

Achieving Global Optimality for Joint Source and Relay Beamforming Design in Two-Hop Relay Channels

Hongying Tang, Wen Chen, *Senior Member, IEEE*, Jun Li, *Member, IEEE*, and Haibin Wan

Abstract—This paper deals with joint source and relay beamforming (BF) design for an amplify-and-forward (AF) multi-antenna multirelay network. Considering that the channel state information (CSI) from relays to a destination is imperfect, we aim to maximize the worst-case received signal-to-noise ratio (SNR). The associated optimization problem is then solved in two steps. In the first step, by fixing the source BF vector, a semi-closed-form solution of the relay BF matrices is obtained, up to a power-allocation factor. In the second step, the global optimal source BF vector is obtained based on the polyblock outer approximation (PA) algorithm. We also propose two low-complexity methods for obtaining the source BF vector, which differ in their complexity and performances. The optimal joint source and relay BF solution obtained by the proposed algorithms serves as the benchmark for evaluating the existing schemes and the proposed low-complexity methods. Simulation results show that the proposed robust design can significantly reduce the sensitivity of the channel uncertainty to the system performance.

Index Terms—Amplify and forward (AF), beamforming (BF), global optimal, multiantenna multirelay system.

I. INTRODUCTION

RELAY communications can extend the coverage of a wireless network and improve the spatial diversity of cooperative systems. There are several cooperative schemes being widely used, e.g., the amplify-and-forward (AF) scheme, the decode-and-forward scheme [1], the filter-and-forward [2]–[4] scheme, etc. Among them, the AF scheme is the simplest scheme, and it has been efficiently used to exploit the benefit of relaying in the two-hop relay channels [5]–[14], the

multiple-access relay channels [15], and the two-way relay channels [2], [3], [16]–[21].

Performing transmit beamforming (BF) at the source and relays can achieve a higher data rate [16], [17]. In particular, AF-BF was considered in [7]–[21]. By maximizing the received SNR, Jing and Jafarkhani [7] gave the analytical solution of the BF design in a single-source and multiple-single-antenna-relay network. Khoshnevis *et al.* [8] considered a multiantenna-source and single-multiantenna-relay network, and gave closed-form solutions for both the source BF vector and the relay BF matrix. By relaxing the single-antenna-source and single-relay assumption, Liang and Schober [9] considered the more general case of a multiantenna-source and multiple-multiantenna-relay network and gave a closed-form solution of the relay BF matrices and a suboptimal solution for the source BF vector.

These studies are all based on the perfect channel state information (CSI) assumption. However, in a practical system, perfect CSI is usually hard to obtain, thus reducing the efficiency of BF design. Therefore, robust design taking imperfect CSI into account has attracted much attention [10]–[14], [16]–[21]. In [10] and [11], a robust distributed BF design in a wireless relay network was considered by minimizing the total relay transmit power and maximizing the received SNR, respectively. In [12], a closed-form solution was obtained for a single-antenna source–destination pair and a multiantenna relay network, and it was discovered that the robust design has a consistent form than a nonrobust design. For the more general work in [14], where the source and the destination are equipped with multiple antennas, it was proven that the robust relay optimization leads to a channel-diagonalizing structure, and a closed-form solution was proposed. Robust design in a two-way relay system were also studied in [18]–[21], on the maximization of SNR criteria, the MMSE criteria, and the minimization of transmit power criteria, respectively.

In this paper, we consider the AF relay networks with one multiantenna source, multiple multiantenna relays, and a single-antenna destination and address the joint BF design of a source and relays under imperfect CSI cases. Joint source and relay BF design has been fully investigated in the two-way relay model in both perfect and imperfect CSI cases [19], [20]. For the two-hop relay networks, however, this problem has not been well solved until now. Even in the perfect CSI case, only a suboptimal solution has been provided for the source BF vector [9]. In the robust case, the situation when the source or the

Manuscript received February 5, 2013; revised December 25, 2013 and February 17, 2014; accepted March 8, 2014. Date of publication March 12, 2014; date of current version November 6, 2014. This work was supported in part by the National Basic Research Program (973 Program) under Project 2012CB316106, by the National Natural Science Foundation of China under Grant 61161130529 and Grant 61328101, by the Science and Technology Commission of Shanghai Municipality through the Science and Technology Innovation Program under Grant 13510711200, and by the Southeast University National Key Laboratory on Mobile Communications under Grant 2013D11. The review of this paper was coordinated by Dr. N.-D. Dao.

H. Tang and W. Chen are with the Department of Electronic Engineering, Shanghai Jiaotong University, Shanghai 200240, China (e-mail: lojordan@sjtu.edu.cn; wenchen@sjtu.edu.cn).

J. Li is with the School of Electrical and Information Engineering, University of Sydney, Sydney, N.S.W. 2006, Australia (e-mail: jun.li@sydney.edu.au).

H. Wan is the College of Computer Science and Electronic Information, Guangxi University, Nanning 530004, China (e-mail: hbwan@gxu.edu.cn).

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/TVT.2014.2311472

relays are equipped with a single antenna has been discussed [11], [12], and the robust relay precoders based on the MMSE receiver and the regularized zero-forcing precoding, without taking into account the effect of the source BF vector, has been investigated [13]. Considering the fact that a practical network may involve a multiantenna source and multiple relays, it is necessary to investigate the joint source and relay BF for these general networks.

The main contributions of this paper are as follows.

- 1) Considering imperfect CSI of the second hop at relays, for a given source BF vector, we derive a semi-closed-form expression of the relay BF matrices, up to a scalar power-allocation factor. Next, we obtain the power-allocation factor through iteration between a Dinkelbach-based approach and a second-order cone programming (SOCP) problem.
- 2) To derive the optimal source BF vector, we transform the original problem into a monotonic problem, which allows us to apply the polyblock outer approximation (PA) algorithm to solve the problem. This PA-based algorithm mainly serves as a benchmark for the performance evaluation, both in the perfect CSI case and the robust case.
- 3) To further reduce the computational complexity, two low-complexity methods are proposed, which differ in their complexity and performances. Simulation results show that the proposed robust design can significantly reduce the sensitivity of the channel uncertainty to the system performance.

This paper is organized as follows. Section II introduces the system model of the multiantenna multirelay channel and gives the problem formulation. In Section III, we give the semi-closed form for the relay BF design under a fixed-source BF vector, up to a power-allocation factor, and then propose a Dinkelbach-based algorithm for determining the corresponding power-allocation factor. In Section IV, the global optimal and suboptimal source BF vectors are obtained. Finally, Section VI provides numerical examples to validate the proposed algorithms.

In this paper, $[\cdot]^*$, $[\cdot]^T$, and $[\cdot]^H$ denote the conjugate, transpose, and conjugate transpose of a matrix or a vector, respectively. \mathbb{R}^N and \mathbb{C}^N denote the N -dimensional real field and complex field, respectively. \mathbf{e}_i denotes a zero vector, except that the i th element is one. $\mathbf{0}_N$ and \mathbf{I}_N denote the N -dimensional zero vector and the identity matrix, respectively. We will use boldface lowercase letters to denote column vectors and boldface uppercase letters to denote matrices. $\|\mathbf{x}\|_2$ and $\|\mathbf{x}\|_1$ denote the Euclidean norm and the absolute sum of vector \mathbf{x} , respectively. $\text{vec}(\mathbf{X})$ stacks the columns of matrix \mathbf{X} into a vector. $|\mathbf{x}| \triangleq [|x_1|, \dots, |x_N|]^T$, and $|\mathbf{x}|^2 \triangleq [|x_1|^2, \dots, |x_N|^2]^T$. The positive semi-definite matrix \mathbf{X} is denoted $\mathbf{X} \succeq 0$. For $\mathbf{x} = [x_1, \dots, x_N]^T$, $\mathbf{y} = [y_1, \dots, y_N]^T \in \mathbb{R}^N$, and $\mathbf{x} \geq \mathbf{y}$ means $x_i \geq y_i$ for $i = 1 \dots N$. The $\text{tr}(\cdot)$ is the trace of a matrix. $\text{diag}[x_1, \dots, x_N]$ denotes a diagonal matrix with the diagonal entries x_1, \dots, x_N . \mathbf{v}^\perp and \mathbf{v}^\parallel denote the unit vectors parallel and perpendicular to \mathbf{v} , respectively. $\mathbf{v}(\mathbf{X})$ denotes the normalized principal eigenvector of \mathbf{X} .

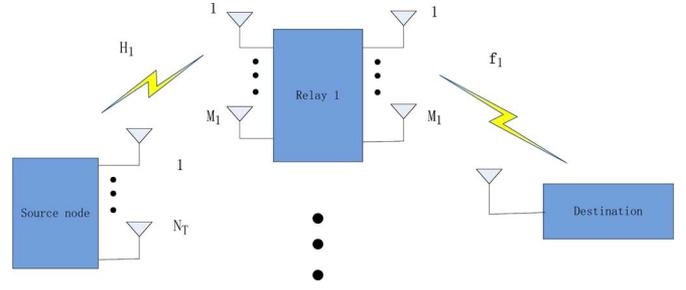


Fig. 1. Two-hop multiantenna multirelay network.

II. PROBLEM STATEMENT

A. System Model

Consider a two-hop AF multiantenna multirelay network, as shown in Fig. 1. The relays process the signals received from the source by using linear operations and then forward the processed signals to the destination. We assume that the source and relay i have N_T and M_i antennas, for $1 \leq i \leq R$, respectively, and the destination only has a single antenna. Note that the direct link between the source and the destination is not taken into account due to large-scale fading. The signal transmission is completed through two hops. In the first hop, the source transmits the N_T -dimensional vector

$$\mathbf{x} = \mathbf{g}d$$

where $\mathbf{g} \in \mathbb{C}^{N_T}$ denotes the BF vector at the source, and d is the transmitted symbol with variance $\sigma_d^2 = \mathcal{E}\{|d|^2\} = 1$. The signal received by the relay i , $1 \leq i \leq R$ is given by

$$\mathbf{q}_i = \mathbf{H}_i \mathbf{x} + \mathbf{n}_i$$

where $\mathbf{H}_i \in \mathbb{C}^{M_i \times N_T}$ denotes the first-hop channel from the source to the i th relay, and $\mathbf{n}_i \in \mathbb{C}^{M_i}$ denotes the additive white Gaussian noise (AWGN) vector with the covariance matrix $\sigma_R^2 \mathbf{I}_{M_i}$ at relay i . By the AF strategy, the signal forwarded by relay i is

$$\mathbf{s}_i = \mathbf{B}_i \mathbf{q}_i$$

where $\mathbf{B}_i \in \mathbb{C}^{M_i \times M_i}$ is the linear precoding matrix of relay i . The received signal at the destination node can be thus expressed as

$$\begin{aligned} r &= \sum_{i=1}^R \mathbf{f}_i^T \mathbf{s}_i + n_D \\ &= \sum_{i=1}^R \mathbf{f}_i^T \mathbf{B}_i \mathbf{H}_i \mathbf{g}d + \sum_{i=1}^R \mathbf{f}_i^T \mathbf{B}_i \mathbf{n}_i + n_D \end{aligned}$$

where \mathbf{f}_i denotes the channel from relay i to the destination, and n_D is the AWGN observed at the destination with variance σ_D^2 .

B. Channel Uncertainty

In a practical wireless communication scenario, perfect CSI is usually difficult to obtain. With only imperfect CSI, the system performance will be deteriorated. This motivates us to investigate the robust design, taking the CSI errors into account. As will be verified in the simulations, our proposed robust scheme will significantly reduce the sensitivity of the system to uncertain CSI.

In this paper, we assume that the uncertainty of the first-hop channel at the source is negligible, and we model the CSI in the second hop at relays to be imperfect; more specifically

$$\mathbf{f}_i = \tilde{\mathbf{f}}_i + \Delta\mathbf{f}_i \quad (1)$$

where $\tilde{\mathbf{f}}_i$ is the available CSI known at the i th relay, and $\Delta\mathbf{f}_i$ is the corresponding CSI error vector. Under the circumstance when the source (e.g., a base station) and the relays are considered fixed, and the destination is moving (e.g., a mobile terminal), the channel statistics of the two hops are different. The first hop is undergoing a slow-fading channel, whereas the second-hop channel may be fast fading due to the mobility of the destination. Then, the CSI feedback from the destination to the relays are usually outdated, and the channel uncertainty must be considered. In [22], this model was also used to exploit the situation when the relays are located closer to the source than to the destination, where this assumption is reasonable because of the high signal quality between the source and the relays.

Many existing studies [11], [18], [23] assume that the CSI error is bounded in a bundled manner, i.e., $\|\Delta\mathbf{f}\|_2 \leq \varepsilon$ for some small $\varepsilon > 0$, where $\Delta\mathbf{f} \triangleq [\Delta\mathbf{f}_1^T, \dots, \Delta\mathbf{f}_R^T]^T$. However, this model is very conservative as the channel between each relay node and the destination experiences independent distribution. In this paper, we adopt a more practical model, assuming that the CSI error vectors are estimated independently, i.e., $\|\Delta\mathbf{f}_i\|_2 \leq \varepsilon_i$, for some small $\varepsilon_i > 0$. We rewrite it as $\Delta\mathbf{f} \in \mathcal{A}$, where

$$\mathcal{A} \triangleq \left\{ \mathbf{a} \mid \mathbf{a} = [\mathbf{a}_1^T, \dots, \mathbf{a}_R^T]^T, \|\mathbf{a}_i\|_2 \leq \varepsilon_i; \mathbf{a}_i \in \mathbb{C}^{M_i} \right\}. \quad (2)$$

We also assume in this paper that the uncertainty error bound is not too large, i.e., $\varepsilon_i \leq \|\tilde{\mathbf{f}}_i\|_2$, which is reasonable since a large error bound would lead to the instability of the system, and any BF design becomes trivial. In this error model, one cannot use the S-lemma to transform the infinitely many constraints of the error vector into a linear matrix inequality [11], [18] as it will degrade into a conservative approach [25]. By contrast, we will use an alternative approach based on the idea of *real-valued implementation* proposed in [26] and prove in Section III-B that only finite realizations of the channel can act as the worst-case channel, thus making the optimization problem tractable again.

C. Problem Formulation

By maximizing the worst-case received SNR over the channel uncertainty region under individual power constraints at the relays and the source, the problem of jointly optimizing

the source BF and the relay BF can be mathematically formulated as

$$\max_{\{\mathbf{B}_i\}_{i=1}^R, \mathbf{g}} \min_{\Delta\mathbf{f} \in \mathcal{A}} \frac{\left| \sum_{i=1}^R \mathbf{f}_i^T \mathbf{B}_i \mathbf{H}_i \mathbf{g} \right|^2}{\sigma_D^2 + \sigma_R^2 \sum_{i=1}^R \|\mathbf{f}_i^T \mathbf{B}_i\|_2^2} \quad (3a)$$

$$\text{s.t. } \|\mathbf{B}_i \mathbf{H}_i \mathbf{g}\|_2^2 + \sigma_R^2 \text{tr}(\mathbf{B}_i^H \mathbf{B}_i) \leq P_i \quad \forall i \quad (3b)$$

$$\|\mathbf{g}\|_2^2 \leq P_s \quad (3c)$$

where P_s is the maximum power at the source, and P_i is the maximum power at relay i . In Section III, we first fix the source BF vector \mathbf{g} and then derive a semi-closed form of the optimal relay BF matrices up to a real-valued power-allocation factor, which can be determined by an SOCP problem. Then, in Section IV, we propose a global optimal algorithm and two suboptimal algorithms to determine \mathbf{g} .

III. OPTIMAL BEAMFORMING MATRICES AT RELAYS

By fixing the source BF vector \mathbf{g} and taking into account the CSI error model (1) and (2), problem (3) becomes

$$\max_{\{\mathbf{B}_i\}_{i=1}^R} \min_{\Delta\mathbf{f} \in \mathcal{A}} \frac{\left| \sum_{i=1}^R (\tilde{\mathbf{f}}_i + \Delta\mathbf{f}_i)^T \mathbf{B}_i \mathbf{u}_i \right|^2}{\sigma_D^2 + \sigma_R^2 \sum_{i=1}^R \left\| (\tilde{\mathbf{f}}_i + \Delta\mathbf{f}_i)^T \mathbf{B}_i \right\|_2^2} \quad (4a)$$

$$\text{s.t. } \|\mathbf{B}_i \mathbf{u}_i\|_2^2 + \sigma_R^2 \text{tr}(\mathbf{B}_i^H \mathbf{B}_i) \leq P_i \quad \forall i \quad (4b)$$

where we defined $\mathbf{u}_i \triangleq \mathbf{H}_i \mathbf{g}$ for convenience. In Section III-A, we will first introduce the related work of problem (4). By fixing the source BF vector \mathbf{g} , a semi-closed-form of \mathbf{B}_i is given in Section III-B, up to a power-allocation factor. Then, in Section III-C, the optimal power-allocation factor is determined via a Dinkelbach-based algorithm.

A. Related Work

Problem (4) has been discussed in [23], where the problem in the multipoint-to-multipoint setting was considered. By vectorizing all \mathbf{B}_i and stacking them to form a column vector as

$$\mathbf{b}_L \triangleq [\text{vec}(\mathbf{B}_1)^T, \dots, \text{vec}(\mathbf{B}_R)^T]^T \in \mathbb{C}^{\sum_{i=1}^R M_i^2}$$

and after some tedious manipulations, (4) can be transformed into a semi-definite programming (SDP) problem with variable $\mathbf{B} \triangleq \mathbf{b}_L \mathbf{b}_L^H \in \mathbb{C}^{(\sum_{i=1}^R M_i^2) \times (\sum_{i=1}^R M_i^2)}$. Obviously, this leads to prohibitively computational complexity. In addition, in some cases, the optimal \mathbf{B} obtained by the SDP solver may not be of rank one, thus leading to suboptimal \mathbf{b}_L . Furthermore, the result in [23] is numerical and can provide no insight into the structure of the optimal relay BF matrices. Therefore, it is necessary to reinvestigate problem (4).

Recently, a closed-form solution of (4), when $R = 1$, is derived in [12]. Adopting the *saddle point* theorem, it has been proven that the worst-case CSI uncertainty can be uniquely determined. Additionally, it has been shown that the robust relay BF matrix has a consistent form as that in the perfect CSI

case. However, when the multiple relay channel is considered, the analysis becomes much more difficult, and the extension of the saddle-point-based technique is no longer straightforward. In the following, we will prove that the robust relay BF matrices in (4) also have a similar form as that in the perfect CSI case, and the worst-case CSI uncertainty is one of the 2^R possible channel errors (see Theorem 1).

B. Semi-Closed form of Optimal Relay BF Matrices

We first introduce the following result given in [7] and [9] under perfect CSI assumption; based on this, we show the result of robust design.

Lemma 1 [9]: With perfect CSI assumption, i.e., $\Delta \mathbf{f}_i = \mathbf{0}_{M_i}$, the optimal relay BF matrices in (4) are given by

$$\mathbf{B}_i = c_i^\# \hat{\mathbf{f}}_i^* \hat{\mathbf{u}}_i^H \quad (5)$$

where $\hat{\mathbf{u}}_i \triangleq \mathbf{u}_i / \|\mathbf{u}_i\|_2$, and $\hat{\mathbf{f}}_i \triangleq \mathbf{f}_i / \|\mathbf{f}_i\|_2$. The real-valued power-allocation vector $\mathbf{c}^\# \triangleq [c_1^\#, \dots, c_R^\#]^T$ is determined by

$$\mathbf{c}^\# = \arg \max_{\mathbf{c}=[c_1, \dots, c_R]} \frac{\left(\sum_{i=1}^R c_i \|\mathbf{f}_i\|_2 \|\mathbf{u}_i\|_2 \right)^2}{\sigma_R^2 \sum_{i=1}^R c_i^2 \|\mathbf{f}_i\|_2^2 + \sigma_D^2} \quad (6a)$$

$$\text{s.t. } c_i \leq \sqrt{\frac{P_i}{\|\mathbf{u}_i\|_2^2 + \sigma_R^2}}, \quad 1 \leq i \leq R. \quad (6b)$$

Corollary 1 [7]: Define $\phi_i \triangleq (\|\mathbf{u}_i\|_2 \sqrt{1 + \|\mathbf{u}_i\|_2^2} / \|\mathbf{f}_i\|_2 \sqrt{P_i})$, for $i = 1, \dots, R$. Let π be a permutation of $\{1, \dots, R\}$ such that $\{\phi_{\pi(i)}\}_{i=1}^R$ are in descending order. Then, $\mathbf{c}^\#$ in (6) has the following analytical solution:

$$c_i^\# = v_i^{(j_0)} \sqrt{\frac{P_i}{\|\mathbf{u}_i\|_2^2 + \sigma_R^2}}$$

where

$$v_i^{(j)} \triangleq \begin{cases} 1, & i = \pi_1, \dots, \pi_j \\ \lambda_j \phi_i, & i = \pi_{j+1}, \dots, \pi_R \end{cases}$$

$\lambda_j \triangleq (1 + \sum_{m=1}^j a_{\pi_m}^2 / \sum_{m=1}^j b_{\pi_m})$, $a_j \triangleq (\|\mathbf{f}_j\|_2 \sqrt{P_j} / \sqrt{1 + \|\mathbf{u}_j\|_2^2})$, $b_j \triangleq (\|\mathbf{f}_j\|_2 \|\mathbf{u}_j\|_2 \sqrt{P_j} / \sqrt{1 + \|\mathbf{u}_j\|_2^2})$, and j_0 is the smallest j , such that $\lambda_j < \phi_{\pi_{j+1}}^{-1}$ for $1 \leq j \leq R$.

Define $\mathcal{B} \triangleq \{\mathbf{a} | \mathbf{a} = [a_1, \dots, a_R]^T, a_i = \|\tilde{\mathbf{f}}_i\|_2 \pm \varepsilon_i\}$, and $\mathbf{f}_\eta \triangleq [f_{\eta 1}, \dots, f_{\eta R}]^T$. Now, we present the optimal robust relay BF matrices in Theorem 1.

Theorem 1: The optimal robust relay BF matrices in (4) are given by

$$\mathbf{B}_i = c_i^\# \hat{\mathbf{f}}_i^* \hat{\mathbf{u}}_i^H \quad (7)$$

where $\hat{\mathbf{f}}_i \triangleq \tilde{\mathbf{f}}_i / \|\tilde{\mathbf{f}}_i\|_2$, and the real-valued $\mathbf{c}^\#$ is the optimal solution to the following problem:

$$\max_{\mathbf{c}} \min_{\mathbf{f}_\eta \in \mathcal{B}} \frac{\left(\sum_{i=1}^R f_{\eta i} c_i \|\mathbf{u}_i\|_2 \right)^2}{\sigma_R^2 \sum_{i=1}^R f_{\eta i}^2 c_i^2 + \sigma_D^2} \quad (8a)$$

$$\text{s.t. } c_i \leq \sqrt{\frac{P_i}{\sigma_R^2 + \|\mathbf{u}_i\|_2^2}}, \quad 1 \leq i \leq R. \quad (8b)$$

Notice that, in Section II-B, we have assumed that $\varepsilon_i \leq \|\tilde{\mathbf{f}}_i\|_2$. Thus, any vector $\mathbf{f}_\eta \in \mathcal{B}$ has nonnegative real-valued elements. From Theorem 1, one can observe that problem (8) is only optimized over the discrete set \mathcal{B} with 2^R elements. By contrast, the original problem (4) is optimized over the continuous region \mathcal{A} with infinite channel realizations. This important step significantly reduces the computational complexity and turns problem (4) in a more tractable form.

To prove Theorem 1, we first discuss the structure of the optimal \mathbf{B}_i , whose expression is given in the following lemma.

Lemma 2: The optimal \mathbf{B}_i in (4) must have the form $\mathbf{B}_i = \mathbf{b}_i \hat{\mathbf{u}}_i^H$ for some $\mathbf{b}_i \in \mathbb{C}^R$. Denote $\mathbf{f} \triangleq [\mathbf{f}_1^T, \dots, \mathbf{f}_R^T]^T = [(\tilde{\mathbf{f}}_1 + \Delta \mathbf{f}_1)^T, \dots, (\tilde{\mathbf{f}}_R + \Delta \mathbf{f}_R)^T]^T$. Then, (4) becomes

$$\max_{\mathbf{b}_i} \min_{\Delta \mathbf{f} \in \mathcal{A}} \text{SNR}(\mathbf{b}_i, \mathbf{f}) \triangleq \frac{\left| \sum_{i=1}^R (\tilde{\mathbf{f}}_i + \Delta \mathbf{f}_i)^T \mathbf{b}_i \|\mathbf{u}_i\|_2 \right|^2}{\sigma_R^2 \sum_{i=1}^R \|(\tilde{\mathbf{f}}_i + \Delta \mathbf{f}_i)^T \mathbf{b}_i\|_2^2 + \sigma_D^2} \quad (9a)$$

$$\text{s.t. } \|\mathbf{b}_i\|_2 \leq \sqrt{\frac{P_i}{\sigma_R^2 + \|\mathbf{u}_i\|_2^2}}. \quad (9b)$$

Proof: See Appendix A. \blacksquare

Then, we come to determine the optimal \mathbf{b}_i . To proceed, we first discuss a particular case $\mathbf{b}_i = c_i \hat{\mathbf{f}}_i^*$ for some $c_i \in \mathbb{C}$, as in the following lemma.

Lemma 3: If $\mathbf{b}_i = c_i \hat{\mathbf{f}}_i^*$ for some $c_i \in \mathbb{C}$, then the optimal c_i of problem (9) must be real valued and problem (9) can be transformed into

$$\max_{\mathbf{c}} \min_{\mathbf{f}_\eta \in \mathcal{B}} \frac{\left(\sum_{i=1}^R f_{\eta i} c_i \|\mathbf{u}_i\|_2 \right)^2}{\sigma_R^2 \sum_{i=1}^R f_{\eta i}^2 c_i^2 + \sigma_D^2} \quad (10a)$$

$$\text{s.t. } c_i \leq \sqrt{\frac{P_i}{\sigma_R^2 + \|\mathbf{u}_i\|_2^2}}, \quad 1 \leq i \leq R. \quad (10b)$$

Proof: See Appendix B. \blacksquare

Proof of Theorem 1: Denote the optimal solution of (10) as $\mathbf{c}^\# \triangleq [c_1^\#, \dots, c_R^\#]$ and $\mathbf{f}_\eta^\# \triangleq [f_{\eta 1}^\#, \dots, f_{\eta R}^\#]^T$. In Appendix B, we have shown that when $\mathbf{b}_i = c_i^\# \hat{\mathbf{f}}_i^*$, the corresponding *worst channel* is $\mathbf{f}^\# \triangleq [f_{\eta 1}^\# \hat{\mathbf{f}}_1^T, \dots, f_{\eta R}^\# \hat{\mathbf{f}}_R^T]^T$. When we use the term *worst channel*, we mean the channel \mathbf{f} with the minimum SNR over $\mathbf{f}_\eta \in \mathcal{B}$ under a fixed \mathbf{c} in (10). Hence, we have

$$\min_{\mathbf{f}_\eta \in \mathcal{B}} \text{SNR} \left(c_i^\# \hat{\mathbf{f}}_i^*, \mathbf{f} \right) = \text{SNR} \left(c_i^\# \hat{\mathbf{f}}_i^*, \mathbf{f}^\# \right). \quad (11)$$

Consider the received SNR in (9) with any \mathbf{b}_i under the particular channel $\mathbf{f}^\#$. We can decompose $\mathbf{b}_i \in \mathbb{C}^{M_i}$ as $\mathbf{b}_i = c_i (\hat{\mathbf{f}}_i^\parallel)^* + d_i (\hat{\mathbf{f}}_i^\perp)^*$, where $c_i, d_i \in \mathbb{C}$, and $\sqrt{|c_i|^2 + |d_i|^2} = \|\mathbf{b}_i\|_2$. Then, we have

$$\begin{aligned} \text{SNR}(\mathbf{b}_i, \mathbf{f}^\#) &= \frac{\left| \sum_{i=1}^R f_{\eta i}^\# \hat{\mathbf{f}}_i^T \left(c_i (\hat{\mathbf{f}}_i^\parallel)^* + d_i (\hat{\mathbf{f}}_i^\perp)^* \right) \|\mathbf{u}_i\|_2 \right|^2}{\sigma_R^2 \sum_{i=1}^R \left\| f_{\eta i}^\# \hat{\mathbf{f}}_i^T \left(c_i (\hat{\mathbf{f}}_i^\parallel)^* + d_i (\hat{\mathbf{f}}_i^\perp)^* \right) \right\|_2^2 + \sigma_D^2} \end{aligned}$$

$$\begin{aligned}
&= \frac{\left| \sum_{i=1}^R f_{\eta_i}^{\#} c_i \|\mathbf{u}_i\|_2 \right|^2}{\sigma_R^2 \sum_{i=1}^R f_{\eta_i}^{\#2} |c_i|^2 + \sigma_D^2} \\
&\stackrel{(a)}{\leq} \frac{\left| \sum_{i=1}^R f_{\eta_i}^{\#} c_i^{\#} \|\mathbf{u}_i\|_2 \right|^2}{\sigma_R^2 \sum_{i=1}^R f_{\eta_i}^{\#2} |c_i^{\#}|^2 + \sigma_D^2} \\
&= \text{SNR} \left(c_i^{\#} \hat{\mathbf{f}}_i^*, \mathbf{f}^{\#} \right) \quad (12)
\end{aligned}$$

where (a) is due to the fact that the optimal solution in (10) is $\mathbf{c}^{\#}$. Since $\mathbf{f}^{\#}$ is only a particular channel, there must be

$$\min_{\mathbf{f}_\eta \in \mathcal{B}} \text{SNR}(\mathbf{b}_i, \mathbf{f}) \leq \text{SNR}(\mathbf{b}_i, \mathbf{f}^{\#}). \quad (13)$$

By combining (11)–(13), we have

$$\begin{aligned}
\min_{\mathbf{f}_\eta \in \mathcal{B}} \text{SNR}(\mathbf{b}_i, \mathbf{f}) &\leq \text{SNR}(\mathbf{b}_i, \mathbf{f}^{\#}) \\
&\stackrel{(a)}{\leq} \text{SNR} \left(c_i^{\#} \hat{\mathbf{f}}_i^*, \mathbf{f}^{\#} \right) \\
&\stackrel{(b)}{=} \min_{\mathbf{f}_\eta \in \mathcal{B}} \text{SNR} \left(c_i^{\#} \hat{\mathbf{f}}_i^*, \mathbf{f} \right) \quad (14)
\end{aligned}$$

where (a) is due to (12), and (b) is due to (11). Equation (14) shows that the optimal $\mathbf{b}_i^{\#} = c_i^{\#} \hat{\mathbf{f}}_i^*$. With this discussion and by combining Lemmas 2 and 3, we get the semi-closed form of \mathbf{B}_i as in (7), up to a power-allocation factor \mathbf{c} determined by (8). ■

C. Dinkelbach-Based Algorithm for Solving the Optimal Power-Allocation factor \mathbf{c}

In Lemma 1, under perfect CSI assumption, \mathbf{c} is obtained by a closed-form solution in Corollary 1. However, in the robust case, such explicit analytical result is difficult to be derived. Here, we will present a Dinkelbach-based algorithm for solving \mathbf{c} in (8).

Introducing a slack variable γ , problem (8) can be transferred into the following equivalent problem:

$$\max_{\mathbf{c}, \gamma} \gamma \quad (15a)$$

$$\text{s.t.} \quad \frac{\left(\sum_{i=1}^R f_{\eta_i} c_i \|\mathbf{u}_i\|_2 \right)^2}{\sigma_R^2 \sum_{i=1}^R f_{\eta_i}^2 c_i^2 + \sigma_D^2} \geq \gamma, \quad \mathbf{f}_\eta \in \mathcal{B} \quad (15b)$$

$$c_i \leq \sqrt{\frac{P_i}{\sigma_R^2 + \|\mathbf{u}_i\|_2^2}} \quad (15c)$$

which can be solved by checking feasibility for a fixed γ iteratively. To find the maximum value of γ , the conventional method is to use the bisection approach [24]: In the κ th iteration, assume that the optimal value γ lies in the interval $[\gamma_l^{(\kappa)}, \gamma_u^{(\kappa)}]$. Set $\gamma = (\gamma_l^{(\kappa)} + \gamma_u^{(\kappa)})/2$, and solve (15). If this problem is found to be feasible, update the interval bounds as $\gamma_l^{(\kappa+1)} = \gamma$ and $\gamma_u^{(\kappa+1)} = \gamma_u^{(\kappa)}$; otherwise, update the interval bounds as $\gamma_l^{(\kappa+1)} = \gamma_l^{(\kappa)}$ and $\gamma_u^{(\kappa+1)} = \gamma$. This iteration is repeated until some threshold is achieved.

Since (8) is a generalized fractional programming problem, it can be alternatively solved with the Dinkelbach-based algorithm as in [27] and [28]. Unlike the bisection-based algorithm, the Dinkelbach-based algorithm does not need to shrink the interval iteratively. By contrast, it exploits the inherent property of the fractional programming problem and approaches to the optimal γ from the left side, e.g., $\gamma^{(\kappa)} \leq \gamma$. The advantage of the Dinkelbach-based algorithm lies in the fact that it has a quotient-superlinear convergence, which is obviously faster than the linear convergence of the bisection-based algorithm [28]. Basically, the Dinkelbach-based algorithm aims to solve a sequence of problem, which is shown in the following, at the κ th iteration:

$$\begin{aligned}
\max_{\mathbf{c}} \min_{\mathbf{f}_\eta \in \mathcal{B}} &\sum_{i=1}^R f_{\eta_i} c_i \|\mathbf{u}_i\|_2 - \sqrt{\gamma^{(\kappa)}} \sigma_R^2 \\
&\times \sum_{i=1}^R f_{\eta_i}^2 c_i^2 - \sqrt{\gamma^{(\kappa)}} \sigma_D^2 \quad (16a)
\end{aligned}$$

$$\text{s.t.} \quad c_i \leq \sqrt{\frac{P_i}{\sigma_R^2 + \|\mathbf{u}_i\|_2^2}}. \quad (16b)$$

By introducing a slack variable τ , problem (16) becomes

$$\max_{\mathbf{c}, \tau} \tau \quad (17a)$$

$$\begin{aligned}
\text{s.t.} \quad \min_{\mathbf{f}_\eta \in \mathcal{B}} &\sum_{i=1}^R f_{\eta_i} c_i \|\mathbf{u}_i\|_2 - \sqrt{\gamma^{(\kappa)}} \sigma_R^2 \\
&\times \sum_{i=1}^R f_{\eta_i}^2 c_i^2 - \sqrt{\gamma^{(\kappa)}} \sigma_D^2 \geq \tau \quad (17b)
\end{aligned}$$

$$c_i \leq \sqrt{\frac{P_i}{\sigma_R^2 + \|\mathbf{u}_i\|_2^2}} \quad (17c)$$

which is obviously equivalent to a SOCP problem, i.e.,

$$\min_{\mathbf{c}, \tau} -\tau \quad (18a)$$

$$\begin{aligned}
\text{s.t.} \quad &\sum_{i=1}^R f_{\eta_i} c_i \|\mathbf{u}_i\|_2 - \sqrt{\gamma^{(\kappa)}} \sigma_R^2 \\
&\times \sum_{i=1}^R f_{\eta_i}^2 c_i^2 - \sqrt{\gamma^{(\kappa)}} \sigma_D^2 \geq \tau, \quad \mathbf{f}_\eta \in \mathcal{B} \quad (18b)
\end{aligned}$$

$$c_i \leq \sqrt{\frac{P_i}{\sigma_R^2 + \|\mathbf{u}_i\|_2^2}} \quad (18c)$$

and can be solved in polynomial time by interior point method. Then, the solution $\mathbf{c}^{(\kappa)}$ from (18) is used to update $\gamma^{(\kappa+1)}$, i.e.,

$$\gamma^{(\kappa+1)} = \min_{\mathbf{f}_\eta \in \mathcal{B}} \frac{\left(\sum_{i=1}^R f_{\eta_i} c_i^{(\kappa)} \|\mathbf{u}_i\|_2 \right)^2}{\sigma_R^2 \sum_{i=1}^R f_{\eta_i}^2 c_i^{(\kappa)2} + \sigma_D^2}. \quad (19)$$

When $\tau = 0$, this iteration stops. We summarize the Dinkelbach-based algorithm in Algorithm I in Table I.

TABLE I
ALGORITHM I: DINKELBACH-BASED ALGORITHM
FOR DETERMINING THE OPTIMAL \mathbf{c} IN (8)

1	Choose δ_1 as the desired threshold. Set $\kappa = 0$ and $c_i^{(\kappa)} = \sqrt{\frac{P_i}{\sigma_R^2 + \ \mathbf{u}_i\ _2^2}}$ as the initial power allocation factor.
2	With given $\mathbf{c}^{(\kappa)}$, set $\gamma^{(\kappa+1)}$ as in (19).
3	Solve the SOCP problem in (18) to obtain $\mathbf{c}^{(\kappa+1)}$.
4	If $\min_{\mathbf{f}_\eta \in \mathcal{B}} \sum_{i=1}^R f_{\eta i} c_i^{(\kappa)} \ \mathbf{u}_i\ _2 - \sqrt{\gamma^{(\kappa)} \sigma_R^2 \sum_{i=1}^R f_{\eta i}^2 c_i^{(\kappa)2} - \sqrt{\gamma^{(\kappa)} \sigma_D^2}} \leq \delta_1$, go to Step 5. Otherwise, $\kappa = \kappa + 1$, and go to Step 2.
5	Return $\mathbf{c}^\# = \mathbf{c}^{(\kappa)}$.

Remark 1: In the perfect CSI case, c_i is obtained by Corollary 1, which can be any value between 0 and its maximal value. However, as pointed out in [7], there is at least one relay that uses its full power. The same phenomenon holds true in the robust case. This can be explained as follows. Suppose that none of the relays use its full power. Then, there exists a real-valued $\chi > 1$ defined as

$$\chi \triangleq \min_{i \in \{1, \dots, R\}} \left\{ \sqrt{\frac{P_i}{c_i^{\#2} (\|\mathbf{u}_i\|_2^2 + \sigma_R^2)}} \right\}.$$

It is easy to see that $\chi c_i^\#$ also satisfies the power constraints in (10b). However

$$\min_{\mathbf{f}_\eta \in \mathcal{B}} \frac{\chi^2 \left(\sum_{i=1}^R f_{\eta i} c_i^\# \|\mathbf{u}_i\|_2 \right)^2}{\chi^2 \sigma_R^2 \sum_{i=1}^R f_{\eta i}^2 c_i^{\#2} + \sigma_D^2} > \min_{\mathbf{f}_\eta \in \mathcal{B}} \frac{\left(\sum_{i=1}^R f_{\eta i} c_i^\# \|\mathbf{u}_i\|_2 \right)^2}{\sigma_R^2 \sum_{i=1}^R f_{\eta i}^2 c_i^{\#2} + \sigma_D^2}.$$

Then, the new coefficient $\chi c_i^\#$ leads to a higher SNR, which contradicts to the assumption that $c_i^\#$ is the optimal solution.

IV. OPTIMAL BEAMFORMING VECTOR AT THE SOURCE

By Theorem 1, the optimization variables of problem (4) have been transformed into \mathbf{c} and \mathbf{g} . According to Sections III-B and C, by fixing \mathbf{g} , which is the optimal solution of \mathbf{c} , and can be obtained from Algorithm I, i.e., $\mathbf{c}^\# = \mathbf{c}^\#(\mathbf{g})$, the remaining challenge is to determine the optimal \mathbf{g} , which is the solution of

$$\max_{\mathbf{g}} \min_{\mathbf{f}_\eta \in \mathcal{B}} \frac{\left| \sum_{i=1}^R f_{\eta i} c_i^\#(\mathbf{g}) \|\mathbf{u}_i\|_2 \right|^2}{\sigma_R^2 \sum_{i=1}^R f_{\eta i}^2 c_i^{\#2}(\mathbf{g}) + \sigma_D^2} \quad (20a)$$

$$\text{s.t.} \quad \|\mathbf{g}\|_2^2 \leq P_s. \quad (20b)$$

Due to the nonconvex nature of (20), it seems impossible to derive the optimal solution. Even in the perfect CSI case, Liang and Schober [9] only propose a suboptimal algorithm based on the gradient method. However, by exploiting the hidden monotonic property of problem (20), we propose an efficient algorithm based on the PA algorithm to determine the global optimal \mathbf{g} . We also find that the global optimal \mathbf{g} is parallel to the principal eigenvector of $\sum_{i=1}^R \mu_i \mathbf{H}_i^H \mathbf{H}_i$, for some $\sum_{i=1}^R \mu_i \leq 1$, $\mu_i \geq 0$, in Section IV-A. Our result covers the special case discussed in [8] that $\mathbf{g} = \sqrt{P_s} \mathbf{v}(\mathbf{H}_1^H \mathbf{H}_1)$ when $R = 1$.

A. Monotonic Optimization

Let \mathbb{R}_+^N be the N -dimensional nonnegative real set. A set $\mathcal{H} \subset \mathbb{R}_+^N$ is called normal if, for any point $\mathbf{x} \in \mathcal{H}$, any point \mathbf{x}' with $\mathbf{0} \leq \mathbf{x}' \leq \mathbf{x}$ must satisfy $\mathbf{x}' \in \mathcal{H}$. An optimization problem is the *monotonic optimization* problem if it can be expressed as

$$\max_{\mathbf{x}} \Phi(\mathbf{x}) \quad \text{s.t.} \quad \mathbf{x} \in \mathcal{H}$$

where $\mathcal{H} \subset \mathbb{R}_+^N$ is a nonempty normal closed set, and the function $\Phi(\mathbf{x})$ is an increasing function with respect to $\mathbf{x} \in \mathcal{H}$.

To exploit the monotonic property of problem (20), we define

$$\mathbf{w} \triangleq [w_1, \dots, w_R]^T \triangleq [\|\mathbf{H}_1 \mathbf{g}\|_2^2, \dots, \|\mathbf{H}_R \mathbf{g}\|_2^2]^T.$$

Then, the worst-case SNR becomes a function of the new variable \mathbf{w} , i.e.,

$$\text{SNR}(\mathbf{w}) \triangleq \min_{\mathbf{f}_\eta \in \mathcal{B}} \frac{\left| \sum_{i=1}^R f_{\eta i} c_i^\#(\mathbf{w}) \sqrt{w_i} \right|^2}{\sigma_R^2 \sum_{i=1}^R f_{\eta i}^2 c_i^{\#2}(\mathbf{w}) + \sigma_D^2} \quad (21)$$

where $c_i^\#(\mathbf{w})$ is the optimal solution of the following problem for given \mathbf{w} :

$$\max_{\mathbf{c}} \min_{\mathbf{f}_\eta \in \mathcal{B}} \frac{\left(\sum_{i=1}^R f_{\eta i} c_i \sqrt{w_i} \right)^2}{\sigma_R^2 \sum_{i=1}^R f_{\eta i}^2 c_i^2 + \sigma_D^2} \quad (22a)$$

$$\text{s.t.} \quad c_i \leq \sqrt{\frac{P_i}{\sigma_R^2 + w_i}}. \quad (22b)$$

Denote

$$\mathcal{U} \triangleq \left\{ \mathbf{w} \mid \mathbf{w} = [\text{tr}(\mathbf{H}_1^T \mathbf{H}_1 \mathbf{G}), \dots, \text{tr}(\mathbf{H}_R^T \mathbf{H}_R \mathbf{G})]^T, \mathbf{G} \succeq \mathbf{0}, \text{tr}(\mathbf{G}) \leq P_s \right\}.$$

Then, we have the following proposition.

Proposition 1: Problem (20) is equivalent to the following *monotonic optimization* problem:

$$\max_{\mathbf{w}} \text{SNR}(\mathbf{w}) \quad \text{s.t.} \quad \mathbf{w} \in \mathcal{U} \quad (23)$$

where the optimal \mathbf{w} of (23) must be on the Pareto boundary¹ of \mathcal{U} , and the associated \mathbf{G} must be of rank one.

Proof: See Appendix C. ■

Suppose that $\mathbf{G}^\#$ is associated with the optimal \mathbf{w} of (23). According to Proposition 1, $\mathbf{G}^\#$ must be of rank one. By eigenvalue decomposition $\mathbf{G}^\# = \mathbf{g}^\# \mathbf{g}^{\#H}$, we can obtain $\mathbf{g}^\#$. Specifically, the structure of the global optimum $\mathbf{g}^\#$ can be derived in Corollary 2 by following a similar argument as that in [32].

Corollary 2: The global optimal \mathbf{g} has the following structure:

$$\mathbf{g} = \sqrt{P_s} \mathbf{v} \left(\sum_{i=1}^R \mu_i \mathbf{H}_i^H \mathbf{H}_i \right), \quad \boldsymbol{\mu} \triangleq [\mu_1, \dots, \mu_R] \in \mathcal{V}$$

¹ \mathbf{x} is called the Pareto boundary (or Pareto optimal) of a region \mathcal{H} if there is no other vector $\mathbf{x}' \in \mathcal{H}$ such that $\mathbf{x}' > \mathbf{x}$.

where

$$\mathcal{V} \triangleq \left\{ \boldsymbol{\mu} \mid \sum_{i=1}^R \mu_i = 1, \mu_i \geq 0 \right\}.$$

Remark 2: According to Corollary 2, by implementing the grid search in \mathcal{V} , one can asymptotically achieve the optimal SNR if the grid is sufficiently fine. By setting the search step as 0.01, one has to compare 100 points for $R = 2$. When $R = 3, 4, 5$, this number rises to 5000, 250 000, and 12 500 000, respectively. It can be seen that the complexity of this grid search increases with R rapidly. Hence, in the following, we will propose an efficient PA-based algorithm for solving the optimal \mathbf{w} by taking advantage of the monotonic property of (23).

Remark 3: It is worth pointing out that, for some special cases, the optimal source BF vector \mathbf{g} has following expressions.

- Case 1: If $N_T = 1$, then $g = \sqrt{P_s}$.
- Case 2: If $R = 1$, then $\mathbf{g} = \sqrt{P_s} \mathbf{v}(\mathbf{H}_1^H \mathbf{H}_1)$.
- Case 3: If $M_1 = M_2 = 1$, then $\mathbf{g} = \sqrt{P_s} \sin \theta (\Pi_{\mathbf{h}_2} \mathbf{h}_1 / \|\Pi_{\mathbf{h}_2} \mathbf{h}_1\|_2) + \sqrt{P_s} \cos \theta (\Pi_{\mathbf{h}_2}^\perp \mathbf{h}_1 / \|\Pi_{\mathbf{h}_2}^\perp \mathbf{h}_1\|_2)$, where $\theta \in [0, (\pi/2)]$ and can be obtained by 1-D search, $\Pi_{\mathbf{x}} \triangleq \mathbf{x}(\mathbf{x}^H \mathbf{x})^{-1} \mathbf{x}^H$ is the orthogonal projection onto the column space of \mathbf{x} , and $\Pi_{\mathbf{x}}^\perp \triangleq \mathbf{I} - \Pi_{\mathbf{x}}$ is the orthogonal projection onto the orthogonal complement of the column space of \mathbf{x} .

B. Polyblock Outer Approximation Algorithm

In the literature, two general algorithms are widely used for solving monotonic problems: the PA algorithm from [29] and the branch–reduce–bound (BRB) algorithm from [30] and [31]. Here, we will briefly introduce the PA algorithm and then propose a PA-based algorithm to solve the optimal \mathbf{w} in (23), which automatically results in the solution of global optimal source BF vector \mathbf{g} . Performance comparison between the PA and BRB algorithms will be given in our simulation part (see [29] and [31] for more details on the PA algorithm).

A set \mathcal{P} is called a polyblock if it is the union of a finite number of boxes.² The main idea of PA is to approximate \mathcal{U} by constructing a sequence of polyblocks $\mathcal{P}^{(\kappa)}$ with increasing accuracy. At each iteration, a refined outer approximation $\mathcal{P}^{(\kappa)}$ of \mathcal{U} is generated, such that $\mathcal{P}^{(1)} \supset \mathcal{P}^{(2)} \supset \dots \supset \mathcal{U}$. Let $\mathcal{Z}^{(\kappa)}$ denote the set containing all the vertices of the polyblock $\mathcal{P}^{(\kappa)}$. Since the optimal \mathbf{w} must be on the Pareto boundary of \mathcal{U} , we will try to find that point in a shrinking search region. The vertex that achieves the maximum SNR in $\mathcal{Z}^{(\kappa)}$ is defined by $\tilde{\mathbf{z}}^{(\kappa)}$, i.e., $\tilde{\mathbf{z}}^{(\kappa)} = \arg \max_{\mathbf{z} \in \mathcal{Z}^{(\kappa)}} \text{SNR}(\mathbf{z})$, which is chosen for determining the next Pareto boundary point on \mathcal{U} . Define $\lambda \tilde{\mathbf{z}}^{(\kappa)}$ as the line that connects the points $\mathbf{0}$ and $\tilde{\mathbf{z}}^{(\kappa)} \triangleq [\tilde{z}_1^{(\kappa)}, \dots, \tilde{z}_{N_T}^{(\kappa)}]^T$. Then, the next feasible point $\mathbf{w}^{(\kappa)} \triangleq [w_1^{(\kappa)}, \dots, w_{N_T}^{(\kappa)}]^T$ is computed as the intersection point on the

Pareto boundary of \mathcal{U} with the line $\lambda \tilde{\mathbf{z}}^{(\kappa)}$. The following method is used to generate N_T new vertices adjacent to $\tilde{\mathbf{z}}^{(\kappa)}$:

$$\mathbf{z}^{(\kappa), i} = \tilde{\mathbf{z}}^{(\kappa)} - \left(\tilde{z}_i^{(\kappa)} - w_i^{(\kappa)} \right) \mathbf{e}_i, \quad i = 1, \dots, N_T \quad (24)$$

where $\mathbf{z}^{(\kappa), i}$ denotes the i th new vertex generated at the κ th iteration. Then, the new vertex set can be expressed as

$$\mathcal{Z}^{(\kappa+1)} = \left(\mathcal{Z}^{(\kappa)} \setminus \tilde{\mathbf{z}}^{(\kappa)} \right) \cup \left\{ \mathbf{z}^{(\kappa), 1}, \dots, \mathbf{z}^{(\kappa), N_T} \right\}. \quad (25)$$

Each vertex $\mathbf{z} \in \mathcal{Z}^{(\kappa+1)}$ defines a box; thus, the new polyblock $\mathcal{P}^{(\kappa+1)}$ is the union of all these boxes. The upper and lower bounds are refined as follows. The current upper bound is $f_{\max}^{(\kappa+1)} = \max_{\mathbf{z} \in \mathcal{Z}^{(\kappa+1)}} \text{SNR}(\mathbf{z})$, and the current lower bound is the maximum SNR among all the feasible points found so far: $f_{\min}^{(\kappa+1)} = \max_{\kappa} \text{SNR}(\mathbf{w}^{(\kappa)})$. The algorithm terminates when the gap between $f_{\min}^{(\kappa+1)}$ and $f_{\max}^{(\kappa+1)}$ meets some threshold. The optimal \mathbf{w} is the feasible point $\mathbf{w}^{(\kappa)}$ that achieves $f_{\min}^{(\kappa+1)}$.

Now, the only remaining problem is how to determine the intersection point $\mathbf{w}^{(\kappa)}$, which will be addressed in the following.

C. Finding Intersection Points by the Rate Profile Approach

Here, we show how to determine the intersection point $\mathbf{w}^{(\kappa)}$ on the Pareto boundary of \mathcal{U} with the line $\lambda \tilde{\mathbf{z}}^{(\kappa)}$ to apply PA Algorithm. To proceed, we first introduce the following lemma, which is important for obtaining $\mathbf{w}^{(\kappa)}$.

Lemma 4: For any \mathbf{w} on the Pareto boundary of \mathcal{U} , the corresponding \mathbf{G} satisfies $\text{tr}(\mathbf{G}) = P_s$.

Proof: Suppose that

$$\mathbf{w}^\sharp \triangleq [\text{tr}(\mathbf{H}_1^H \mathbf{H}_1 \mathbf{G}^\sharp), \dots, \text{tr}(\mathbf{H}_R^H \mathbf{H}_R \mathbf{G}^\sharp)]^T$$

is on the Pareto boundary of \mathcal{U} . If $\text{tr}(\mathbf{G}^\sharp) < P_s$, we can scale \mathbf{G}^\sharp to \mathbf{G}' such that $\mathbf{G}' = \beta \mathbf{G}^\sharp$ for some $\beta > 1$, and $\text{tr}(\mathbf{G}^\sharp) < \text{tr}(\mathbf{G}') \leq P_s$. Then, $\mathbf{w}' \triangleq [\text{tr}(\mathbf{H}_1^H \mathbf{H}_1 \mathbf{G}'), \dots, \text{tr}(\mathbf{H}_R^H \mathbf{H}_R \mathbf{G}')]^T > \mathbf{w}^\sharp$, which contradicts to the assumption that \mathbf{w}^\sharp is on the Pareto boundary of \mathcal{U} . Therefore, we have $\text{tr}(\mathbf{G}^\sharp) = P_s$. ■

Lemma 4 states that any Pareto boundary point $\mathbf{w} \in \mathcal{U}$ must have its corresponding \mathbf{G} satisfying $\text{tr}(\mathbf{G}) = P_s$. As shown in Fig. 2, any point $\mathbf{w} \in \mathcal{U}$ corresponds to a profile vector $\boldsymbol{\omega} \triangleq [\omega_1, \dots, \omega_R] = \mathbf{w} / \|\mathbf{w}\|_1$ or, equivalently, the slope of the line $\lambda \tilde{\mathbf{z}}^{(\kappa)}$. Consequently, the intersection point $\mathbf{w}^{(\kappa)}$ can be expressed as $\boldsymbol{\omega} Q^\sharp$, where Q^\sharp is the optimal value of the following problem:

$$\max_{\mathbf{G}, Q} Q \quad (26a)$$

$$\text{s.t.} \quad \text{tr}(\mathbf{H}_i^H \mathbf{H}_i \mathbf{G}) = \omega_i Q, \quad i = 1, \dots, R \quad (26b)$$

$$\text{tr}(\mathbf{G}) = P_s \quad (26c)$$

$$\mathbf{G} \succeq \mathbf{0}. \quad (26d)$$

The above approach to find $\mathbf{w}^{(\kappa)}$ is known as *rate profile* [29].

Equation (26) is an SDP problem and can be efficiently solved using the MATLAB tool package such as CVX [33].

²For given $\mathbf{b} \in \mathbb{R}_+^{N_T}$, the set of all \mathbf{x} such that $\mathbf{0} \leq \mathbf{x} \leq \mathbf{b}$ is called a box with vertex \mathbf{b} .

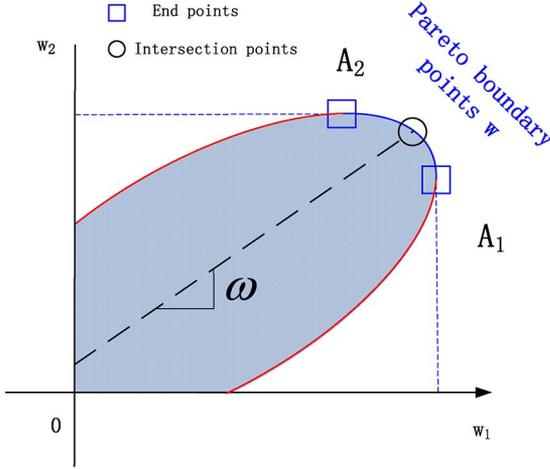


Fig. 2. Example of \mathcal{U} when $R = 2$. The Pareto boundary is only a part of the boundary of \mathcal{U} . Two end points A_i are determined by $\mathbf{G}_{A_i} \triangleq \arg \max_{\text{tr}(\mathbf{G})=P_s} \text{tr}(\mathbf{H}_i^H \mathbf{H}_i \mathbf{G}) = P_s \mathbf{v}(\mathbf{H}_i^H \mathbf{H}_i) [\mathbf{v}(\mathbf{H}_i^H \mathbf{H}_i)]^H$, for $i = 1, 2$. Then, point $A_i = (P_s \|\mathbf{H}_1 \mathbf{v}(\mathbf{H}_i^H \mathbf{H}_i)\|_2^2, P_s \|\mathbf{H}_2 \mathbf{v}(\mathbf{H}_i^H \mathbf{H}_i)\|_2^2)$. The dashed line uniquely determines the ratio ω between each element of the intersection point \mathbf{w} and its 1-norm $\|\mathbf{w}\|_1$.

TABLE II
ALGORITHM II: PA-BASED ALGORITHM FOR
DETERMINING THE OPTIMAL \mathbf{g}

1	Set $\kappa = 0$ and δ_2 as the given threshold. Initialize $\mathcal{Z}^{(\kappa)} = \{\mathbf{b}_0\}$, $\tilde{\mathbf{z}}^{(\kappa)} = \mathbf{b}_0$, $\mathcal{V}_z^{(\kappa)} = \{y y = \text{SNR}(\mathbf{z}), \mathbf{z} \in \mathcal{Z}^{(\kappa)}\}$, $f_{\min}^{(\kappa)} = 0$, and $f_{\max}^{(\kappa)} = \max_{y \in \mathcal{V}_z^{(\kappa)}} y$, where $\mathbf{b}_0 \triangleq [P_s \lambda_{\max}(\mathbf{H}_1^H \mathbf{H}_1), \dots, P_s \lambda_{\max}(\mathbf{H}_R^H \mathbf{H}_R)]^T$. The initial $\mathbf{g}^{(\kappa)}$ is the nonrobust beamforming vector in [9].
2	Compute the intersection point $\mathbf{w}^{(\kappa)}$ on the Pareto boundary of \mathcal{U} with the line $\lambda \tilde{\mathbf{z}}^{(\kappa)}$ and obtain the corresponding $\mathbf{g}^{(\kappa)}$.
3	Compute N_T new vertices that are adjacent to $\mathbf{w}^{(\kappa)}$ by (24) and update $\mathcal{Z}^{(\kappa+1)}$ by (25). Let $\mathcal{V}^{(\kappa+1)} = \{\mathcal{V}^{(\kappa)} \setminus \text{SNR}(\tilde{\mathbf{z}}^{(\kappa)})\} \cup \{\text{SNR}(\tilde{\mathbf{z}}^{(\kappa), i})\}$, $i = 1, \dots, N_T$.
4	Update the lower bound and upper bound $f_{\min}^{(\kappa+1)} = \max_{\kappa} \text{SNR}(\mathbf{w}^{(\kappa)})$, $f_{\max}^{(\kappa+1)} = \max_{y \in \mathcal{V}^{(\kappa+1)}} y$. Let $\kappa_0 \triangleq \arg \max_{\kappa} \text{SNR}(\mathbf{w}^{(\kappa)})$ and $\tilde{\mathbf{z}}^{(\kappa+1)}$ be the associate $\mathbf{z} \in \mathcal{Z}^{(\kappa+1)}$ that achieves $f_{\max}^{(\kappa+1)}$.
5	If $f_{\max}^{(\kappa+1)} - f_{\min}^{(\kappa+1)} \leq \delta_2$, go to Step 6. Otherwise, let $\kappa = \kappa + 1$, and go to Step 2.
6	Return $\mathbf{g}^\sharp = \mathbf{g}^{(\kappa_0)}$.

Denote the optimal solution as $\mathbf{G}^{(\kappa)}$. According to Proposition 1, $\mathbf{G}^{(\kappa)}$ must be of rank one. Then, the intersection point $\mathbf{w}^{(\kappa)} = [\text{tr}(\mathbf{H}_1^H \mathbf{H}_1 \mathbf{G}^{(\kappa)}), \dots, \text{tr}(\mathbf{H}_R^H \mathbf{H}_R \mathbf{G}^{(\kappa)})]^T$, and the corresponding $\mathbf{g}^{(\kappa)}$ is obtained by eigenvalue decomposition of $\mathbf{G}^{(\kappa)}$ as $\mathbf{G}^{(\kappa)} = \mathbf{g}^{(\kappa)} \mathbf{g}^{(\kappa)H}$.

Remark 4: It should be mentioned that we adopted a different approach compared with [29] for determining the feasible point $\mathbf{w}^{(\kappa)}$. In [29], the solution involves iterations between a bisection algorithm and an SDP problem. This paper, however, presents direct approach for obtaining $\mathbf{w}^{(\kappa)}$, hence bypassing any bisection approach.

D. Overall Algorithm for Determining Global Optimal \mathbf{g}

The PA-based algorithm for solving (20) is summarized as Algorithm II in Table II.

E. Low-Complexity Suboptimal Methods for Determining \mathbf{g}

The optimal solution obtained from Algorithm II is of high complexity. In practice, it can be observed that computing the global optimal solution is practically feasible for a small number of relays. Thus, we treat Algorithm II mainly as a benchmark for performance evaluation. For practical implementation, here, we propose two low-complexity suboptimal methods for determining the source BF vector \mathbf{g} , which provides a tradeoff between the computational complexity and the system performance.

1) *Robust Gradient Method:* The first method applies the gradient method in [9, Tab. I] with $\text{grad}_{\bar{\mathbf{g}}}$ determined by the following gradient estimate:

$$\text{grad}_{\bar{\mathbf{g}}} = \frac{1}{2\delta} [(\text{SNR}(\bar{\mathbf{g}} + \delta \mathbf{e}_1) - \text{SNR}(\bar{\mathbf{g}} - \delta \mathbf{e}_1)), \dots, (\text{SNR}(\bar{\mathbf{g}} + \delta \mathbf{e}_{2N_T}) - \text{SNR}(\bar{\mathbf{g}} - \delta \mathbf{e}_{2N_T}))]^T \quad (27)$$

where $\bar{\mathbf{g}} \triangleq [\text{Re}\{\mathbf{g}\}^T, \text{Im}\{\mathbf{g}\}^T]^T$, δ is a small positive constant, and the SNR in (27) is expressed as a function of $\bar{\mathbf{g}}$. Note that our gradient estimate in (27) is different from that in [9], where they compute it in an analytical form for each $\bar{\mathbf{g}}$. By comparison, for evaluating the gradient estimate in (27), it has to apply Algorithm I for all $4N_T$ vectors $\bar{\mathbf{g}}_k \pm \delta \mathbf{e}_i$, $1 \leq i \leq 2N_T$. As will be shown in Section VI, this method preserves the optimality to some extent.

2) *Simplified Robust Method:* In this method, We choose \mathbf{g} as the nonrobust solution in [9] and the power-allocation factor \mathbf{c} as the solution of (22) for given \mathbf{g} . Since this method utilizes Algorithm I only once, it is far less complex than the *robust gradient method*. However, as verified in Section VI, it shows a near-optimal performance.

V. IMPLEMENTATION ISSUES AND COMPLEXITY

Here, we discuss implementation issues and computational complexity for the proposed algorithms. For computing the source BF vector \mathbf{g} , the source needs the CSI \mathbf{H}_i of the first hops and the available channel magnitudes $\|\tilde{\mathbf{f}}_i\|_2$ of the second hops, which can be fed back by each relay. After computing \mathbf{g} and the real-valued optimal power-allocation factor \mathbf{c}^\sharp at the source, they will be broadcast to each relay node. For determining the relay BF matrices, each relay node only requires the local CSI and the \mathbf{g} and \mathbf{c}^\sharp from the source.

In Algorithm I, one only needs to determine a vector with R real variables c_i rather than R matrices $\mathbf{B}_i \in \mathbb{C}^{M_i \times M_i}$ in the conventional method [23]. According to [36], the design complexity of solving the SOCP problem (18) can be approximated as $\mathcal{O}((2^R)^{(3/2)} R^3 \log(1/\theta))$, given solution accuracy $\theta > 0$. Hence, the complexity of Algorithm I is $\mathcal{O}((2^R)^{(3/2)} R^3 \log(1/\theta))$ times the number of iterations. By contrast, using the method in [23], the complexity of the SDP solver is $\mathcal{O}((N^2(N^2 + 1)/2)^3 \log(1/\theta))$ with $N = \sum_{i=1}^R M_i^2$ times the number of iterations, which is fairly high. As shown earlier, our complexity of the SOCP problem in each iteration is much lower than that in [23]. In Section VI, we will further show that the iteration number by the Dinkelbach-based algorithm is less than the bisection-based algorithm in [23].

The major computing step of Algorithm II in iteration κ is solving problem (26) for determining $\mathbf{w}^{(\kappa)}$ and computing $N_T + 1$ worst-case SNR(z), including the intersection point $\mathbf{w}^{(\kappa)}$ and N_T new vertices. According to [34], the complexity of solving SDP problem (26) can be approximated as $\mathcal{O}(\max(N_T, R + 1)^4 \sqrt{N_T} \log(1/\theta))$. Notice that, in the perfect CSI case, SNR(z) is directly obtained by Lemma 1 and Corollary 1; in the robust case, SNR(z) is obtained by Algorithm I in Section III-C. Section VI shows the average iteration time of Algorithms I and II.

VI. SIMULATIONS AND DISCUSSION

Here, we provide numerical results to validate the proposed algorithms in this paper, using the numerical convex optimization solver CVX [33]. First, the convergence of Algorithms I and II is illustrated, compared with the bisection approach and the BRB algorithm, respectively. Then, the performance evaluation of our robust design is addressed.

The channel fading is modeled as Rayleigh fading, and each channel entry satisfies the complex normal distribution $\mathcal{CN}(0, 1)$. The noise at each node is assumed zero-mean unit variance complex Gaussian random variables. We set the power consumed at the source as 10 dB. In our simulations, we set ε_i as $\varepsilon_i^2 = \rho \|\hat{\mathbf{f}}_i\|_2^2$ with $\rho \in [0, 1)$. The larger ρ is, the poorer CSI quality will be. We also set the convergence thresholds of Algorithms I and II, respectively, as $\delta_1 = 0.01$, and $\delta_2 = 0.1$. All results are averaged over 100 channel realizations.

The following benchmarks are compared through simulations here. *Perfect optimal method* is obtained by our proposed method in Section IV under perfect CSI assumption. *Perfect gradient method* is obtained by the gradient method in [9] under perfect CSI assumption. *The robust optimal method* is the robust optimal design method proposed in Algorithm II. *The robust gradient method and simplified robust method* are proposed in Section IV-E. *The nonrobust method* was proposed in [9] using imperfect CSI.

A. Convergence Evaluation

First, we study the convergence performance of Algorithm I. Fig. 3 shows the average iteration time of Algorithm I and the bisection approach to achieve the predefined accuracy δ_1 for $R \in \{2, 4, 6\}$. The initial upper bound $\gamma_u^{(0)}$ and lower bound $\gamma_l^{(0)}$ of the bisection approach are specified as the worst-case received SNR of the *perfect optimal method* and that of the *non-robust method*, respectively. It can be observed that Algorithm I takes less than half iteration numbers of the bisection approach for most of the SNR regime. Thus, Algorithm I is more efficient.

Then, we evaluate the convergence behavior of Algorithm II and the BRB-based algorithm in [30]. We set $(N_T, M_1, M_2) = (2, 2, 2)$ and fix the relay power as 30 dB. Fig. 4 shows the average iteration numbers to achieve the certain accuracy levels both in the perfect case and in the robust case, where we set $\rho = 0.3$. The accuracy of the lower and upper bounds are defined as $(f_{\min} - f_{\text{opt}})/f_{\text{opt}}$ and $(f_{\max} - f_{\text{opt}})/f_{\text{opt}}$, respectively, where f_{opt} is the optimal value of the worst-case

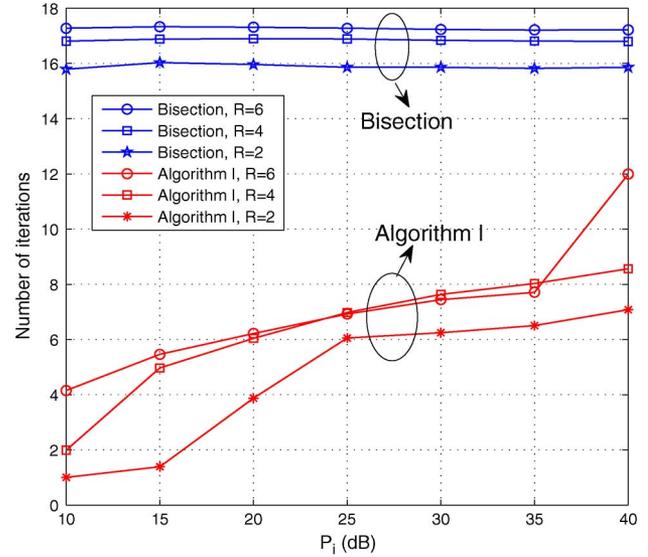


Fig. 3. Average iteration time comparison for Algorithm I and the bisection approach.

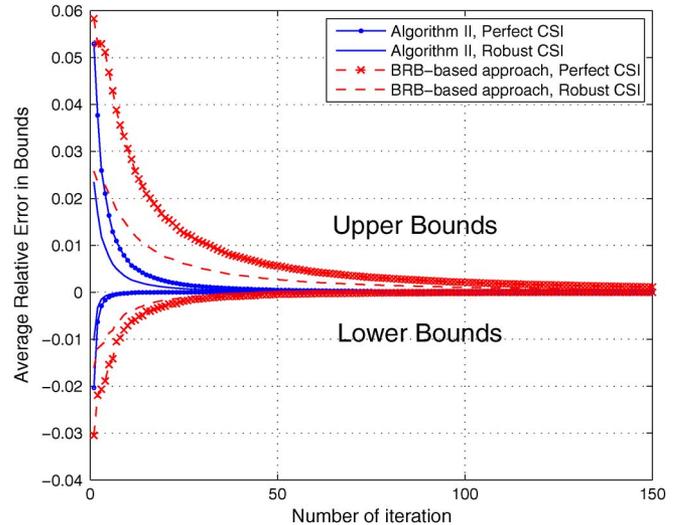


Fig. 4. Relative error of lower and upper bounds on the SNR value versus the number of iteration.

SNR. It can be seen that both algorithms quickly achieve the optimal solutions, but more iterations of the BRB algorithm are needed to achieve certain accuracy. Thus, we claim that, in our problem, Algorithm II is more efficient than the BRB-based algorithm. Notice that the convergence performance of the BRB and PA algorithms is also illustrated in [30] and [31], showing that different algorithms are superior in different scenarios.

Another observation in Fig. 4 is that ,in the robust case, both Algorithm II and BRB algorithm converges more quickly than that in the perfect case. This phenomenon is further shown in Fig. 5, which compares the lower and upper bounds of the proposed PA algorithm under different ρ assumption. It is shown that larger ρ leads to a smaller gap between the upper bound and the lower bound in each iteration. This can be explained as the maximum value over the vertices of the polyblock $\mathcal{P}^{(\kappa)}$ is lower for larger ρ .

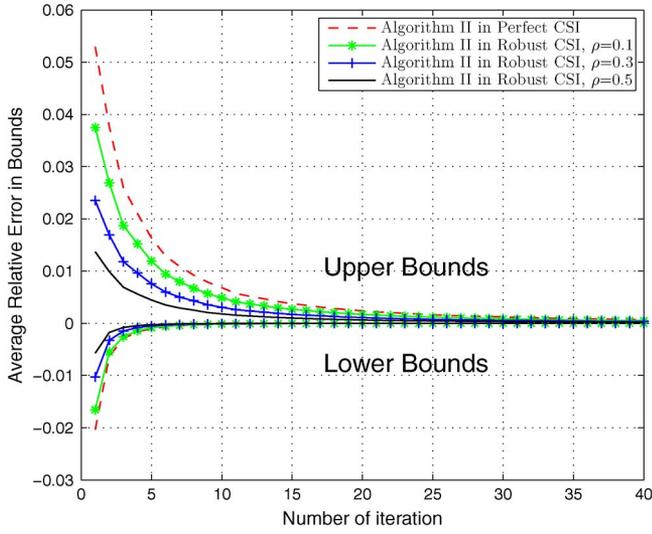


Fig. 5. Relative error of lower and upper bounds on the SNR value versus the number of iterations for different ρ .

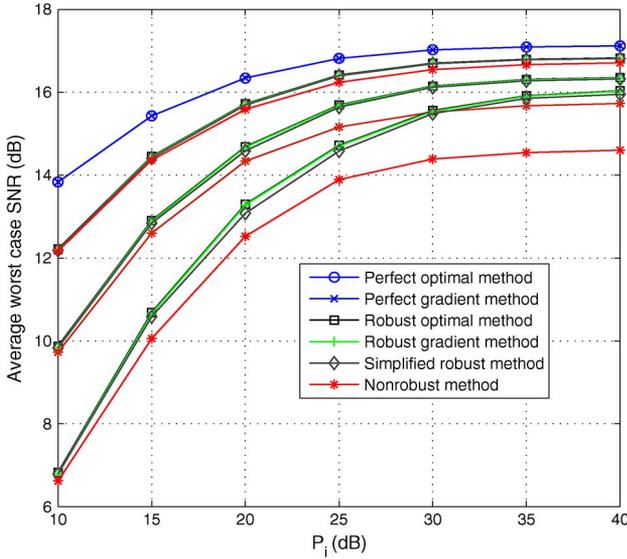


Fig. 6. Average worst-case SNR versus different relay power levels in different error bound cases.

B. Performance Comparison With the Existing Schemes

We now compare our robust BF design with some existing schemes. The parameters are set as $(N_T, M_1, M_2) = (2, 2, 2)$. Fig. 6 shows the average worst-case received SNR versus individual relay power. Simulations reveal that the *nonrobust method* will cause increasing performance loss with the increment of channel uncertainty, compared with the perfect CSI case. Even when the relay power is very large, this loss cannot be compensated. It can be observed that, when the relay power is 40 dB and the channel uncertainty ratio $\rho = 0.5$, this performance degradation is about 2.5 dB. On the other hand, the robust design can improve the performance for any channel uncertainty ratio. Although the gradient method only achieves local optimality in theory, it behaves well in our simulations and has a close-to-optimal performance in both the perfect and the robust cases. It is also shown in Fig. 6 that, as a simple yet efficient method, the *simplified robust method* has a

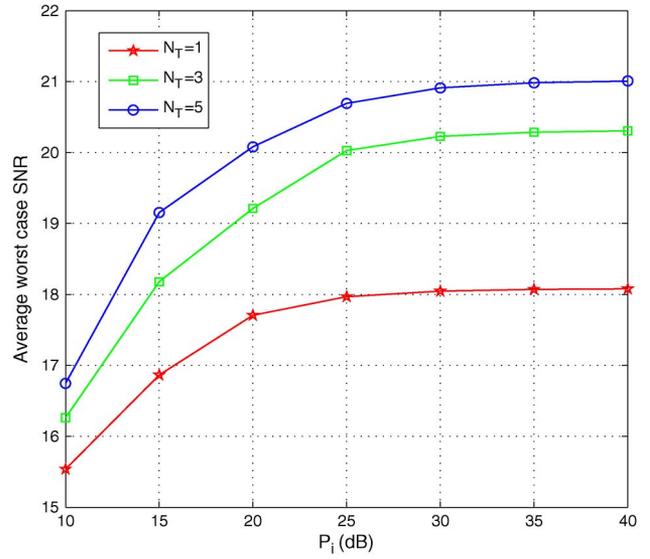


Fig. 7. Average worst-case SNR versus different relay power levels with different numbers of transmit antennas.

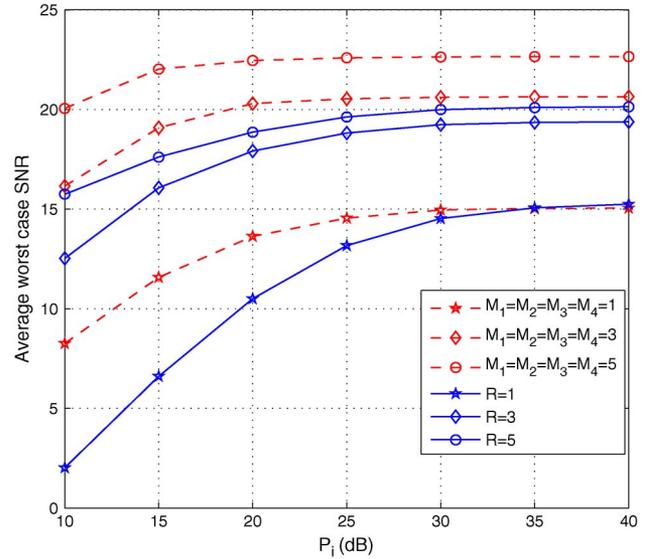


Fig. 8. Average worst-case SNR versus different relay power levels with different relay numbers and different relay antennas.

near optimal performance, which greatly facilitates the practical application of the robust design.

C. Performance Evaluation With Different Network Configurations

We investigate the impact of different network configurations. We set $\rho = 0.3$, $R = 3$, and $M_i = 3$ for $i = 1, 2, 3$. One can see that increasing the source antennas N_T from 1 to 3 brings the most benefit to the system, and the SNR improves to 2.3 dB when the relay power is 40 dB. However, the improvement is not so apparent if the source antenna number further increases to 5, where the SNR only improves to 0.7 dB. The results in Fig. 7 indicate that small increment of the source antenna number can greatly improve the system performance.

Fig. 8 compares the average worst-case SNR versus the individual relay power by the proposed robust design in [11]

and [12]. Here, we assume $N_T = 1$, since the method in [11] and [12] cannot be applied to the general case $N_T \geq 2$. The channel uncertainty parameter ρ is set to be 0.3. When we consider the network with $R = 4$ relays, we consider the cases $M_i = 1, 3, 5$, respectively, for $i = 1, 2, 3, 4$. The $M_i = 1$ case corresponds to the method proposed in [11]. In Fig. 8, when the antenna number at each relay increases from 1 to 3, the average worst-case SNR increases to 5.5 dB. By contrast, when this number further increases to 5, the average worst-case SNR increases to 2 dB. We then investigate the impact of the relay number in the system. Here, we assume that each relay is equipped with two antennas. The $R = 1$ curve corresponds to the method used in [12]. Similar result can be observed when we vary the relay number from 1 to 3, where the SNR increases to about 4.1 dB. If we further increase the relay number to 5, the SNR increases to about 0.7 dB. Fig. 8 shows that increasing the relay number and relay antenna number are both beneficial. Moreover, one can greatly improve the system performance by slightly increasing the relay number or relay antenna number, which validated the importance of this paper.

VII. CONCLUSION

In this paper, we have considered a multiantenna multirelay channel with one source and one destination. Assuming that the relay only amplifies and forwards its received signals, we present a global optimal BF design in the robust case. To maximize the worst-case received SNR, we aim to jointly design the BF matrices at the source and the relays under individual power constraints at the source and the relays. We give a semi-closed form of the relay BF matrices up to a power scalar factor. The optimal and suboptimal algorithms for solving the source BF vector are also proposed. Numerical results verify the advantage of the proposed algorithm over the existing methods.

APPENDIX A PROOF OF LEMMA 2

Suppose that the singular value decomposition of \mathbf{u}_i is

$$\mathbf{u}_i = \mathbf{U}_i \begin{bmatrix} \|\mathbf{u}_i\|_2 \\ \mathbf{0}_{M_i-1} \end{bmatrix} \triangleq \mathbf{U}_i \boldsymbol{\Sigma}_i \quad (28)$$

where the unitary matrix $\mathbf{U}_i \in \mathbb{C}^{M_i \times M_i}$. Then, we can express the relay BF matrices as

$$\mathbf{B}_i = \mathbf{Y}_i \mathbf{U}_i^H \quad (29)$$

where $\mathbf{Y}_i \in \mathbb{C}^{M_i \times M_i}$ is a matrix to be determined. Upon substituting (29) and (28) into (4), the max-min SNR problem subject to the individual power constraints is given by

$$\max_{\mathbf{Y}_i} \min_{\Delta \mathbf{f} \in \mathcal{A}} \frac{\left| \sum_{i=1}^R \mathbf{f}_i^T \mathbf{Y}_i \boldsymbol{\Sigma}_i \right|^2}{\sigma_D^2 + \sigma_R^2 \sum_{i=1}^R \mathbf{f}_i^T \mathbf{Y}_i \mathbf{Y}_i^H \mathbf{f}_i^*} \quad (30a)$$

$$\text{s.t. } \text{tr}(\mathbf{Y}_i (\boldsymbol{\Sigma}_i \boldsymbol{\Sigma}_i^H + \sigma_R^2) \mathbf{Y}_i^H) \leq P_i. \quad (30b)$$

We can further partition \mathbf{Y}_i as

$$\mathbf{Y}_i = [\mathbf{b}_i \quad \mathbf{Z}_{yi}]$$

where $\mathbf{b}_i \in \mathbb{C}^{M_i \times 1}$, and $\mathbf{Z}_{yi} \in \mathbb{C}^{M_i \times (M_i-1)}$. Then, we have

$$\mathbf{Y}_i \boldsymbol{\Sigma}_i = [\mathbf{b}_i \quad \mathbf{Z}_{yi}] \begin{bmatrix} \|\mathbf{u}_i\|_2 \\ \mathbf{0} \end{bmatrix} = \|\mathbf{u}_i\|_2 \mathbf{b}_i. \quad (31)$$

Upon substituting (31) into (30), we have the received SNR at the destination as

$$\begin{aligned} \text{SNR} &= \frac{\left| \sum_{i=1}^R \mathbf{f}_i^T \mathbf{b}_i \|\mathbf{u}_i\|_2 \right|^2}{\sigma_D^2 + \sigma_R^2 \sum_{i=1}^R \|\mathbf{f}_i^T \mathbf{Y}_i\|_2^2} \\ &= \frac{\left| \sum_{i=1}^R \mathbf{f}_i^T \mathbf{b}_i \|\mathbf{u}_i\|_2 \right|^2}{\sigma_D^2 + \sigma_R^2 \sum_{i=1}^R \left(\|\mathbf{f}_i^T \mathbf{b}_i\|_2^2 + \|\mathbf{f}_i^T \mathbf{Z}_{yi}\|_2^2 \right)} \end{aligned} \quad (32)$$

and the individual relay power becomes

$$\begin{aligned} \text{tr}(\mathbf{Y}_i [\boldsymbol{\Sigma}_i \boldsymbol{\Sigma}_i^H + \sigma_R^2] \mathbf{Y}_i^H) &= \text{tr}(\mathbf{b}_i (\|\mathbf{u}_i\|_2^2 + \sigma_R^2) \mathbf{b}_i^H) + \sigma_R^2 \text{tr}(\mathbf{Z}_{yi} \mathbf{Z}_{yi}^H) \\ &= (\|\mathbf{u}_i\|_2^2 + \sigma_R^2) \|\mathbf{b}_i\|_2^2 + \sigma_R^2 \text{tr}(\mathbf{Z}_{yi} \mathbf{Z}_{yi}^H). \end{aligned}$$

From (32), to achieve maximum SNR with respect to \mathbf{Y}_i , we must minimize the denominator of SNR by forcing $\mathbf{Z}_{yi} = \mathbf{0}$. Then, we can express \mathbf{B}_i as

$$\mathbf{B}_i = \mathbf{b}_i (\mathbf{U}_i)_1^H = \mathbf{b}_i \hat{\mathbf{u}}_i^H \quad (33)$$

where $(\mathbf{U}_i)_1$ denotes the first column of \mathbf{U}_i . By substituting (33) into (4), we get (9).

APPENDIX B PROOF OF LEMMA 3

When $\mathbf{b}_i = c_i \hat{\mathbf{f}}_i^*$, we have $\mathbf{B}_i = c_i \hat{\mathbf{f}}_i^* \hat{\mathbf{u}}_i^H$ by Lemma 2. Then, the objective function of (4) becomes

$$\begin{aligned} & \frac{\left| \sum_{i=1}^R (\tilde{\mathbf{f}}_i + \eta_i \tilde{\mathbf{f}}_i^{\parallel} + \tau_i \tilde{\mathbf{f}}_i^{\perp})^T c_i \hat{\mathbf{f}}_i^* \|\mathbf{u}_i\|_2 \right|^2}{\sigma_R^2 \sum_{i=1}^R \left\| \left(\tilde{\mathbf{f}}_i + \eta_i \tilde{\mathbf{f}}_i^{\parallel} + \tau_i \tilde{\mathbf{f}}_i^{\perp} \right)^T c_i \hat{\mathbf{f}}_i^* \right\|_2^2 + \sigma_D^2} \\ &= \frac{\left| \sum_{i=1}^R (\|\tilde{\mathbf{f}}_i\|_2 + \eta_i) c_i \|\mathbf{u}_i\|_2 \right|^2}{\sigma_R^2 \sum_{i=1}^R \left(\|\tilde{\mathbf{f}}_i\|_2 + \eta_i \right)^2 |c_i|^2 + \sigma_D^2} \end{aligned} \quad (34)$$

where we have decomposed $\Delta \mathbf{f}_i = \eta_i \tilde{\mathbf{f}}_i^{\parallel} + \tau_i \tilde{\mathbf{f}}_i^{\perp}$, with $|\eta_i|^2 + |\tau_i|^2 \leq \varepsilon_i^2$ and $\eta_i, \tau_i \in \mathbb{C}$. Equation (34) implies that, when $\mathbf{b}_i = c_i \hat{\mathbf{f}}_i^*$, only η_i affects the minimum value of (34). Then, we can focus on $\Delta \mathbf{f}_i = \eta_i \tilde{\mathbf{f}}_i^{\parallel}$ or $\mathbf{f}_i = (\|\tilde{\mathbf{f}}_i\|_2 + \eta_i) \tilde{\mathbf{f}}_i^{\parallel}$, with $|\eta_i| \leq \varepsilon_i$. Thus, (4) is equivalent to

$$\max_c \min_{|\eta_i| \leq \varepsilon_i} \frac{\left| \sum_{i=1}^R (\|\tilde{\mathbf{f}}_i\|_2 + \eta_i) c_i \|\mathbf{u}_i\|_2 \right|^2}{\sigma_R^2 \sum_{i=1}^R \left(\|\tilde{\mathbf{f}}_i\|_2 + \eta_i \right)^2 |c_i|^2 + \sigma_D^2} \quad (35a)$$

$$\text{s.t. } |c_i| \leq \sqrt{\frac{P_i}{\sigma_R^2 + \|\mathbf{u}_i\|_2^2}}. \quad (35b)$$

It is worth noting that, in (35), η_i is a complex value. We will show in the following that (35) can be transformed into a problem with real-valued variable η_i , which is further

limited to $\pm\varepsilon_i$. This paper comes from the idea of *real-valued implementation* that has recently been proposed in [26]. Define

$$\begin{aligned} f_{\eta_i} &\triangleq \|\tilde{\mathbf{f}}_i\|_2 + \eta_i \\ \mathbf{R}_s &\triangleq (\mathbf{u} \odot \mathbf{f}_\eta)^* (\mathbf{u} \odot \mathbf{f}_\eta)^T \\ \mathbf{R}_n &\triangleq \sigma_R^2 \text{diag} [\|\mathbf{f}_\eta\|^2] \end{aligned}$$

where $\mathbf{u} \triangleq [\|\mathbf{u}_1\|_2, \dots, \|\mathbf{u}_R\|_2]^T$, and the operator \odot denotes the point-wise multiplication of two vectors. Then, we can write the objective of (35) as

$$\text{SNR} = \frac{\mathbf{c}^H \mathbf{R}_s \mathbf{c}}{\mathbf{c}^H \mathbf{R}_n \mathbf{c} + \sigma_D^2}. \quad (36)$$

Note that \mathbf{R}_n is a real-valued diagonal matrix, whereas \mathbf{R}_s is in general complex valued. The *real-valued implementation* idea [26] aims to transform \mathbf{R}_s into a real-valued matrix. First, we can write $\mathbf{u} \odot \mathbf{f}_\eta = \mathbf{u} \odot |\mathbf{f}_\eta| \odot \varphi$, where $\varphi \triangleq [e^{j\varphi_1}, \dots, e^{j\varphi_R}]^T$, $e^{j\varphi_i}$ denotes the phase of f_{η_i} , and $j = \sqrt{-1}$. Then, for any complex vector \mathbf{c} , one can always decompose it into $\mathbf{c} = \tilde{\mathbf{c}} \odot \tilde{\varphi}$, where $\tilde{\varphi} \triangleq [e^{-j\varphi_1}, \dots, e^{-j\varphi_R}]^T$, and $\tilde{\mathbf{c}}$ is determined by element-wise division between \mathbf{c} and $\tilde{\varphi}$. By referring to (36), the objective of (35) is given by

$$\begin{aligned} \text{SNR} &= \frac{(\tilde{\mathbf{c}} \odot \tilde{\varphi})^H (\mathbf{u} \odot |\mathbf{f}_\eta| \odot \varphi)^* (\mathbf{u} \odot |\mathbf{f}_\eta| \odot \varphi)^T (\tilde{\mathbf{c}} \odot \tilde{\varphi})}{(\tilde{\mathbf{c}} \odot \varphi)^H \mathbf{R}_n (\tilde{\mathbf{c}} \odot \varphi) + \sigma_D^2} \\ &= \frac{\tilde{\mathbf{c}}^H (\mathbf{u} \odot |\mathbf{f}_\eta|)^* (\mathbf{u} \odot |\mathbf{f}_\eta|)^T \tilde{\mathbf{c}}}{\tilde{\mathbf{c}}^H \mathbf{R}_n \tilde{\mathbf{c}} + \sigma_D^2} \\ &= \frac{\tilde{\mathbf{c}}^H \bar{\mathbf{R}}_s \tilde{\mathbf{c}}}{\tilde{\mathbf{c}}^H \mathbf{R}_n \tilde{\mathbf{c}} + \sigma_D^2} \end{aligned}$$

where $\bar{\mathbf{R}}_s \triangleq (\mathbf{u} \odot |\mathbf{f}_\eta|)^* (\mathbf{u} \odot |\mathbf{f}_\eta|)^T$ is a real-valued matrix. Notice that, for any real-valued $\bar{\mathbf{R}}_s$ and \mathbf{R}_n , by maximizing the received SNR, the corresponding $\tilde{\mathbf{c}}$ must be real valued [11], [26]. Now, (35) can be rewritten as

$$\max_{\tilde{\mathbf{c}} \in \mathbb{R}^R} \min_{|\eta_i| \leq \varepsilon_i} \frac{\tilde{\mathbf{c}}^H \bar{\mathbf{R}}_s \tilde{\mathbf{c}}}{\tilde{\mathbf{c}}^H \mathbf{R}_n \tilde{\mathbf{c}} + \sigma_D^2} \quad (37a)$$

$$\text{s.t. } |\tilde{c}_i| \leq \sqrt{\frac{P_i}{\sigma_R^2 + \|\mathbf{u}_i\|_2^2}}. \quad (37b)$$

In [26], since φ is fixed in their perfect CSI assumption, they find the optimal solution $\tilde{\mathbf{c}}$ in (37) and obtain \mathbf{c} by $\mathbf{c} = \tilde{\mathbf{c}} \odot \tilde{\varphi}$. In our case, things are a bit different. The value of $\tilde{\varphi}$ is not important in this problem. From (37), we can see that it is $|\mathbf{f}_\eta|$ rather than $\tilde{\varphi}$ that affects the value of worst-case SNR. For any given $|\eta_i| \leq \varepsilon_i$, $\eta_i \in \mathbb{C}$, we can find a real-valued $|\tilde{\eta}_i| \leq \varepsilon_i$, $\tilde{\eta}_i \in \mathbb{R}$, such that $\|\|\tilde{\mathbf{f}}_i\|_2 + \eta_i\| = \|\|\tilde{\mathbf{f}}_i\|_2 + \tilde{\eta}_i\|$, as shown in Fig. 9. Therefore, considering real-valued η_i will not lose the optimality of (37) or, equivalently, (35). By slight abuse of notation \mathbf{c} instead of $\tilde{\mathbf{c}}$, we can transform (35) into

$$\max_{\mathbf{c}} \min_{-\varepsilon_i \leq \eta_i \leq \varepsilon_i} \frac{\left(\sum_{i=1}^R f_{\eta_i} c_i \|\mathbf{u}_i\|_2\right)^2}{\sigma_R^2 \sum_{i=1}^R f_{\eta_i}^2 c_i^2 + \sigma_D^2} \quad (38a)$$

$$\text{s.t. } c_i \leq \sqrt{\frac{P_i}{\sigma_R^2 + \|\mathbf{u}_i\|_2^2}}. \quad (38b)$$

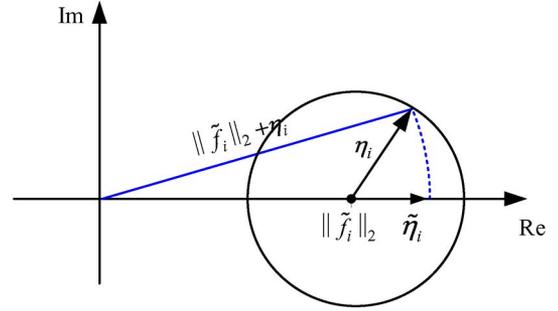


Fig. 9. By rotating $\|\tilde{\mathbf{f}}_i\|_2 + \eta_i$ to the real axis, one can always find a real-valued $\tilde{\eta}_i$ such that $\|\|\tilde{\mathbf{f}}_i\|_2 + \eta_i\| = \|\|\tilde{\mathbf{f}}_i\|_2 + \tilde{\eta}_i\|$.

Introducing a slack variable γ , problem (38) is transformed into

$$\max_{\mathbf{c}, \gamma} \gamma \quad (39a)$$

$$\text{s.t. } \min_{-\varepsilon_i \leq \eta_i \leq \varepsilon_i} \frac{\left(\sum_{i=1}^R f_{\eta_i} c_i \|\mathbf{u}_i\|_2\right)^2}{\sigma_R^2 \sum_{i=1}^R f_{\eta_i}^2 c_i^2 + \sigma_D^2} \geq \gamma \quad (39b)$$

$$c_i \leq \sqrt{\frac{P_i}{\sigma_R^2 + \|\mathbf{u}_i\|_2^2}} \quad (39c)$$

which is equivalent to the following problem:

$$\max_{\mathbf{c}, \gamma} \gamma \quad (40a)$$

$$\text{s.t. } \frac{\left(\sum_{i=1}^R f_{\eta_i} c_i \|\mathbf{u}_i\|_2\right)^2}{\sigma_R^2 \sum_{i=1}^R f_{\eta_i}^2 c_i^2 + \sigma_D^2} \geq \gamma, \quad -\varepsilon_i \leq \eta_i \leq \varepsilon_i \quad (40b)$$

$$c_i \leq \sqrt{\frac{P_i}{\sigma_R^2 + \|\mathbf{u}_i\|_2^2}}. \quad (40c)$$

Let

$$f(\mathbf{f}_\eta) \triangleq -\sum_{i=1}^R f_{\eta_i} c_i \|\mathbf{u}_i\|_2 + \sqrt{\gamma \left[\sigma_R^2 \sum_{i=1}^R f_{\eta_i}^2 c_i^2 + \sigma_D^2 \right]}. \quad (41)$$

Equation (40b) is equivalent to $\max_{-\varepsilon_i \leq \eta_i \leq \varepsilon_i} f(\mathbf{f}_\eta) \leq 0$. Note that $f(\mathbf{f}_\eta)$ is convex in \mathbf{f}_η and reaches the maximization at the vertices [11]. Hence, the optimal solution of problem (40) can be obtained by enumerating 2^R possibilities of \mathbf{f}_η or, i.e., $\mathbf{f}_\eta \in \mathcal{B}$, with each one corresponding to an SOCP constraint. Equivalently

$$\max_{\mathbf{c}, \gamma} \gamma \quad (42a)$$

$$\text{s.t. } \frac{\left(\sum_{i=1}^R f_{\eta_i} c_i \|\mathbf{u}_i\|_2\right)^2}{\sigma_R^2 \sum_{i=1}^R f_{\eta_i}^2 c_i^2 + \sigma_D^2} \geq \gamma, \quad \mathbf{f}_\eta \in \mathcal{B} \quad (42b)$$

$$c_i \leq \sqrt{\frac{P_i}{\sigma_R^2 + \|\mathbf{u}_i\|_2^2}}. \quad (42c)$$

Notice that (42) is equivalent to the form in (8); our proof is completed.

APPENDIX C
PROOF OF PROPOSITION 1

Here, we will first prove that problem (23) belongs to the class of monotonic optimization problem, or more specifically, $\text{SNR}(\mathbf{w})$ is an increasing function with respect to $\mathbf{w} \in \mathcal{U}$. Then, we will show that problems (20) and (23) are equivalent.

In (21), we have expressed the worst-case SNR as a function of \mathbf{w} , where the power-allocation factor $\mathbf{c}^\#$ is adaptively determined as an optimal solution of (22) with respect to \mathbf{w} . For convenience, we further define $\widetilde{\text{SNR}}(\mathbf{c}, \mathbf{w})$ as a function of \mathbf{w} and \mathbf{c} , where \mathbf{c} is only one possible power-allocation option rather than the optimal choice or, i.e.,

$$\widetilde{\text{SNR}}(\mathbf{c}, \mathbf{w}) \triangleq \min_{\mathbf{f}_\eta \in \mathcal{B}} \frac{\left(\sum_{i=1}^R f_{\eta i} c_i \sqrt{w_i} \right)^2}{\sigma_R^2 \sum_{i=1}^R f_{\eta i}^2 c_i^2 + \sigma_D^2} \quad (43a)$$

$$\text{s.t. } c_i \leq \sqrt{\frac{P_i}{\sigma_R^2 + w_i}}. \quad (43b)$$

Then, by definition, one can easily see that $\text{SNR}(\mathbf{w}) = \widetilde{\text{SNR}}(\mathbf{c}^\#, \mathbf{w})$. Suppose $\mathbf{w}' \geq \mathbf{w}''$, where $\mathbf{w}' \triangleq [w'_1, \dots, w'_R]^T$ and $\mathbf{w}'' \triangleq [w''_1, \dots, w''_R]^T$. Let $\mathbf{c}^\# \triangleq [c_1^\#, \dots, c_R^\#]^T$ and $\mathbf{c}'' \triangleq [c_1'', \dots, c_R'']^T$ be the optimal solution of (22) for given \mathbf{w}' and \mathbf{w}'' , respectively. We will show that $\text{SNR}(\mathbf{w}') \geq \text{SNR}(\mathbf{w}'')$ or, equivalently

$$\widetilde{\text{SNR}}(\mathbf{c}^\#, \mathbf{w}') \geq \widetilde{\text{SNR}}(\mathbf{c}''^\#, \mathbf{w}''). \quad (44)$$

Choose one special relay power-allocation factor for the given \mathbf{w}' as $\tilde{\mathbf{c}} \triangleq [\tilde{c}_1, \dots, \tilde{c}_R]^T$, such that

$$\tilde{c}_i^2 (w_i + \sigma_R^2) = c_i''^2 (w_i'' + \sigma_R^2), \quad i = 1, \dots, R. \quad (45)$$

By this condition, the relay power is unchanged; thus, the power constraints in (43b) are not violated. Since $w_i' \geq w_i''$, we have $w_i' + \sigma_R^2 \geq w_i'' + \sigma_R^2$. Then, $\tilde{c}_i^2 \leq c_i''^2$ by (45), which implies that $\tilde{c}_i^2 \sigma_R^2 \leq c_i''^2 \sigma_R^2$. Then, by (45), we have

$$\tilde{c}_i^2 w_i' \geq c_i''^2 w_i''. \quad (46)$$

Let $\Gamma_1(\mathbf{f}_\eta) \triangleq (\sum_{i=1}^R f_{\eta i} \tilde{c}_i \sqrt{w_i})^2 / (\sigma_R^2 \sum_{i=1}^R f_{\eta i}^2 \tilde{c}_i^2 + \sigma_D^2)$ and $\Gamma_2(\mathbf{f}_\eta) \triangleq (\sum_{i=1}^R f_{\eta i} c_i'' \sqrt{w_i''})^2 / (\sigma_R^2 \sum_{i=1}^R f_{\eta i}^2 c_i''^2 + \sigma_D^2)$ for $\mathbf{f}_\eta \in \mathcal{B}$. Then, we have

$$\widetilde{\text{SNR}}(\tilde{\mathbf{c}}, \mathbf{w}') = \min_{\mathbf{f}_\eta \in \mathcal{B}} \Gamma_1(\mathbf{f}_\eta) \quad (47)$$

$$\widetilde{\text{SNR}}(\mathbf{c}''^\#, \mathbf{w}'') = \min_{\mathbf{f}_\eta \in \mathcal{B}} \Gamma_2(\mathbf{f}_\eta). \quad (48)$$

We first fix some $\mathbf{f}_\eta \in \mathcal{B}$. Note that $|\eta_i| \leq \varepsilon_i \leq \|\tilde{\mathbf{f}}_i\|_2$; thus, we have $f_{\eta i} = \|\tilde{\mathbf{f}}_i\|_2 + \eta_i \geq 0$. Then, for fixed \mathbf{f}_η , the numerator of $\Gamma_1(\mathbf{f}_\eta)$ is larger than that of $\Gamma_2(\mathbf{f}_\eta)$ due to (46), whereas the denominator of $\Gamma_1(\mathbf{f}_\eta)$ is smaller than that of $\Gamma_2(\mathbf{f}_\eta)$ due to $\tilde{c}_i^2 \leq c_i''^2$. Hence, for any $\mathbf{f}_\eta \in \mathcal{B}$, we have

$$\Gamma_1(\mathbf{f}_\eta) \geq \Gamma_2(\mathbf{f}_\eta). \quad (49)$$

Suppose that the minimum value of $\Gamma_1(\mathbf{f}_\eta)$ over $\mathbf{f}_\eta \in \mathcal{B}$ is achieved at \mathbf{f}'_η , i.e., $\min_{\mathbf{f}_\eta \in \mathcal{B}} \Gamma_1(\mathbf{f}_\eta) = \Gamma_1(\mathbf{f}'_\eta)$. Then, we have

$$\min_{\mathbf{f}_\eta \in \mathcal{B}} \Gamma_1(\mathbf{f}_\eta) = \Gamma_1(\mathbf{f}'_\eta) \stackrel{(a)}{\geq} \Gamma_2(\mathbf{f}'_\eta) \geq \min_{\mathbf{f}_\eta \in \mathcal{B}} \Gamma_2(\mathbf{f}_\eta) \quad (50)$$

where (a) is due to (49). Then, (47), (48), and (50) lead to

$$\widetilde{\text{SNR}}(\tilde{\mathbf{c}}, \mathbf{w}') \geq \widetilde{\text{SNR}}(\mathbf{c}''^\#, \mathbf{w}''). \quad (51)$$

Since $\tilde{\mathbf{c}}$ is just chosen to satisfy (45) and may not be optimal for $\mathbf{w} = \mathbf{w}'$, we have

$$\widetilde{\text{SNR}}(\mathbf{c}^\#, \mathbf{w}') \geq \widetilde{\text{SNR}}(\tilde{\mathbf{c}}, \mathbf{w}'). \quad (52)$$

By (51) and (52), we have (44), which implies that $\text{SNR}(\mathbf{w})$ is a monotonic increasing function with respect to \mathbf{w} .

On the other hand, \mathcal{U} has been proven to be convex [32]. Consequently, \mathcal{U} is normal due to the property of the convex region [35]. Following the similar lines in [32], it can be shown that \mathcal{U} is nonempty and closed. Thus, (23) is a monotonic optimization problem.

As compared with other nonconvex problems, monotonic problems have the important property that its optimal solution is attained on the Pareto boundary of the feasible region, which can be utilized for solving the problem efficiently.

According to [32], any Pareto boundary of \mathcal{U} must be achieved by some rank-one matrix \mathbf{G} ; we claim that problems (20) and (23) are equivalent.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewer for their great constructive comments that improved this paper.

REFERENCES

- [1] Y. Liu and W. Chen, "Adaptive resource allocation for improved DF aided downlink multi-user OFDM systems," *IEEE Wireless Commun. Lett.*, vol. 1, no. 6, pp. 557–560, Dec. 2012.
- [2] H. Chen, S. Shahbazpanahi, and A. B. Gershman, "Filter-and-forward distributed beamforming for two-way relay networks with frequency selective channels," *IEEE Trans. Signal Process.*, vol. 60, no. 4, pp. 1927–1941, Apr. 2012.
- [3] Y. Liang, A. Ikhlef, W. H. Gerstacker, and R. Schober, "Two-way filter-and-forward beamforming for frequency-selective channels," *IEEE Trans. Wireless Commun.*, vol. 10, no. 12, pp. 4172–4183, Dec. 2011.
- [4] Y. Liang, A. Ikhlef, W. H. Gerstacker, and R. Schober, "Cooperative filter-and-forward beamforming for frequency-selective channels with equalization," *IEEE Trans. Wireless Commun.*, vol. 10, no. 1, pp. 228–239, Jan. 2011.
- [5] Z. Wang, W. Chen, and J. Li, "Efficient beamforming for MIMO relaying broadcast channel with imperfect channel estimation," *IEEE Trans. Veh. Technol.*, vol. 61, no. 1, pp. 419–426, Jan. 2012.
- [6] Y. Zhang, H. Luo, and W. Chen, "Efficient relay beamforming design with SIC detection for dual-hop MIMO relay networks," *IEEE Trans. Veh. Technol.*, vol. 59, no. 8, pp. 4192–4197, Oct. 2010.
- [7] Y. Jing and H. Jafarkhani, "Network beamforming using relays with perfect channel information," *IEEE Trans. Inf. Theory*, vol. 55, no. 6, pp. 2499–2517, Jun. 2009.
- [8] B. Khoshnevis, W. Yu, and R. Adve, "Grassmannian beamforming for MIMO amplify-and-forward relaying," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 8, pp. 1397–1408, Oct. 2008.
- [9] Y. Liang and R. Schober, "Cooperative amplify-and-forward beamforming with multiple multi-antenna relays," *IEEE Trans. Commun.*, vol. 59, no. 9, pp. 2605–2615, Sep. 2011.

- [10] P. Ubaidulla and A. Chockalingam, "Robust distributed beamforming for wireless relay networks," in *Proc. IEEE 20th Int. Symp. Pers., Indoor Mobile Radio Commun.*, Sep. 2009, pp. 2345–2349.
- [11] G. Zheng, K. K. Wong, A. Paulraj, and B. Ottersten, "Robust collaborative-relay beamforming," *IEEE Trans. Signal Process.*, vol. 57, no. 8, pp. 3130–3143, Aug. 2009.
- [12] H. Shen, W. Xu, J. Wang, and C. Zhao, "A worst case robust beamforming design for multi-antenna AF relaying," *IEEE Commun. Lett.*, vol. 17, no. 4, pp. 1089–1092, Apr. 2013.
- [13] Z. Wang and W. Chen, "Relay beamforming design with SIC detection for MIMO multi-relay networks with imperfect CSI," *IEEE Trans. Veh. Technol.*, vol. 62, no. 8, pp. 3774–3785, Oct. 2013.
- [14] H. Shen, J. Wang, B. C. Levy, and C. Zhao, "Robust optimization for amplify-and-forward MIMO relaying from a worst-case perspective," *IEEE Trans. Signal Process.*, vol. 61, no. 21, pp. 5458–5471, Nov. 2013.
- [15] H. Wan and W. Chen, "Joint source and relay design for multi-user MIMO non-regenerative relay networks with direct links," *IEEE Trans. Veh. Technol.*, vol. 61, no. 6, pp. 2871–2876, Jul. 2012.
- [16] R. Wang, M. Tao, and Z. Xiang, "Nonlinear precoding design for MIMO amplify and forward two-way relay systems," *IEEE Trans. Veh. Technol.*, vol. 61, no. 9, pp. 3984–3995, Nov. 2012.
- [17] R. Wang, M. Tao, and Y. Huang, "Linear precoding designs for amplify-and-forward multiuser two-way relay systems," *IEEE Trans. Wireless Commun.*, vol. 11, no. 12, pp. 4457–4469, Dec. 2012.
- [18] M. Tao and R. Wang, "Robust relay beamforming for two-way relay networks," *IEEE Commun. Lett.*, vol. 16, no. 7, pp. 1052–1055, Jul. 2012.
- [19] R. Wang and M. Tao, "Joint source and relay precoding designs for MIMO two-way relaying based on MSE criterion," *IEEE J. Sel. Areas Commun.*, vol. 60, no. 3, pp. 1352–1365, Mar. 2012.
- [20] J. Zou, H. Luo, M. Tao, and R. Wang, "Joint source and relay optimization for non-regenerative MIMO two-way relay systems with imperfect CSI," *IEEE Trans. Wireless Commun.*, vol. 11, no. 9, pp. 3305–3315, Sep. 2012.
- [21] A. Aziz, Z. Meng, Z. Jianwei, C. N. Georghiadis, and C. Shuguang, "Robust beamforming with channel uncertainty for two-way relay networks," in *Proc. IEEE ICC*, Jun. 2012, pp. 3632–3636.
- [22] C. Kuo, S. Wu, and C. Tseng, "Robust linear beamformer designs for coordinated multi-point AF relaying in downlink multi-cell networks," *IEEE Trans. Veh. Technol.*, vol. 11, no. 7, pp. 3272–3283, Sep. 2012.
- [23] B. K. Chalise and L. Vandendorpe, "Optimization of MIMO relays for multipoint-to-multipoint communications: Nonrobust and robust designs," *IEEE Trans. Signal. Process.*, vol. 58, no. 12, pp. 6355–6368, Dec. 2010.
- [24] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
- [25] A. Nemirovski, *Lectures on Modern Convex Optimization*. Philadelphia, PA, USA: SIAM, 2011.
- [26] L. Zhang, W. Liu, and J. Li, "Low-complexity distributed beamforming for relay networks with real-valued implementation," *IEEE Trans. Signal. Process.*, vol. 61, no. 20, pp. 5039–5048, Oct. 2013.
- [27] W. Wang, S. Jin, and F. Zheng, "Maximin SNR beamforming strategies for two-way relay channels," *IEEE Commun. Lett.*, vol. 16, no. 7, pp. 1006–1009, Jul. 2012.
- [28] Z. Fang, X. Wang, and X. Yuan, "Beamforming design for multiuser two-way relaying: A unified approach via max-min SINR," *IEEE Trans. Signal. Process.*, vol. 61, no. 23, pp. 5841–5852, Dec. 2013.
- [29] L. Liu, R. Zhang, and K. C. Chua, "Achieving global optimality for weighted sum-rate maximization in the K-user Gaussian interference channel with multiple antennas," *IEEE Trans. Wireless Commun.*, vol. 11, no. 5, pp. 1933–1945, May 2012.
- [30] E. Björnson, G. Zheng, M. Bengtsson, and B. Ottersten, "Robust monotonic optimization framework for multicell MISO systems," *IEEE Trans. Signal. Process.*, vol. 60, no. 5, pp. 2508–2523, May 2012.
- [31] E. Björnson and E. Jorswieck, "Optimal resource allocation in coordinated multi-cell systems," *Found. Trends Commun. Inf. Theory*, vol. 9, no. 2/3, pp. 113–381, 2013.
- [32] R. Mochaourab and E. A. Jorswieck, "Optimal beamforming in interference networks with perfect local channel information," *IEEE Trans. Signal. Process.*, vol. 59, no. 3, pp. 1128–1141, Mar. 2011.
- [33] M. Grant and S. Boyd, *CVX' Users' Guide 2009*. [Online]. Available: <http://cvxr.com/cvx/doc/index.html>
- [34] 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.
- [35] J. Brehmer, *Utility Maximization in Nonconvex Wireless Systems*. Berlin, Germany: Springer-Verlag, 2012.
- [36] M. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret, "Applications of second-order cone programming," *Linear Algebra Appl.*, vol. 284, no. 1–3, pp. 193–228, Nov. 1998.



techniques.

Hongying Tang received the B.S. and M.S. degrees in electronic engineering from the University of Electronic Science and Technology of China, Chengdu, China, and Shanghai Jiao Tong University, Shanghai, China, in 2007 and 2009, respectively. She is currently working toward the Ph.D. degree with the Network Coding and Transmission Laboratory, Department of Electronic Engineering, Shanghai Jiao Tong University.

Her current research interests include cooperative networks and multiple-input–multiple-output



Wen Chen (M'03–SM'11) received the B.S. and M.S. degrees from Wuhan University, Wuhan, China, in 1990 and 1993, respectively, and the Ph.D. degree from The University of Electro-Communications, Tokyo, Japan, in 1999.

From 1991 to 2001, he was a Researcher with the Japan Society for the Promotion of Sciences. In 2001, he was with the University of Alberta, Edmonton, AB, Canada, first as a Postdoctoral Fellow with the Information Research Laboratory and then as a Research Associate with the Department of Electrical and Computer Engineering. Since 2006, he has been a Full Professor with the Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai, China, where he is also the Director of the Institute for Signal Processing and Systems. His research interests include network coding, cooperative communications, cognitive radio, and multiple-input–multiple-output orthogonal frequency-division multiplexing systems.



Jun Li (M'09) received the Ph.D. degree in electronic engineering from Shanghai Jiao Tong University, Shanghai, China, in 2009.

From January 2009 to June 2009, he was with the Department of Research and Innovation, Alcatel-Lucent Shanghai Bell Company Ltd., as a Research Scientist. From June 2009 to April 2012, he was a Postdoctoral Fellow with the School of Electrical Engineering and Telecommunications, the University of New South Wales, Sydney, Australia. Since April 2012, he has been a Research Fellow with the School of Electrical Engineering, the University of Sydney. His research interests include network information theory, channel coding theory, wireless network coding, and cooperative communications.

Dr. Li served as a Technical Program Committee member for several international conferences, such as the Asia–Pacific Conference on Communications in 2009, 2010, and 2013; the IEEE Vehicular Technology Conference in Spring 2011 and Fall 2014; the IEEE International Conference on Communications in 2011 and 2014; and the IEEE Region 10 Conference in 2012.

Dr. Li served as a Technical Program Committee member for several international conferences, such as the Asia–Pacific Conference on Communications in 2009, 2010, and 2013; the IEEE Vehicular Technology Conference in Spring 2011 and Fall 2014; the IEEE International Conference on Communications in 2011 and 2014; and the IEEE Region 10 Conference in 2012.



Haibin Wan received the Ph.D. degree in electronic engineering from Shanghai Jiao Tong University, Shanghai, China, in 2013.

He is currently with the College of Computer Science and Electronic Information, Guangxi University, Nanning, China. His research interests include cooperative communications, resource allocation, physical-layer security, and massive multiple-input–multiple-output techniques.