > 0$ and $\tilde{\boldsymbol{h}}_k^{\mathrm{H}}(-\mu_{s_k}^\star \boldsymbol{I})\tilde{\boldsymbol{h}}_k - \sigma_k^2 - \mu_{s_k}^\star \epsilon_k^2 < 0$. Since $\mathrm{Rank}\left(\boldsymbol{X}_{s_k}^\star + \mu_{s_k}^\star \boldsymbol{I}\right) \geq 1$, for all $k \in \{1, \ldots, K\}$, the Moore-Penrose inverses [42] of $\boldsymbol{X}_{s_k} + \mu_{s_k} \boldsymbol{I}$, denoted as $(\boldsymbol{X}_{s_k} + \mu_{s_k}\boldsymbol{I})^\dagger$, must exist.

Using Lemma 3, (18c) of $\mathbf{P1-SDR}$ can be equivalently expressed as

$$\left[\boldsymbol{I} - \left(\boldsymbol{X}_{s_k} + \mu_{s_k}\boldsymbol{I}\right)\left(\boldsymbol{X}_{s_k} + \mu_{s_k}\boldsymbol{I}\right)^\dagger\right] \boldsymbol{X}_{s_k}\tilde{\boldsymbol{h}}_k = \boldsymbol{0}, \quad \text{(C.3)}$$

$$\tilde{\boldsymbol{h}}_k^{\mathrm{H}} \boldsymbol{X}_{s_k}\tilde{\boldsymbol{h}}_k - \sigma_k^2 - \mu_{s_k}\epsilon_k^2 - \tilde{\boldsymbol{h}}_k^{\mathrm{H}} \boldsymbol{X}_{s_k}^{\mathrm{H}}\left(\boldsymbol{X}_{s_k} + \mu_{s_k}\boldsymbol{I}\right)^\dagger \boldsymbol{X}_{s_k}\tilde{\boldsymbol{h}}_k \geq 0, \quad \text{(C.4)}$$

$$\boldsymbol{X}_{s_k} + \mu_{s_k}\boldsymbol{I} \succeq \boldsymbol{0}. \quad \text{(C.5)}$$

Next, we show that the constraint (C.3) can be eliminated. Define $\tilde{\boldsymbol{H}}_k = \tilde{\boldsymbol{h}}_k \tilde{\boldsymbol{h}}_k^{\mathrm{H}}$. Since $\boldsymbol{X}_{s_k} = \boldsymbol{X}_{s_k}^{\mathrm{H}}$, (C.4) is equivalent to

$$\begin{aligned} &\tilde{\boldsymbol{h}}_k^{\mathrm{H}} \boldsymbol{X}_{s_k}\tilde{\boldsymbol{h}}_k - \sigma_k^2 - \mu_{s_k}\epsilon_k^2 - \tilde{\boldsymbol{h}}_k^{\mathrm{H}} \boldsymbol{X}_{s_k}^{\mathrm{H}}\left(\boldsymbol{X}_{s_k} + \mu_{s_k}\boldsymbol{I}\right)^\dagger \boldsymbol{X}_{s_k}\tilde{\boldsymbol{h}}_k \\ =&\mathrm{Tr}\left\{\left(\tilde{\boldsymbol{H}}_k \boldsymbol{X}_{s_k}\right)\left[\boldsymbol{I} - \left(\boldsymbol{X}_{s_k} + \mu_{s_k}\boldsymbol{I}\right)^\dagger \boldsymbol{X}_{s_k}\right]\right\} - \sigma_k^2 - \mu_{s_k}\epsilon_k^2 \\ =&\mu_{s_k}\mathrm{Tr}\left[\tilde{\boldsymbol{H}}_k \boldsymbol{X}_{s_k}\left(\boldsymbol{X}_{s_k} + \mu_{s_k}\boldsymbol{I}\right)^\dagger\right] - \sigma_k^2 - \mu_{s_k}\epsilon_k^2 \geq 0, \quad \text{(C.6)} \end{aligned}$$

where in particular, there must be $\tilde{\boldsymbol{H}}_k \boldsymbol{X}_{s_k} \neq \boldsymbol{0}$. Moreover, from (21), there is

$$\boldsymbol{X}_{s_k} = \sum_{i=1}^{K+1}\left(\frac{P_{s_k,i}}{\gamma_k} - \sum_{j\neq k} P_{s_j,i} - P_{q,i}\right)\boldsymbol{u}_i\boldsymbol{u}_i^{\mathrm{H}}. \quad \text{(C.7)}$$

Since $\tilde{\boldsymbol{H}}_k \boldsymbol{X}_{s_k} \neq \boldsymbol{0}$ we have $P_{s_k,k}/\gamma_k - \sum_{j\neq k} P_{s_j,k} - P_{q,k} > 0$. Consequently, the kth singular value of $\boldsymbol{I} - \left(\boldsymbol{X}_{s_k} + \mu_{s_k}\boldsymbol{I}\right)\left(\boldsymbol{X}_{s_k} + \mu_{s_k}\boldsymbol{I}\right)^\dagger$ is equal to zero, which implies that (C.3) and (C.4) are equivalent to (C.6). Thus, (18c) of $\mathbf{P1-SDR}$ can be expressed as (C.5) and (C.6).

Substituting (C.7) into (C.5) and (C.6), we know that (18c) can be equivalently expressed as

$$\frac{\mu_{s_k}\|\tilde{\boldsymbol{h}}_k\|^2\left(P_{s_k,k}/\gamma_k - \sum_{i\neq k} P_{s_i,k} - P_{q,k}\right)}{\mu_{s_k} + P_{s_k,k}/\gamma_k - \sum_{i\neq k} P_{s_i,k} - P_{q,k}} - \sigma_k^2 - \mu_{s_k}\epsilon_k^2 \geq 0,$$

$$\frac{P_{s_k,i}}{\gamma_k} - \sum_{j\neq k} P_{s_j,i} - P_{q,i} + \mu_{s_k} \geq 0, \quad \forall 1 \leq i \leq K+1.$$

The proof of Proposition 3 is completed.

Appendix D
Proof of Theorem 1

Assuming $\mu_q^\star, \{\mu_{s_k}^\star\}, \boldsymbol{Q}^\star, \{\boldsymbol{S}_k^\star\}$ are the optimal solutions of $\mathbf{P1-SDR}$, it has been shown in Proposition 1 that $\boldsymbol{Q}^\star, \{\boldsymbol{S}_k^\star\}$ must have the following form

$$\begin{cases} \boldsymbol{Q}^\star = \sum_{i=1}^{K} P_{q,i}^\star \frac{\tilde{\boldsymbol{h}}_i \tilde{\boldsymbol{h}}_i^{\mathrm{H}}}{\|\tilde{\boldsymbol{h}}_i\|^2} + P_{q,K+1}^\star \frac{\tilde{\boldsymbol{h}}_q \tilde{\boldsymbol{h}}_q^{\mathrm{H}}}{\|\tilde{\boldsymbol{h}}_q\|^2}, \\ \boldsymbol{S}_k^\star = \sum_{i=1}^{K} P_{s_k,i}^\star \frac{\tilde{\boldsymbol{h}}_i \tilde{\boldsymbol{h}}_i^{\mathrm{H}}}{\|\tilde{\boldsymbol{h}}_i\|^2} + P_{s_k,K+1}^\star \frac{\tilde{\boldsymbol{h}}_q \tilde{\boldsymbol{h}}_q^{\mathrm{H}}}{\|\tilde{\boldsymbol{h}}_q\|^2}. \end{cases} \quad \text{(D.1)}$$

When $\min\{\mu_{s_k}^\star\} \geq \sum_{i=1}^K P_{s_i,K+1}^\star + P_{q,K+1}^\star$, we will prove $\mathrm{Rank}\left(\boldsymbol{Q}^\star\right) \leq 1$ and $\mathrm{Rank}\left(\boldsymbol{S}_k^\star\right) \leq 1$, $\forall k \in \{1, \ldots, K\}$ via contradiction. Assume $\mathrm{Rank}\left(\boldsymbol{Q}^\star\right) \geq 2$ or $\mathrm{Rank}\left(\boldsymbol{S}_k^\star\right) \geq 2$. From (D.1) this means that there are at least two nonzero coefficients in $\{P_{q,i}^\star, 1 \leq i \leq K+1\}$ or in $\{P_{s_k,i}^\star, 1 \leq i \leq K+1\}$.

Let the transmit power for the energy user be

$$P_q^* = \sum_{i=1}^{K} P_{s_i,K+1}^\star + P_{q,K+1}^\star, \quad \text{(D.2)}$$

and the transmit power for the information users be

$$P_k^* = P_{s_k,k}^\star - \sum_{i\neq k}^{K} P_{s_i,k}^\star - P_{q,k}^\star. \quad \text{(D.3)}$$

Construct a sequence of new solutions as follows

$$\boldsymbol{Q}^* = P_q^* \frac{\tilde{\boldsymbol{h}}_q \tilde{\boldsymbol{h}}_q^{\mathrm{H}}}{\|\tilde{\boldsymbol{h}}_q\|^2}, \boldsymbol{S}_k^* = P_k^* \frac{\tilde{\boldsymbol{h}}_k \tilde{\boldsymbol{h}}_k^{\mathrm{H}}}{\|\tilde{\boldsymbol{h}}_k\|^2}, k = 1, 2, \ldots, K, \quad \text{(D.4)}$$

which satisfy $\mathrm{Rank}\left(\boldsymbol{Q}^*\right) \leq 1$ and $\mathrm{Rank}\left(\boldsymbol{S}_k^*\right) \leq 1$, for all $k \in \{1, \ldots, K\}$. With (D.4), $\boldsymbol{X}_q^* = \sum_{k=1}^{K} \boldsymbol{S}_k^* + \boldsymbol{Q}^*$ and $\boldsymbol{X}_{s_k}^* = \frac{1}{\gamma_k}\boldsymbol{S}_k^* - \sum_{i\neq k} \boldsymbol{S}_i^* - \boldsymbol{Q}^*$ can be expressed as

$$\boldsymbol{X}_q^* = \sum_{i=1}^{K} P_i^* \boldsymbol{u}_i \boldsymbol{u}_i^{\mathrm{H}} + P_q^* \boldsymbol{u}_{K+1}\boldsymbol{u}_{K+1}^{\mathrm{H}}, \quad \text{(D.5)}$$

$$\begin{aligned} \boldsymbol{X}_{s_k}^* = &-\sum_{i=1}^{k-1} P_i^* \boldsymbol{u}_i \boldsymbol{u}_i^{\mathrm{H}} + P_k^*/\gamma_k \boldsymbol{u}_k \boldsymbol{u}_k^{\mathrm{H}} - \sum_{i=k+1}^{K} P_i^* \boldsymbol{u}_i \boldsymbol{u}_i^{\mathrm{H}} \\ &- P_q^* \boldsymbol{u}_{K+1}\boldsymbol{u}_{K+1}^{\mathrm{H}}. \end{aligned} \quad \text{(D.6)}$$

Substituting (D.2)-(D.6) into $\mathbf{P1-SDR}$, we obtain from (18a) that

$$\mathrm{Tr}(\boldsymbol{Q}^*) + \sum_{k=1}^{K}\mathrm{Tr}(\boldsymbol{S}_k^*) \leq \mathrm{Tr}(\boldsymbol{Q}^\star) + \sum_{k=1}^{K}\mathrm{Tr}(\boldsymbol{S}_k^\star), \quad \text{(D.7)}$$

where $\mathrm{Tr}(\boldsymbol{Q}^*) + \sum_{k=1}^{K}\mathrm{Tr}(\boldsymbol{S}_k^*) < \mathrm{Tr}(\boldsymbol{Q}^\star) + \sum_{k=1}^{K}\mathrm{Tr}(\boldsymbol{S}_k^\star)$ if $\mathrm{Rank}\left(\boldsymbol{Q}^\star\right) \geq 2$.

Using Proposition 2, we know from (18b) that

$$\begin{aligned}&\mu_q^\star \mathrm{Tr}\left[\tilde{\boldsymbol{H}}_q \boldsymbol{X}_q^* \left(\boldsymbol{X}_q^* + \mu_q^\star \boldsymbol{I}\right)^{-1}\right] + \sigma_q^2 - Q/\zeta - \mu_q^\star \epsilon_q^2 \\ =&\frac{\mu_q^\star \rho_q \left(P_{q,K+1}^\star + \sum_{k=1}^K P_{s_k,K+1}^\star\right)}{\mu_q^\star + P_{q,K+1}^\star + \sum_{k=1}^K P_{s_k,K+1}^\star} + \sigma_q^2 - Q/\zeta - \mu_q^\star \epsilon_q^2 \\ =&\mu_q^\star \mathrm{Tr}\left[\tilde{\boldsymbol{H}}_q \boldsymbol{X}_q^\star \left(\boldsymbol{X}_q^\star + \mu_q^\star \boldsymbol{I}\right)^{-1}\right] + \sigma_q^2 - Q/\zeta - \mu_q^\star \epsilon_q^2 \geq 0. \end{aligned} \tag{D.8}$$

Using Proposition 3, we know from (18c) that

$$\begin{aligned}&\mu_{s_k}^\star \mathrm{Tr}\left[\tilde{\boldsymbol{H}}_k \boldsymbol{X}_{s_k}^* \left(\boldsymbol{X}_{s_k}^* + \mu_{s_k}^\star \boldsymbol{I}\right)^{\dagger}\right] - \sigma_k^2 - \mu_{s_k}^\star \epsilon_k^2 \\ =&\frac{\mu_{s_k}^\star \rho_k \left(P_{s_k,k}^\star - \sum_{i\neq k}^K P_{s_i,k}^\star - P_{q,k}^\star\right)/\gamma_k}{\mu_{s_k}^\star + \left(P_{s_k,k}^\star - \sum_{i\neq k}^K P_{s_i,k}^\star - P_{q,k}^\star\right)/\gamma_k} - \sigma_k^2 - \mu_{s_k}^\star \epsilon_k^2 \\ \geq&\frac{\mu_{s_k}^\star \rho_k \left(P_{s_k,k}^\star/\gamma_k - \sum_{i\neq k} P_{s_i,k}^\star - P_{q,k}^\star\right)}{\mu_{s_k}^\star + P_{s_k,k}^\star/\gamma_k - \sum_{i\neq k} P_{s_i,k}^\star - P_{q,k}^\star} - \sigma_k^2 - \mu_{s_k}^\star \epsilon_k^2 \\ =&\mu_{s_k}^\star \mathrm{Tr}\left[\tilde{\boldsymbol{H}}_k \boldsymbol{X}_{s_k}^\star \left(\boldsymbol{X}_{s_k}^\star + \mu_{s_k}^\star \boldsymbol{I}\right)^{\dagger}\right] - \sigma_k^2 - \mu_{s_k}^\star \epsilon_k^2 \geq 0, \end{aligned} \tag{D.9}$$

where $\mu_{s_k}^\star \mathrm{Tr}\left[\tilde{\boldsymbol{H}}_k \boldsymbol{X}_{s_k}^* \left(\boldsymbol{X}_{s_k}^* + \mu_{s_k}^\star \boldsymbol{I}\right)^{\dagger}\right] - \sigma_k^2 - \mu_{s_k}^\star \epsilon_k^2 > \mu_{s_k}^\star \mathrm{Tr}\left[\tilde{\boldsymbol{H}}_k \boldsymbol{X}_{s_k}^\star \left(\boldsymbol{X}_{s_k}^\star + \mu_{s_k}^\star \boldsymbol{I}\right)^{\dagger}\right] - \sigma_k^2 - \mu_{s_k}^\star \epsilon_k^2$ if $\mathrm{Rank}\left(\boldsymbol{Q}^\star\right) \geq 2$ or $\mathrm{Rank}\left(\boldsymbol{S}_k^\star\right) \geq 2$.

Moreover, using Proposition 3 and from (18c), we obtain that

$$\begin{aligned}\boldsymbol{X}_{s_k}^* + \mu_{s_k}^\star \boldsymbol{I} =& -\sum_{i=1}^{k-1} P_i^* \boldsymbol{u}_i \boldsymbol{u}_i^{\mathrm{H}} + P_k^*/\gamma_k \boldsymbol{u}_k \boldsymbol{u}_k^{\mathrm{H}} \\ &- \sum_{i=k+1}^{K} P_i^* \boldsymbol{u}_i \boldsymbol{u}_i^{\mathrm{H}} - P_q^* \boldsymbol{u}_{K+1} \boldsymbol{u}_{K+1}^{\mathrm{H}} + \mu_{s_k}^\star \boldsymbol{I}, \end{aligned}$$

where

(a) For $j = k$, since $P_k^*/\gamma_k \geq P_{s_k,k}^\star/\gamma_k - \sum_{i\neq k} P_{s_i,k}^\star - P_{q,k}^\star$, there is $P_k^*/\gamma_k + \mu_{s_k} \geq \frac{P_{s_k,k}}{\gamma_k} - \sum_{j\neq k} P_{s_j,k} - P_{q,k} + \mu_{s_k} \geq 0$

(b) For $j = 1, \ldots, k, k+1, \ldots, K$, since $\gamma_k > 1$, there is $-P_j^* + \mu_{s_k}^\star \geq P_{s_k,j}^\star/\gamma_k - \sum_{i\neq k} P_{s_i,j}^\star - P_{q,j}^\star + \mu_{s_k}^\star \geq 0$.

(c) For $j = K+1$, using the assumption that $\min\{\mu_{s_k}^\star\} \geq \sum_{i=1}^K P_{s_i,K+1}^\star + P_{q,K+1}^\star$, there is $-P_q^* + \mu_{s_k}^\star \geq 0$.

Consequently, there holds

$$\boldsymbol{X}_{s_k}^* + \mu_{s_k}^\star \boldsymbol{I} \succeq \boldsymbol{0}. \tag{D.10}$$

From (18d)-(18f) and (20), there are

$$\boldsymbol{X}_q^* = \sum_{k=1}^K \boldsymbol{S}_k^* + \boldsymbol{Q}^*,\ \boldsymbol{X}_{s_k}^* = \frac{1}{\gamma_k}\boldsymbol{S}_k^* - \sum_{i\neq k} \boldsymbol{S}_i^* - \boldsymbol{Q}^*, \tag{D.11}$$

$$\mu_q^\star \geq 0, \quad \mu_{s_k}^\star \geq 0, \quad \boldsymbol{Q}^* \succeq \boldsymbol{0}, \quad \boldsymbol{S}_k^* \succeq \boldsymbol{0}, \tag{D.12}$$

where $k = 1, 2, \ldots, K$. Based on (D.7)$\sim$(D.12), we know that $\boldsymbol{Q}^*$ and $\{\boldsymbol{S}_k^*\}$ satisfy all the constraints of $\mathbf{P1}-\mathbf{SDR}$ and provide a smaller trace than $\boldsymbol{Q}^\star$ and $\{\boldsymbol{S}_k^\star\}$. Thus, $\boldsymbol{Q}^*$ and $\{\boldsymbol{S}_k^*\}$ are better solutions, which contradicts the assumption that $\boldsymbol{Q}^\star$ and $\{\boldsymbol{S}_k^\star\}$ are optimal solutions. Consequently, $\mathrm{Rank}\left(\boldsymbol{Q}^\star\right) \leq 1$ and $\mathrm{Rank}\left(\boldsymbol{S}_k^\star\right) \leq 1$, for all $k \in \{1, \ldots, K\}$, completing the proof of Theorem 1.

APPENDIX E
PROOF OF PROPOSITION 4

The Lagrangian of $\mathbf{P2}$ is expressed as

$$\begin{aligned}&\mathcal{L}\left(\mu_q, \{\mu_{s_k}\}, P_q, \{P_k\}, \varpi, \{\xi_k\}, \varrho_q, \{\varrho_k\}, \upsilon_q, \{\upsilon_k\}\right) \\ =&P_q + \sum_{k=1}^K P_k - \varpi\left(\frac{\mu_q P_q \|\tilde{\boldsymbol{h}}_q\|^2}{\mu_q + P_q} + \sigma_q^2 - Q/\zeta - \mu_q \epsilon_q^2\right) \\ &- \sum_{k=1}^K \left[\xi_k \left(\frac{\mu_{s_k} P_k \|\tilde{\boldsymbol{h}}_k\|^2}{\mu_{s_k}\gamma_k + P_k} - \sigma_k^2 - \mu_{s_k}\epsilon_k^2\right)\right] \\ &- \varrho_p \mu_q - \varrho_k \mu_{s_k} - \upsilon_q P_q - \upsilon_k P_k, \end{aligned}$$

where $\varpi \geq 0$ and $\{\xi_k \geq 0\}$ are the dual variables associated with the constraints in (23b) and (23c), respectively, $\varrho_q \geq 0$, $\{\varrho_k \geq 0\}$, $\upsilon_q \geq 0$ and $\{\upsilon_k \geq 0\}$ are the dual variables associated with the constraints in (23d). The KKT conditions related to μ_q, $\{\mu_{s_k}\}$, P_q and $\{P_k\}$ can be formulated as

$$\frac{\partial \mathcal{L}}{\partial \mu_q^\star} = \varpi^\star \epsilon_q^2 - \frac{\varpi^\star P_q^{\star 2} \|\tilde{\boldsymbol{h}}_q\|^2}{(\mu_q^\star + P_q^\star)^2} - \varrho_q^\star = 0, \tag{E.1}$$

$$\frac{\partial \mathcal{L}}{\partial \mu_{s_k}^\star} = \xi_k^\star \epsilon_k^2 - \frac{\xi_k^\star P_k^{\star 2} \|\tilde{\boldsymbol{h}}_k\|^2}{(\mu_{s_k}^\star \gamma_k + P_k^\star)^2} - \varrho_k^\star = 0, \tag{E.2}$$

$$\frac{\partial \mathcal{L}}{\partial P_q^\star} = 1 - \frac{\varpi^\star \mu_q^{\star 2} \|\tilde{\boldsymbol{h}}_q\|^2}{(\mu_q^\star + P_q^\star)^2} - \upsilon_q^\star = 0, \tag{E.3}$$

$$\frac{\partial \mathcal{L}}{\partial P_k^\star} = 1 - \frac{\xi_k^\star \mu_{s_k}^{\star 2} \gamma_k \|\tilde{\boldsymbol{h}}_k\|^2}{(\mu_{s_k}^\star \gamma_k + P_k^\star)^2} - \upsilon_k^\star = 0, \tag{E.4}$$

$$\varpi^\star \left(\frac{\mu_q^\star P_q^\star \|\tilde{\boldsymbol{h}}_q\|^2}{\mu_q^\star + P_q^\star} + \sigma_q^2 - Q/\zeta - \mu_q^\star \epsilon_q^2\right) = 0, \tag{E.5}$$

$$\xi_k^\star \left(\frac{\mu_{s_k}^\star P_k^\star \|\tilde{\boldsymbol{h}}_k\|^2}{\mu_{s_k}^\star \gamma_k + P_k^\star} - \sigma_k^2 - \mu_{s_k}^\star \epsilon_k^2\right) = 0, \tag{E.6}$$

$$\varrho_q^\star \mu_q^\star = 0, \quad \varrho_k^\star \mu_{s_k}^\star = 0, \quad \upsilon_q^\star P_q^\star = 0, \quad \upsilon_k^\star P_k^\star = 0, \tag{E.7}$$

where $1 \leq k \leq K$, $\mu_q^\star > 0$, $\{\mu_{s_k}^\star > 0\}$, $P_q^\star > 0$ and $\{P_k^\star > 0\}$ are the optimal primal variables, while $\varpi^\star \geq 0$, $\{\xi_k^\star \geq 0\}$, $\varrho_q^\star = 0$, $\{\varrho_k^\star = 0\}$, $\upsilon_q^\star = 0$, and $\{\upsilon_k^\star = 0\}$ are the optimal dual variables.

From (E.3) and (E.4), we must have $\varpi^\star > 0$ and $\{\xi_k^\star > 0\}$. From (E.5) and (E.6), there are

$$\frac{\mu_q^\star P_q^\star \|\tilde{\boldsymbol{h}}_q\|^2}{\mu_q^\star + P_q^\star} + \sigma_q^2 - Q/\zeta - \mu_q^\star \epsilon_q^2 = 0, \tag{E.8}$$

$$\frac{\mu_{s_k}^\star P_k^\star \|\tilde{\boldsymbol{h}}_k\|^2}{\mu_{s_k}^\star \gamma_k + P_k^\star} - \sigma_k^2 - \mu_{s_k}^\star \epsilon_k^2 = 0,\ 1 \leq k \leq K. \tag{E.9}$$

Moreover, it is easily derived from (E.1) and (E.2) that

$$\mu_q^\star = \frac{P_q^\star \|\tilde{\boldsymbol{h}}_q\| - P_q^\star \epsilon_q}{\epsilon_q}, \mu_{s_k}^\star = \frac{P_k^\star \|\tilde{\boldsymbol{h}}_k\| - P_k^\star \epsilon_k}{\gamma_k \epsilon_k}. \tag{E.10}$$

Substituting (E.10) into (E.8) and (E.9), respectively, we obtain

$$P_q^\star = \frac{Q/\zeta - \sigma_q^2}{(\|\tilde{\boldsymbol{h}}_q\| - \epsilon_q)^2},\ P_k^\star = \frac{\gamma_k \sigma_k^2}{(\|\tilde{\boldsymbol{h}}_k\| - \epsilon_k)^2}. \tag{E.11}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2019.2916405, IEEE Transactions on Wireless Communications

Finally, plugging $P_k^\star$ and (E.10) into (24), it is easily known that the solutions in (E.11) are also optimal for $\mathbf{P1}$ when

$$\min\left\{\frac{\sigma_k^2}{(\|\tilde{\boldsymbol{h}}_k\|-\epsilon_k)\epsilon_k}\right\} \geq \frac{Q/\zeta-\sigma_q^2}{(\|\tilde{\boldsymbol{h}}_q\|-\epsilon_q)^2}, \tag{E.12}$$

which completing the proof of Proposition 4.

References

[1] L. R. Varshney, "Transporting information and energy simultaneously," in *IEEE Int. Symp. Inf. Th.*, pp. 1612–1616, 2008.

[2] R. Zhang and C. K. Ho, "MIMO broadcasting for simultaneous wireless information and power transfer," *IEEE Trans. Wireless Commun.*, vol. 12, pp. 1989–2001, May 2013.

[3] X. Zhou, C. K. Ho, and R. Zhang, "Wireless power meets energy harvesting: a joint energy allocation approach in OFDM-based system," *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 3481–3491, May 2016

[4] C. Zhong, H. Liang, H. Lin, H. A. Suraweera, F. Qu, and Z. Zhang, "Energy beamformer and time split design for wireless powered two-way relaying systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 6, pp. 3723–3736, June 2018.

[5] N. Zhao, Y. Cao, F. R. Yu, Y. Chen, M. Jin, and V. C. M. Leung, "Artificial noise assisted secure interference networks with wireless power transfer," *IEEE Trans. Veh. Technol.*, vol. 67, no. 2, pp. 1087–1098, Feb. 2018.

[6] Y. Chen, N. Zhao, and M. Alouini, "Wireless energy harvesting using signals from multiple fading channels," *IEEE Trans. Commun.*, vol. 65, no. 11, pp. 5027–5039, Nov. 2017.

[7] A. Nasir, X. Zhou, S. Durrani, and R. Kennedy, "Relaying protocols for wireless energy harvesting and information processing," *IEEE Trans. Wireless Commun.*, vol. 12, no. 7, pp. 3622–3636, Jul. 2013.

[8] G. Zhu, C. Zhong, H. A. Suraweera, G. K. Karagiannidis, Z. Zhang, and T. A. Tsiftsis, "Wireless information and power transfer in relay systems with multiple antennas and interference," *IEEE Trans. Commun.*, vol. 63, no. 4, pp. 1400–1418, Apr. 2015.

[9] Q. Shi, L. Liu, W. Xu, and R. Zhang, "Joint transmit beamforming and receive power splitting for MISO SWIPT systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 6, pp. 3269–3280, Jun. 2014.

[10] D. W. K. Ng, E. S. Lo, and R. Schober, "Robust beamforming for secure communication in systems with wireless information and power transfer," *IEEE Trans. Wireless Commun.*, vol. 13, no. 8, pp. 4599–4615, Aug. 2014.

[11] M. R. A. Khandaker, K. K. Wong, Y. Zhang, and Z. Zheng, "Probabilistically robust SWIPT for secrecy MISOME systems," *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 1, pp. 211–226, Jan. 2017.

[12] F. Wang, T. Peng, Y. Huang, and X. Wang, "Robust transceiver optimization for power-splitting based downlink MISO SWIPT systems," *IEEE Signal Process. Lett.*, vol. 22, no. 9, pp. 1492–1496, Sep. 2015.

[13] J. Liao, M. R. A. Khandaker, and K. K. Wong, "Robust power-splitting SWIPT beamforming for broadcast channels," *IEEE Commun. Lett.*, vol. 20, no. 1, pp. 181–184, Jan. 2016.

[14] R. Feng, Q. Li, Q. Zhang, and J. Qin, "Robust secure transmission in MISO simultaneous wireless information and power transfer system," *IEEE Trans. Veh. Technol.*, vol. 64, no. 1, pp. 400–405, Jan. 2015.

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

[16] C. Xing, S. Ma, and Y. Zhou, "Matrix-monotonic optimization for MIMO systems," *IEEE Trans. Signal Process.*, vol. 63, no.25, pp. 334–348, Jan. 2015.

[17] S. Jin, X. Liang, K. K. Wong, X. Gao, and Q. Zhu,"Ergodic rate analysis for multipair massive MIMO two-way relay networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 3, pp. 1480–1491, Nov. 2015.

[18] L. Fan, S. Jin, C. K. Wen, and H. Zhang, "Uplink achievable rate for massive MIMO systems with low-resolution ADC," *IEEE Commun. Lett.*, vol. 19, no. 12, pp. 2186–2189, Oct. 2015.

[19] S. Jin, X. Wang, Z. Li, K. K. Wong, Y. Huang, and X. Tang, "On massive MIMO zero-forcing transceiver using time-shifted pilots," *IEEE Trans. Veh. Technol.*, vol. 65, no. 1, pp. 59–74, Jan. 2016.

[20] G. Yang, C. K. Ho, R. Zhang, and Y. L. Guan, "Throughput optimization for massive MIMO systems powered by wireless energy transfer," *IEEE J. Sel Topics Commun.*, vol. 33, no. 8, pp. 1640–1650, Aug. 2015.

[21] L. Zhao, X. Wang, and K. Zheng, "Downlink hybrid information and energy transfer with massive MIMO," *IEEE Trans. Wireless Commun.*, vol. 15, no. 2, pp. 1309–1322, Feb. 2016.

[22] G. Amarasuriya, E. G. Larsson, and H. V. Poor, "Wireless information and power transfer in multiway Massive MIMO relay networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 6, pp. 3837–3855, June 2016.

[23] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. Soong, and J. C. Zhang, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.

[24] 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.

[25] 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.

[26] 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.

[27] 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.

[28] 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.

[29] 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. 2017.

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

[31] 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, Mar. 2016.

[32] C. Xing, S. Ma, and Y.-C. Wu, "Robust joint design of linear relay precoder and destination equalizer for dual-hop amplify-and-forward MIMO relay systems," *IEEE Trans. Signal Process.*, vol. 58, no. 4, pp. 2273–2283, Apr. 2010.

[33] C. Xing, S. Ma, Z. Fei, Y. C. Wu, and H. Vincent Poor, "A general robust linear transceiver design for multi-hop amplify-and-forward MIMO relaying systems," *IEEE Trans. Signal Process.*, vol. 61, no. 5, pp. 1196–1201, Mar. 2013

[34] F. Zhu, F. Gao, S. Jin, H. Lin, and M. Yao, "Robust downlink beamforming for BDMA massive MIMO system," *IEEE Trans. Commun.*, vol. 66, no. 4, pp. 1496–1507, Dec. 2017.

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

[36] S. Jin, M. R. McKay, C. Zhong, and K. K. Wong, "Ergodic capacity analysis of amplify-and-forward MIMO dual-hop systems," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2204–2224, Apr. 2010.

[37] Q. Zhang, S. Jin, K. K. Wong, H. Zhu, and M. Matthaiou, "Power scaling of uplink massive MIMO systems with arbitrary-rank channel means," *IEEE J. Sel Topics Signal Process.*, vol. 8, no. 5, pp. 966–981, May 2014.

[38] S. Boyd, L. E. Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*. Philadelphia, PA: SIAM Studies in Applied Mathematics, 1994.

[39] F. Zhu, F. Gao, T. Zhang, K. Sun, and M. Yao, "Physical-layer security for full duplex communications with self-interference mitigation," *IEEE Trans. Wireless Commun.*, vol. 15, no. 1, pp. 329–340, Jan. 2016.

[40] F. Zhu, F. Gao, M. Yao, and H. Zou, "Joint information- and jamming-beamforming for physical layer security with full duplex base station," *IEEE Trans. Signal Process.*, vol. 62, no. 24, pp. 6391–6401, Dec. 2014.

[41] 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.

[42] R. Penrose, "A generalized inverse for matrices," in *Proc. Cambridge Philos. Soc.*, 1955, vol. 51, pp. 406–413.

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

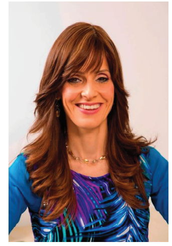
This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2019.2916405, IEEE Transactions on Wireless Communications

Fenghao 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 also 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 National University of Singapore, Singapore in 2007. He was a Research Fellow with the Institute for Infocomm Research (I2R), 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.

Prof. Gao's 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/coauthored more than 120 refereed IEEE journal papers and more than 120 IEEE conference proceeding papers.

Prof. Gao has served as an Editor of IEEE Transactions on Wireless Communications, IEEE Transactions on Cognitive Communications and Networking, IEEE Signal Processing Letters, IEEE Communications Letters, IEEE Wireless Communications Letters, International Journal on Antennas and Propagations, and China Communications. He has also serves as the symposium co-chair for 2018 IEEE Vehicular Technology Conference Spring (VTC), 2015 IEEE Conference on Communications (ICC), 2014 IEEE Global Communications Conference (GLOBECOM), 2014 IEEE Vehicular Technology Conference Fall (VTC), as well as Technical Committee Members for many other IEEE conferences.



Yonina C. Eldar (S'98,M'02,SM'07,F'12) received the B.Sc. degree in Physics in 1995 and the B.Sc. degree in Electrical Engineering in 1996 both from Tel-Aviv University (TAU), Tel-Aviv, Israel, and the Ph.D. degree in Electrical Engineering and Computer Science in 2002 from the Massachusetts Institute of Technology (MIT), Cambridge. She is currently a Professor in the Department of Mathematics and Computer Science, Weizmann Institute of Science, Rehovot, Israel. She was previously a Professor in the Department of Electrical Engineering at the Technion, where she held the Edwards Chair in Engineering. She is also a Visiting Professor at MIT, a Visiting Scientist at the Broad Institute, and an Adjunct Professor at Duke University and was a Visiting Professor at Stanford. She is a member of the Israel Academy of Sciences and Humanities (elected 2017), an IEEE Fellow and a EURASIP Fellow. Her research interests are in the broad areas of statistical signal processing, sampling theory and compressed sensing, learning and optimization methods, and their applications to biology and optics.

Dr. Eldar has received numerous awards for excellence in research and teaching, including the IEEE Signal Processing Society Technical Achievement Award (2013), the IEEE/AESS Fred Nathanson Memorial Radar Award (2014), and the IEEE Kiyo Tomiyasu Award (2016). She was a Horev Fellow of the Leaders in Science and Technology program at the Technion and an Alon Fellow. She received the Michael Bruno Memorial Award from the Rothschild Foundation, the Weizmann Prize for Exact Sciences, the Wolf Foundation Krill Prize for Excellence in Scientific Research, the Henry Taub Prize for Excellence in Research (twice), the Hershel Rich Innovation Award (three times), the Award for Women with Distinguished Contributions, the Andre and Bella Meyer Lectureship, the Career Development Chair at the Technion, the Muriel & David Jacknow Award for Excellence in Teaching, and the Technion's Award for Excellence in Teaching (twice). She received several best paper awards and best demo awards together with her research students and colleagues including the SIAM outstanding Paper Prize and the IET Circuits, Devices and Systems Premium Award, and was selected as one of the 50 most influential women in Israel. She was a member of the Young Israel Academy of Science and Humanities and the Israel Committee for Higher Education. She is the Editor in Chief of Foundations and Trends in Signal Processing, a member of the IEEE Sensor Array and Multichannel Technical Committee and serves on several other IEEE committees. In the past, she was a Signal Processing Society Distinguished Lecturer, member of the IEEE Signal Processing Theory and Methods and Bio Imaging Signal Processing technical committees, and served as an associate editor for the IEEE Transactions On Signal Processing, the EURASIP Journal of Signal Processing, the SIAM Journal on Matrix Analysis and Applications, and the SIAM Journal on Imaging Sciences. She was Co-Chair and Technical Co-Chair of several international conferences and workshops.

Gongbin Qian received the B.Eng. degree from Harbin Institute of Technology, Harbin, China in 1990, the M.Sc. degree from Harbin Institute of Technology, Harbin, China in 1993, and the Ph.D. degree from South China University of Technology, Guangzhou, China in 2009. He is currently an Associate Professor in Shenzhen University. Prof. Qian's research areas include communication theory, signal processing for communications, array signal processing.