

Robust Joint Source-Relay-Destination Design Under Per-Antenna Power Constraints

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

Abstract—This paper deals with joint source-relay-destination beamforming (BF) design for an amplify-and-forward (AF) relay network. Considering the channel state information (CSI) from the relay to the destination is imperfect, we first aim to maximize the worst case received SNR under per-antenna power constraints. The associated optimization problem is then solved in two steps. In the first step, by revealing the rank-one property of the optimal relay BF matrix, we establish the semi-closed form solution of the joint optimal BF design that only depends on a vector variable. Based on this result, in the second step, we propose a low-complexity iterative algorithm to obtain the remaining unknown variable. We also study the problem for minimizing the maximum per-antenna power at the relay while ensuring a received signal-to-noise ratio (SNR) target, and show that it reduces to the SNR maximization problem. Thus the same methods can be applied to solve it. The differences between our result and the existing related work are also discussed in details. In particular, we show that in the perfect CSI case, our algorithm has the same performance but with much lower cost of computational complexity than the existing method. Finally, in the simulation part, we investigate the impact of imperfect CSI on the system performance to verify our analysis.

Index Terms—Amplify-and-forward, beamforming, multi-antenna relay system, per-antenna power.

I. INTRODUCTION

RELAY communications have been studied extensively over the past decades as a means of extending the coverage of wireless network and improving the spatial diversity of the system. Transmission schemes at the relay can be categorized into several groups, i.e., the Amplify-and-Forward (AF) scheme, the Decode-and-Forward scheme [1] etc. Among them, the AF scheme is the most simple scheme, as the relay only performs linear processing on the received signal and re-transmits it to the destination. AF scheme has been effi-

ciently used to exploit the benefit of relaying in the multiple access relay channels [2], the broadcast relay channels [3], and the two-hop relay channels [4]–[6], [11].

Performing joint source-relay linear beamforming (BF) can achieve higher data rate [2], [3]. For the relay BF design, many previous works concentrate on the sum power constraints. It is shown in [5], [6] that for the single user with single relay case, the optimal relay BF matrix \mathbf{W} can be seen as the combination of a maximum-ratio-combining (MRC) equalizer \mathbf{w}_r and a maximum-ratio-transmission (MRT) equalizer \mathbf{w}_t , or i.e., $\mathbf{W} = \mathbf{w}_r \mathbf{w}_t^H$. By using this BF matrix, the received SNR can be maximized. However, this optimal design matrix would result in different elemental power allocations on each antenna, which is undesirable from the power amplifier design perspective [7]. Considering the fact that each antenna usually uses the same type of power amplifier and consequently has the same power dynamic range and peak power, this would bring some difficulties on the power amplifier designs. For the relay node (usually low cost, with low-profile power amplifier) in a wireless network, setting a peak power threshold under per-antenna power control can successfully solve this problem and thus relax the power amplifier design effort. This is the motivation that we explore the per-antenna power constraints in a relay network in this paper.

Several papers have considered the per-antenna power constraints in different problem setups. In a MISO channel discussed by [8], a closed-form solution was derived for transmit BF design. A novel transceiver design under mixed power constraints (including the sum power constraints as well as per-antenna power constraints) for MIMO systems was investigated in [9], where the authors proposed their analytical method by exploiting the hidden physical meaning of the problem. For the multiuser downlink channel [10], the authors proposed a framework of efficient optimization technique for determining the downlink BF vector via a dual uplink problem. This uplink-downlink duality was recently extended into the relay system by [11], where a semi-closed form solution for the optimal relay BF matrix that depends on a set of dual variables was derived. However, this numerical result cannot provide any insight in designing the relay BF matrix and the complexity is also very high.

In general, perfect CSI is usually hard to obtain, due to many factors such as inaccurate channel estimation, channel quantization and feedback delay. Since the performance of a relay system is sensitive to the accuracy of available CSI, robust design taking channel uncertainty into account has attracted much attention. Generally, there are two widely used CSI uncertainty model in the literatures: the stochastic model and the deterministic model. For the statistical CSI uncertainty model, which as-

Manuscript received September 04, 2014; revised December 10, 2014; accepted March 01, 2015. Date of publication March 20, 2015; date of current version April 14, 2015. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Gustau Camps-Valls. This work was supported by the National 973 Project #2012CB316106, by NSF China #61328101, by the STCSM Science and Technology Innovation Program #13510711200, and by the SEU National Key Lab on Mobile Communications #2013D11.

H. Tang and W. Chen are with the Department of Electronic Engineering, Shanghai Jiaotong University, Shanghai 200240, China. They are also with the School of Electronic Engineering and Automation, Guilin University of Electronic Technology, Guilin, 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, New South Wales 2006, Australia (e-mail: jun.li1@sydney.edu.au).

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/TSP.2015.2414892

sumes the distribution of the CSI to be known and seeks to enhance the average system performance, the optimal relay precoder design was obtained in [12]. Later, this work was extended to the multi-hop relay channel in [14], [13]. In contrast, the deterministic CSI uncertainty model, assumes that the instantaneous value of CSI error is norm-bounded, and aims to yield worst-case guarantees. Under this model, references [15], [16] studied the corresponding robust problem in [5], [6], respectively, and obtained the optimal solution. Unfortunately, the extension from the aforementioned works to the relay BF design under per-antenna power constraints is not straightforward, and the existing designs are not applicable any more. To the best of our knowledge, the robust joint source-relay-destination BF design has not been studied in the existing literatures, and even in the perfect CSI case, the optimal solution of this problem is not yet known.

In this paper, we consider the AF-relay networks with a single source-destination pair and a single relay, and address the joint source-relay-destination BF design problem under deterministic imperfect CSI model. With per-antenna power constraints at the relay node, we adopt two widely used performance metrics, the maximization of the received SNR and the minimization of per-antenna power at the relay with a given required SNR. The main contributions of this paper are as follows:

- We first aim to maximize the received SNR from the worst case perspective. By fixing the source BF vector, we determine the rank one property of the optimal relay BF matrix, revealing that it can be decomposed as the combination of an MRC equalizer and an equalizer \mathbf{w} to be determined. Therefore, the original complicated matrix-valued problem has been converted into a much simpler problem with variable \mathbf{w} . Based on the above results, the optimal source BF vector is derived, and the destination BF vector is given as a function of \mathbf{w} . We further propose a low-complexity iterative algorithm for solving \mathbf{w} . Simulation results show that our robust design can significantly reduce the sensitivity of the channel uncertainty to the system performance.
- Then we consider minimizing per-antenna power at the relay node under a given received SNR target. We prove that it can also be transformed into the SNR maximization problem and thus can be solved by the same method.
- The differences between our result and the existing related work are also discussed in details. In particular, we show that in the perfect CSI case, our method has the same performance but with significantly lower complexity than that in [11].

Although only single data stream is discussed in this paper, our results can still provide some useful insights to the more general case when transmitting multiple data streams. Moreover, for the fifth generation (5G) cellular network, where millimeter-wave (mmWave) bands are used, the conventional microwave architecture where every antenna is connected to a high-rate ADC/DAC is unlikely to be applicable anymore [19]. Therefore lower order MIMO is preferred in this case. We believe that our result can have more applications in the next generation communication systems.

Notations: $[\cdot]^H$ denotes the conjugated transpose of a matrix or a vector. \mathbb{C}^n denotes the n dimensional complex field.

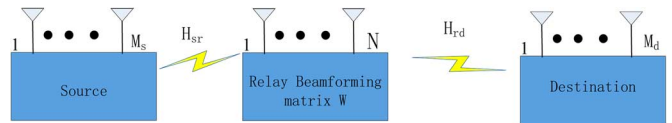


Fig. 1. A two-hop multi-antenna relay network.

We will use boldface lowercase letters to denote column vectors and boldface uppercase letters to denote matrices. $\mathbf{0}$ denotes a vector or matrix with all zeros entries. $\|\cdot\|_2$ denotes the Euclidean norm of a vector and $\|\cdot\|_F$ denotes the Frobenius norm of a matrix. $\mathbf{X} \succeq \mathbf{0}$ means that the matrix \mathbf{X} is symmetric positive semidefinite. $tr(\cdot)$ is the trace of a matrix. $\mathcal{E}(\cdot)$ is the expectation of a random variable. $\lambda_{\max}(\cdot)$ and $\mathbf{v}(\cdot)$ stand for the largest eigenvalue and the corresponding eigenvector of a matrix, respectively.

II. PROBLEM STATEMENT

In this section, we will introduce the system model, the channel uncertainty and the problem formulation.

A. System Model

We consider a two-hop AF multiantenna relay network as shown in Fig. 1, where the source and destination are equipped with M_s and M_d antennas respectively, and the relay is equipped with N antennas. The direct link 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 a symbol s with unit power, and the signal received by the relay is given by

$$\mathbf{y}_r = \sqrt{P_s} \mathbf{H}_{sr} \mathbf{b} s + \mathbf{n}_r,$$

where \mathbf{b} is an $M_s \times 1$ transmit beamforming (BF) vector with unit norm, $\mathbf{H}_{sr} \in \mathbb{C}^{N \times M_s}$ denotes the first hop channel matrix from the source to the relay, P_s is the given transmit power at the source, and \mathbf{n}_r denotes the N dimensional complex additive white Gaussian noise (AWGN) vector with variance matrix $\sigma_r^2 \mathbf{I}_N$ at the relay. By the AF strategy, the signal forwarded by the relay is

$$\mathbf{q} = \mathbf{W} \mathbf{y}_r,$$

where $\mathbf{W} \in \mathbb{C}^{N \times N}$ is the linear BF matrix of the relay. Then the transmit power at each antenna of the relay is given by

$$\mathcal{E} \left\{ \|\mathbf{W} \mathbf{y}_r\|_i^2 \right\} = \left[P_s \mathbf{W} \mathbf{H}_{sr} \mathbf{b} \mathbf{b}^H \mathbf{H}_{sr}^H \mathbf{W}^H + \sigma_r^2 \mathbf{W} \mathbf{W}^H \right]_{i,i},$$

where $[\mathbf{a}]_i$ denotes the i th element of the vector \mathbf{a} , and $[\mathbf{A}]_{i,i}$ denotes the i th diagonal entry of the matrix \mathbf{A} . The received signal at the destination node can be expressed as

$$\mathbf{y}_d = \sqrt{P_s} \mathbf{r}^H \mathbf{H}_{rd} \mathbf{W} \mathbf{H}_{sr} \mathbf{b} s + \mathbf{r}^H \mathbf{H}_{rd} \mathbf{W} \mathbf{n}_r + \mathbf{r}^H \mathbf{n}_d,$$

where $\mathbf{H}_{rd} \in \mathbb{C}^{M_d \times N}$ denotes the second hop channel matrix from the relay to the destination, \mathbf{r} is an $M_d \times 1$ receive BF vector with unit norm, and \mathbf{n}_d is an AWGN vector observed

at the destination with variance matrix $\sigma_d^2 \mathbf{I}_{M_d}$. The received signal-to-noise ratio (SNR) at the destination is then given by

$$\text{SNR} = \frac{P_s |\mathbf{r}^H \mathbf{H}_{rd} \mathbf{W} \mathbf{H}_{sr} \mathbf{b}|^2}{\sigma_r^2 \|\mathbf{r}^H \mathbf{H}_{rd} \mathbf{W}\|_2^2 + \sigma_d^2}.$$

B. Channel Uncertainty Model

In a practical wireless communication system, with only imperfect channel state information at the transmitter side (CSIT), the system performance will be deteriorated. This motivates us to investigate the robust design which takes the channel state information (CSI) errors into account.

In this paper, we assume that the CSI in the second hop varies much faster than that of the first hop, for example, the positions of the source (e.g. a base station) and the relay are fixed, and the destination is moving (e.g. a mobile terminal). Then the CSI feedback from the destination to the relay is usually outdated and the CSI error must be considered [20]. On the other hand, the hop between the source and the relay undergoes slow fading channel due to their fixed positions. When the relay transmits signals with pilots in the second time slot, it is possible for the source to estimate the first hop CSI nearly perfectly via the reciprocal channel, if the training SNR is high [21]. Hence the channel error in the first hop can be neglected. Based on the above assumptions, we only consider the CSIT uncertainty in the second hop. The authors in [22] also used this model for exploiting the situation when the relays are located closer to the source than to the destination, while this assumption is reasonable because of the high signal quality between the source and the relays. This model has also been widely used in the literatures such as [4], [15]–[17], [20], [23], [24].

Define the channel error matrix $\Delta \mathbf{H}$ as the difference between the actual channel \mathbf{H}_{rd} and the available channel $\tilde{\mathbf{H}}_{rd} \triangleq [\tilde{\mathbf{h}}_1, \dots, \tilde{\mathbf{h}}_N]$, i.e., $\Delta \mathbf{H} = \mathbf{H}_{rd} - \tilde{\mathbf{H}}_{rd}$. Then under the deterministic channel uncertainty model, $\Delta \mathbf{H}$ can be described as

$$\|\Delta \mathbf{H}\|_F \leq \varepsilon,$$

which indicates that the uncertainty channel matrix of the second hop at the relay node is norm bounded by some small positive number ε .

C. Problem Formulation

In this paper, we consider two widely used performance metrics: the maximization of the received SNR and the minimization of per-antenna power at the relay with a given SNR target.

1) *SNR Maximization*: By maximizing the worst case received SNR over the channel uncertainty region with per-antenna power constraints at the relay, the problem can be formulated as

$$\begin{aligned} \mathcal{P}_1 : \max_{\mathbf{W}, \mathbf{b}, \mathbf{r}} \min_{\|\Delta \mathbf{H}\|_F \leq \varepsilon} & \frac{P_s |\mathbf{r}^H (\tilde{\mathbf{H}}_{rd} + \Delta \mathbf{H}) \mathbf{W} \mathbf{H}_{sr} \mathbf{b}|^2}{\sigma_r^2 \|\mathbf{r}^H (\tilde{\mathbf{H}}_{rd} + \Delta \mathbf{H}) \mathbf{W}\|_2^2 + \sigma_d^2} \\ \text{s.t.} & \left[P_s \mathbf{W} \mathbf{H}_{sr} \mathbf{b} \mathbf{b}^H \mathbf{H}_{sr}^H \mathbf{W}^H + \sigma_r^2 \mathbf{W} \mathbf{W}^H \right]_{i,i} \leq P_r, \\ & \|\mathbf{b}\|_2 = 1, \|\mathbf{r}\|_2 = 1. \end{aligned}$$

where P_r is the maximum per-antenna power consumption of the relay node.

2) *Power Minimization Problem*: The per antenna power minimization problem with a given received SNR target γ_0 over all possible channel uncertainty errors is formulated as

$$\begin{aligned} \mathcal{P}_2 : \min_{\mathbf{W}, \mathbf{b}, \mathbf{r}, P_r} & \\ \text{s.t.} & \min_{\|\Delta \mathbf{H}\|_F \leq \varepsilon} \frac{P_s |\mathbf{r}^H (\tilde{\mathbf{H}}_{rd} + \Delta \mathbf{H}) \mathbf{W} \mathbf{H}_{sr} \mathbf{b}|^2}{\sigma_r^2 \|\mathbf{r}^H (\tilde{\mathbf{H}}_{rd} + \Delta \mathbf{H}) \mathbf{W}\|_2^2 + \sigma_d^2} \geq \gamma_0, \\ & \left[P_s \mathbf{W} \mathbf{H}_{sr} \mathbf{b} \mathbf{b}^H \mathbf{H}_{sr}^H \mathbf{W}^H + \sigma_r^2 \mathbf{W} \mathbf{W}^H \right]_{i,i} \leq P_r, \\ & \|\mathbf{b}\|_2 = 1, \|\mathbf{r}\|_2 = 1. \end{aligned}$$

In the following sections, we will first solve problem \mathcal{P}_1 , giving the closed-form solution of \mathbf{b} and the semi-closed form solution of (\mathbf{W}, \mathbf{r}) as a function of \mathbf{w} . Based on this result, we will further propose an iterative algorithm to obtain \mathbf{w} . Finally, we will show the relationship between problem \mathcal{P}_1 and problem \mathcal{P}_2 .

III. SNR MAXIMIZATION PROBLEM

A. A Semi-closed Form Solution of the Joint BF design

In this section, we will solve problem \mathcal{P}_1 . By fixing the source BF vector \mathbf{b} , problem \mathcal{P}_1 becomes

$$\begin{aligned} \max_{\mathbf{W}, \mathbf{r}} \min_{\|\Delta \mathbf{H}\|_F \leq \varepsilon} & \frac{P_s |\mathbf{r}^H (\tilde{\mathbf{H}}_{rd} + \Delta \mathbf{H}) \mathbf{W} \mathbf{g}|^2}{\sigma_r^2 \|\mathbf{r}^H (\tilde{\mathbf{H}}_{rd} + \Delta \mathbf{H}) \mathbf{W}\|_2^2 + \sigma_d^2} \\ \text{s.t.} & \left[P_s \mathbf{W} \mathbf{g} \mathbf{g}^H \mathbf{W}^H + \sigma_r^2 \mathbf{W} \mathbf{W}^H \right]_{i,i} \leq P_r, \\ & \|\mathbf{r}\|_2 = 1. \end{aligned} \quad (1)$$

where we have defined $\mathbf{g} \triangleq \mathbf{H}_{sr} \mathbf{b}$. To further analyze problem (1), we need to introduce Lemma 1, which verifies the rank-one property of the optimal \mathbf{W} .

Lemma 1 (Rank One Condition): The optimal \mathbf{W} in problem (1) is given by $\mathbf{W} = \frac{\tilde{P}_r}{\|\mathbf{g}\|_2} \mathbf{w} \mathbf{g}^H$, for some $\mathbf{w} \triangleq [w_1, \dots, w_N]^T \in \mathbb{C}^N$, where $\tilde{P}_r \triangleq \sqrt{\frac{P_r}{P_s \|\mathbf{g}\|_2^2 + \sigma_r^2}}$.

Proof: See Appendix A. ■

Lemma 1 establishes the fact that the optimal \mathbf{W} in problem (1) is the combination of two equalizers. One is the MRC equalizer \mathbf{g} , and the other one, \mathbf{w} , is to be determined. By Lemma 1, problem (1) becomes

$$\begin{aligned} \max_{\mathbf{w}, \mathbf{r}} \min_{\|\Delta \mathbf{H}\|_F \leq \varepsilon} & \frac{\tilde{P}_r^2 P_s \|\mathbf{g}\|_2^2 |\mathbf{r}^H (\tilde{\mathbf{H}}_{rd} + \Delta \mathbf{H}) \mathbf{w}|^2}{\tilde{P}_r^2 \sigma_r^2 \|\mathbf{r}^H (\tilde{\mathbf{H}}_{rd} + \Delta \mathbf{H}) \mathbf{w}\|_2^2 + \sigma_d^2}, \\ \text{s.t.} & |w_i| \leq 1, 1 \leq i \leq N, \\ & \|\mathbf{r}\|_2 = 1. \end{aligned} \quad (2)$$

Since the received SNR in (2) is an increasing function with respect to $\|\mathbf{g}\|_2$, it can be immediately concluded that the optimal \mathbf{b} is the principal eigenvector of $\mathbf{H}_{sr}^H \mathbf{H}_{sr}$. Similarly, the received SNR is also an increasing function with respect to

$|\mathbf{r}^H(\tilde{\mathbf{H}}_{rd} + \Delta\mathbf{H})\mathbf{w}|^2$, then the optimal (\mathbf{w}, \mathbf{r}) is the solution of the following problem

$$\begin{aligned} \max_{\mathbf{w}, \mathbf{r}} \quad & \min_{\|\Delta\mathbf{H}\|_F \leq \varepsilon} \left| \mathbf{r}^H(\tilde{\mathbf{H}}_{rd} + \Delta\mathbf{H})\mathbf{w} \right|, \\ \text{s.t.} \quad & |w_i| \leq 1, \|\mathbf{r}\|_2 = 1. \end{aligned} \quad (3)$$

According to [25, Lemma 3.1], the optimal value of the inner minimization problem of (3) is given by $\max\{|\mathbf{r}^H\tilde{\mathbf{H}}_{rd}\mathbf{w}| - \varepsilon\|\mathbf{w}\|_2, 0\}$. However, when $\varepsilon\|\mathbf{w}\|_2 < |\mathbf{r}^H\tilde{\mathbf{H}}_{rd}\mathbf{w}|$, the objective function in (3) is zero, which leads to the invalid transmission. Therefore, we aim at optimizing the following problem

$$\begin{aligned} \max_{\mathbf{w}, \mathbf{r}} \quad & |\mathbf{r}^H\tilde{\mathbf{H}}_{rd}\mathbf{w}| - \varepsilon\|\mathbf{w}\|_2 \\ \text{s.t.} \quad & |w_i| \leq 1, \|\mathbf{r}\|_2 = 1. \end{aligned} \quad (4)$$

For any fixed \mathbf{w} , the optimal \mathbf{r} in problem (4) can be directly given by $\mathbf{r} = \frac{\tilde{\mathbf{H}}_{rd}\mathbf{w}}{\|\tilde{\mathbf{H}}_{rd}\mathbf{w}\|_2}$. Substituting this expression into (4), we get

$$\begin{aligned} \max_{\mathbf{w}} \quad & f(\mathbf{w}) \triangleq \|\tilde{\mathbf{H}}_{rd}\mathbf{w}\|_2 - \varepsilon\|\mathbf{w}\|_2. \\ \text{s.t.} \quad & |w_i| \leq 1. \end{aligned} \quad (5)$$

Notice that the original complicated nonconvex problem \mathcal{P}_1 with matrix-valued variables $(\mathbf{W}, \mathbf{b}, \mathbf{r})$, has now been transformed into a much simpler problem, problem (5), with variable \mathbf{w} . The remaining challenge is to determine the optimal solution and the maximum value in (5), denoted as \mathbf{w}° and f° , respectively. Combining the above discussion and Lemma 1, we directly obtain the following theorem.

Theorem 1: The optimal $(\mathbf{W}, \mathbf{b}, \mathbf{r})$ in problem \mathcal{P}_1 are given by

$$\begin{aligned} \mathbf{W} &= \sqrt{\frac{P_r}{\|\mathbf{H}_{sr}\mathbf{b}\|_2^2 (P_s\|\mathbf{H}_{sr}\mathbf{b}\|_2^2 + \sigma_r^2)}} \mathbf{w}^\circ \mathbf{b}^H \mathbf{H}_{sr}^H, \\ \mathbf{b} &= \boldsymbol{\nu}(\mathbf{H}_{sr}^H \mathbf{H}_{sr}), \\ \mathbf{r} &= \frac{\tilde{\mathbf{H}}_{rd}\mathbf{w}^\circ}{\|\tilde{\mathbf{H}}_{rd}\mathbf{w}^\circ\|_2}, \end{aligned}$$

respectively. The optimal received SNR corresponds to

$$\text{SNR} = \frac{\tilde{P}_r^2 P_s \|\mathbf{H}_{sr}\mathbf{b}\|_2^2 \max\{f^\circ, 0\}^2}{\tilde{P}_r^2 \sigma_r^2 \max\{f^\circ, 0\}^2 + \sigma_d^2}. \quad (6)$$

B. A Low-complexity Iterative Algorithm for Solving \mathbf{w}

Due to the non-convex nature of (5), its optimal solution is difficult to obtain in general. Let us define $\boldsymbol{\Upsilon} \triangleq \mathbf{w}\mathbf{w}^H$. By the equation

$$\|\tilde{\mathbf{H}}_{rd}\mathbf{w}\|_2 = \sqrt{\|\tilde{\mathbf{H}}_{rd}\mathbf{w}\|_2^2} = \sqrt{\text{tr}(\tilde{\mathbf{H}}_{rd}^H \tilde{\mathbf{H}}_{rd} \boldsymbol{\Upsilon})},$$

and dropping the rank one constraint, we can transform problem (5) into a relaxed form as

$$\begin{aligned} \max_{\boldsymbol{\Upsilon}, q_1, q_2} \quad & \sqrt{q_1} - \varepsilon\sqrt{q_2} \\ \text{s.t.} \quad & \text{tr}(\tilde{\mathbf{H}}_{rd}^H \tilde{\mathbf{H}}_{rd} \boldsymbol{\Upsilon}) = q_1, \text{tr}(\boldsymbol{\Upsilon}) = q_2, \\ & \text{tr}(\mathbf{E}_i \boldsymbol{\Upsilon}) \leq 1, 1 \leq i \leq N, \\ & \boldsymbol{\Upsilon} \succeq \mathbf{0}, \end{aligned} \quad (7)$$

where $\mathbf{E}_i \in \mathbb{R}^{N \times N}$ denotes a matrix with all zero entries except for the (i, i) th entry which equals to one.

Note that since $\sqrt{q_1}$ and $\sqrt{q_2}$ are concave, the objective function of problem (7) is the difference of two concave functions. Such a problem is recognized as the difference of two convex functions programming (DC programming) problem, which can be efficiently solved via the POLynomial Time DC (POTDC) approach proposed in [26]. The main idea of the POTDC approach is replacing $\sqrt{q_2}$ by its first order Taylor expansion around some point q_c and then solving the resulting convex problem at the κ th iteration, i.e.,

$$\begin{aligned} \max_{\boldsymbol{\Upsilon}, q_1, q_2} \quad & \sqrt{q_1} - \varepsilon(q_2 - q_c)/2\sqrt{q_c} - \varepsilon\sqrt{q_c} \\ \text{s.t.} \quad & \text{tr}(\tilde{\mathbf{H}}_{rd}^H \tilde{\mathbf{H}}_{rd} \boldsymbol{\Upsilon}) = q_1, \text{tr}(\boldsymbol{\Upsilon}) = q_2, \\ & \text{tr}(\mathbf{E}_i \boldsymbol{\Upsilon}) \leq 1, 1 \leq i \leq N, \\ & \boldsymbol{\Upsilon} \succeq \mathbf{0}, 0 \leq q_2 \leq N. \end{aligned} \quad (8)$$

where the constraint $0 \leq q_2 \leq N$ is added since the feasible region of q_2 is a requirement for the POTDC approach.

Problem (8) is recognized as a semidefinite programming (SDP) problem, and thus can be efficiently solved by MATLAB package such as CVX [27]. The optimal q_2 is obtained for updating q_c . Let $\kappa = \kappa + 1$, then one can start the new iteration until some threshold meets. After obtaining the optimal $\boldsymbol{\Upsilon}^\circ$, one can extract \mathbf{w} from the *rank-1 approximation* of $\boldsymbol{\Upsilon}^\circ$. That is: when $\boldsymbol{\Upsilon}^\circ$ is of rank one, then let $\mathbf{w}_{RA} = \sqrt{\lambda_{\max}(\boldsymbol{\Upsilon}^\circ)} \boldsymbol{\nu}(\boldsymbol{\Upsilon}^\circ)$. Otherwise, do the randomization step [28] to get an approximation solution as

$$\tilde{\mathbf{w}}_{RA} \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Upsilon}^\circ), \quad \boldsymbol{\Upsilon}^\circ \approx \tilde{\mathbf{w}}_{RA} \tilde{\mathbf{w}}_{RA}^H.$$

Let $\mathbf{w}_{RA} = \tilde{\mathbf{w}}_{RA}/\beta$, where β is maximum absolute value among all elements of $\tilde{\mathbf{w}}_{RA}$. Then we get an approximation of f° as $f_{RA} = \|\tilde{\mathbf{H}}_{rd}\mathbf{w}_{RA}\|_2 - \varepsilon\|\mathbf{w}_{RA}\|_2$. The complexity of this POTDC-based approach is about $\mathcal{O}(N^7)$ times the number of iterations, which is fairly high.

We will next propose an iterative algorithm to solve \mathbf{w} . The advantage of this algorithm lies in that it has much lower complexity and is easier to implement since only the arithmetic operation rather than the advanced software package is required. From the simulation results in Section VI, one can see that it can reach a solution that is very close to the POTDC-based algorithm.

Notice that problem (5) is equivalent to problem (4). Define the objective function in (4) as $\Phi(\mathbf{w}, \mathbf{r}) \triangleq |\mathbf{r}^H\tilde{\mathbf{H}}_{rd}\mathbf{w}|_2 - \varepsilon\|\mathbf{w}\|_2$. Then we can determine the variable (\mathbf{w}, \mathbf{r}) by the alternating optimization method: Step 1, for any fixed \mathbf{w} , determine the optimal \mathbf{r} as $\mathbf{r} = \frac{\tilde{\mathbf{H}}_{rd}\mathbf{w}}{\|\tilde{\mathbf{H}}_{rd}\mathbf{w}\|_2}$; Step 2, for fixed \mathbf{r} , determine the optimal \mathbf{w} . Repeat step 1-2 until convergence. Inspired by this sense, we will solve problem (4) with fixed \mathbf{r} , that is,

$$\begin{aligned} \max_{\mathbf{w}} \quad & |\mathbf{r}^H\tilde{\mathbf{H}}_{rd}\mathbf{w}| - \varepsilon\|\mathbf{w}\|_2 \\ \text{s.t.} \quad & |w_i| \leq 1. \end{aligned} \quad (9)$$

Note that the phase term of each w_i has no impact on $\|\mathbf{w}\|_2$. Then in order to maximize $|\mathbf{r}^H\tilde{\mathbf{H}}_{rd}\mathbf{w}| - \varepsilon\|\mathbf{w}\|_2$, it must be

equal to the phase term of $\tilde{\mathbf{h}}_i^H \mathbf{r}$. In this case, problem (9) becomes

$$\begin{aligned} \max_{|w_i|} \quad & \sum_{i=1}^N |w_i| |\tilde{\mathbf{h}}_i^H \mathbf{r}| - \varepsilon \sqrt{\sum_{i=1}^N |w_i|^2}, \\ \text{s.t.} \quad & |w_i| \leq 1. \end{aligned} \quad (10)$$

Obviously, problem (10) is convex, and thus can be solve efficiently. By further simplification, we will give an analytical optimal solution of problem (10). Refer to $\alpha_i \triangleq |\tilde{\mathbf{h}}_i^H \mathbf{r}|$ as the antenna gain for the i th antenna, and denote π as a permutation of $\{1, \dots, N\}$ such that $\alpha_{\pi(i)}$ is sorted in a non-decreasing order, i.e., $\alpha_{\pi(1)} \leq \dots \leq \alpha_{\pi(N)}$. Then it can be shown that $|w_{\pi(i)}^\circ|$ is also sorted in a non-decreasing order. Suppose the opposite that $|w_{\pi(i_1)}^\circ| > |w_{\pi(i_2)}^\circ|$, for $i_1 < i_2$. Then the value of (10) can always be increased by swapping the value of $|w_{\pi(i_1)}^\circ|$ and $|w_{\pi(i_2)}^\circ|$, since $\sum_{i=1}^N |w_{\pi(i)}| \alpha_{\pi(i)}$ becomes larger and $\varepsilon \sqrt{\sum_{i=1}^N |w_{\pi(i)}|^2}$ remains unchanged. This contradicts with the assumption that \mathbf{w}° is the optimal solution. Thereby $|w_{\pi(i)}^\circ|$ is also sorted in a non-decreasing order.

To solve (10), let us assume that at least $N - k$ antennas uses the maximum power, that is, $|w_{\pi(k+1)}| = \dots = |w_{\pi(N)}| = 1$. Denote $\mathcal{S}(k) = \{\pi(1), \dots, \pi(k)\}$ as the set of antennas that are free to choose their optimal power consumption. Then problem (10) can be rewritten as

$$\begin{aligned} \max_{k=1, \dots, N} \max_{|w_i|} \quad & \sum_{i \notin \mathcal{S}(k)} \alpha_i + \sum_{i \in \mathcal{S}(k)} |w_i| \alpha_i - \\ & \varepsilon \sqrt{N - k + \sum_{i \in \mathcal{S}(k)} |w_i|^2}, \\ \text{s.t.} \quad & |w_i| \leq 1, i \in \mathcal{S}(k). \end{aligned} \quad (11)$$

Problem (11) has been decomposed into N subproblems. Denote $\mathbf{w}^* \triangleq (w_{\pi(1)}^*, \dots, w_{\pi(k)}^*)$ and $\Omega(k)$ as the optimal solution and optimal value of the k th subproblem,

$$\begin{aligned} \max_{|w_i|} \quad & \sum_{i \notin \mathcal{S}(k)} \alpha_i + \sum_{i \in \mathcal{S}(k)} |w_i| \alpha_i - \varepsilon \sqrt{N - k + \sum_{i \in \mathcal{S}(k)} |w_i|^2} \\ \text{s.t.} \quad & |w_i| \leq 1, i \in \mathcal{S}(k). \end{aligned} \quad (12)$$

respectively. Then when k increases, $\mathcal{S}(k)$ becomes larger, and more antenna are free to choose their optimal power. Therefore, it can be easily verified that $\Omega(k)$ is a non-decreasing function. Moreover, according to the above discussion, it is easy to know that $|w_{\pi(1)}^*| \leq \dots \leq |w_{\pi(k)}^*| \leq 1$.

Problem (12) is a constrained convex problem, whose optimal solution can be uniquely determined by (KKT) conditions, or i.e.,

$$\alpha_i - \varepsilon \frac{|w_i^*|}{\sqrt{N - k + \sum_{i \in \mathcal{S}(k)} |w_i^*|^2}} - \mu_i = 0, \quad (13)$$

$$\mu_i (|w_i^*| - 1) = 0, \quad (14)$$

where μ_i is a dual variable for the i th power constraint. Suppose that $|w_{\pi(k)}^*| = 1$, then there will be $k - 1$ antennas left to choose their optimal power, and the k th subproblem boils down

to the $k - 1$ th subproblem, i.e., $\Omega(k) = \Omega(k - 1)$. Hence in the following, we concentrate on the case when $|w_{\pi(1)}^*| \leq \dots \leq |w_{\pi(k)}^*| < 1$, which results in $\mu_i = 0$ and reduces equality (13) to

$$\alpha_i^2 = \frac{\varepsilon^2 |w_i^*|^2}{N - k + \sum_{i \in \mathcal{S}(k)} |w_i^*|^2}, i \in \mathcal{S}(k). \quad (15)$$

By adding the above equation for $i \in \mathcal{S}(k)$, we get

$$\sum_{i \in \mathcal{S}(k)} \alpha_i^2 = \frac{\varepsilon^2 \sum_{i \in \mathcal{S}(k)} |w_i^*|^2}{N - k + \sum_{i \in \mathcal{S}(k)} |w_i^*|^2},$$

or equivalently,

$$\sum_{i \in \mathcal{S}(k)} |w_i^*|^2 = \frac{(N - k) \sum_{i \in \mathcal{S}(k)} \alpha_i^2}{\varepsilon^2 - \sum_{i \in \mathcal{S}(k)} \alpha_i^2}. \quad (16)$$

Substituting (16) into (15), we have

$$|w_i^*| = \alpha_i \sqrt{\frac{N - k}{\varepsilon^2 - \sum_{i \in \mathcal{S}(k)} \alpha_i^2}}, i \in \mathcal{S}(k). \quad (17)$$

To make sure that $|w_{\pi(k)}^*| < 1$, we must have

$$\varepsilon^2 > \sum_{i \in \mathcal{S}(k)} \alpha_i^2 + \alpha_{\pi(k)}^2 (N - k) \triangleq \chi(k). \quad (18)$$

It is easy to prove that $\chi(k)$ is a non-decreasing function as we show below

$$\begin{aligned} \chi(k+1) - \chi(k) &= \sum_{i \in \mathcal{S}(k+1)} \alpha_i^2 + \alpha_{\pi(k+1)}^2 (N - k - 1) \\ &\quad - \sum_{i \in \mathcal{S}(k)} \alpha_i^2 - \alpha_{\pi(k)}^2 (N - k) \\ &= \alpha_{\pi(k+1)}^2 + \alpha_{\pi(k+1)}^2 (N - k - 1) - \alpha_{\pi(k)}^2 (N - k) \\ &= (N - k) (\alpha_{\pi(k+1)}^2 - \alpha_{\pi(k)}^2) \geq 0. \end{aligned}$$

Denote k° as the maximum number that satisfies (18), or i.e., $\chi(1) \leq \dots \leq \chi(k^\circ) < \varepsilon^2 \leq \chi(k^\circ + 1) \leq \dots \leq \chi(N)$. Then for any $k > k^\circ$, we have $|w_{\pi(k)}^*| = 1$, and hence the k th subproblem of (11) boils down to the k° th subproblem. More specifically, the maximum value of problem (10) corresponds to

$$\begin{aligned} \max_{k=1, \dots, N} \Omega(k) &= \max_{k=1, \dots, k^\circ} \Omega(k) \stackrel{(a)}{=} \Omega(k^\circ) \\ &\stackrel{(b)}{=} \sum_{i \notin \mathcal{S}(k^\circ)} \alpha_i - \sqrt{(N - k^\circ) \left(\varepsilon^2 - \sum_{i \in \mathcal{S}(k^\circ)} \alpha_i^2 \right)}, \end{aligned} \quad (19)$$

where (a) is due to the fact that $\Omega(k)$ is a non-decreasing function, and (b) is obtained by substituting (16) and (17) into (12).

Equation (19) tells us that the maximum value of $\Omega(k)$ is achieved by $k = k^\circ$. Then by (17), the optimal \mathbf{w} in problem

(9) corresponds to $w_i^\circ = |w_i^\circ| e^{j\phi_i}$, where $e^{j\phi_i}$ is the phase term of $\tilde{\mathbf{h}}_i^H \mathbf{r}$, $j = \sqrt{-1}$, and

$$|w_i^\circ| = \begin{cases} \alpha_i \sqrt{\frac{N-k^\circ}{\varepsilon^2 - \sum_{i \in \mathcal{S}(k^\circ)} \alpha_i^2}}, & i \in \mathcal{S}(k^\circ) \\ 1 & i \notin \mathcal{S}(k^\circ) \end{cases} \quad (20)$$

By using (20), we have solved problem (9) in an analytical way. Then we can propose an alternating optimization method for solving problem (4), as formally described in Algorithm 1. The convergence of Algorithm 1 is guaranteed, which is based on the following relationship

$$\Phi(\mathbf{w}^{(\kappa)}, \mathbf{r}^{(\kappa)}) \leq \Phi(\mathbf{w}^{(\kappa+1)}, \mathbf{r}^{(\kappa)}) \leq \Phi(\mathbf{w}^{(\kappa+1)}, \mathbf{r}^{(\kappa+1)}).$$

It follows that each iteration will increase the value of $\Phi(\mathbf{w}, \mathbf{r})$ and obviously $\Phi(\mathbf{w}, \mathbf{r})$ has an upper bound for finite power constraints. Therefore this iterative algorithm must converge.

Algorithm 1

The alternating optimization method for joint solution of (\mathbf{w}, \mathbf{r})

Require

Set $\mathbf{r}^{(0)}$ as a random vector with unit norm and the solution accuracy ϵ . Let $\kappa = 0$ and $\Phi(\mathbf{w}^{(0)}, \mathbf{r}^{(0)}) = 0$.

repeat

- 1) Update $\mathbf{w}^{(\kappa+1)}$ by (20).
- 2) Update $\mathbf{r}^{(\kappa+1)} = \frac{\tilde{\mathbf{H}}_{rd} \mathbf{w}^{(\kappa+1)}}{\|\tilde{\mathbf{H}}_{rd} \mathbf{w}^{(\kappa+1)}\|_2}$.
- 3) Update $\Phi(\mathbf{w}^{(\kappa+1)}, \mathbf{r}^{(\kappa+1)}) = |\mathbf{r}^{(\kappa+1)H} \tilde{\mathbf{H}}_{rd} \mathbf{w}^{(\kappa+1)}| - \varepsilon \|\mathbf{w}^{(\kappa+1)}\|_2$.
- 4) Let $\kappa = \kappa + 1$.

until $|\Phi(\mathbf{w}^{(\kappa)}, \mathbf{r}^{(\kappa)}) - \Phi(\mathbf{w}^{(\kappa-1)}, \mathbf{r}^{(\kappa-1)})| \leq \epsilon$

return $\mathbf{w}_{AO} = \mathbf{w}^{(\kappa)}$, $\mathbf{r}_{AO} = \mathbf{r}^{(\kappa)}$ and $f_{AO} = \Phi(\mathbf{w}^{(\kappa)}, \mathbf{r}^{(\kappa)})$.

From (18), one can see that k° is related to ε , which makes $\Omega(k^\circ)$ an implicit function with respect to ε . Specifically, when $\varepsilon^2 > \chi(N) = \sum_{i=1}^N |\tilde{\mathbf{h}}_i^H \mathbf{r}|^2$, we have $|w_{\pi(N)}| < 1$, then $k^\circ = N$, which results in $\Omega(k^\circ) = 0$. In this case, the transmission is declared to be invalid. To further investigate the feasible variation range of ε for the valid transmission, we give the following proposition.

Proposition 1: For any fixed \mathbf{r} , the necessary and sufficient condition for $\Omega(k^\circ) > 0$ is

$$\varepsilon < \sqrt{\chi(N)} \leq \sqrt{\lambda_{\max}(\tilde{\mathbf{H}}_{rd} \tilde{\mathbf{H}}_{rd}^H)}. \quad (21)$$

Proof: See Appendix B. ■

Remark 1: Generally speaking, Algorithm 1 may find a local optimal solution, which heavily depends on the initial point. Since problem (4) is not convex, there is no guarantee that this local optimal solution coincides with the global optimum. According to Proposition 1, to satisfy (21), we can choose the initial point $\mathbf{r}^{(0)}$ as the principal eigenvector of $\tilde{\mathbf{H}}_{rd} \tilde{\mathbf{H}}_{rd}^H$. For practical use, we also apply Algorithm 1 several times to find a better solution, with different initial point $\mathbf{r}^{(0)}$ each time. Simulation

results shows that Algorithm 1 performs almost as well as the POTDC-based algorithm.

Remark 2: Consider the case when ε is very small, i.e., $\varepsilon^2 \leq \chi(1)$, we have $|w_{\pi(1)}^\circ| = \dots = |w_{\pi(N)}^\circ| = 1$. Algorithm 1 has a simplified implementation, with \mathbf{w} updated by $\mathbf{w}^\circ \triangleq \mathbf{p}(\mathbf{r})$, where $\mathbf{p}(\mathbf{r}) \triangleq [e^{j\phi_1}, \dots, e^{j\phi_N}]^T$. In this case, the relay BF matrix design is given by

$$\mathbf{W} = \sqrt{\frac{P_r}{\|\mathbf{H}_{sr} \mathbf{b}\|_2^2 (P_s \|\mathbf{H}_{sr} \mathbf{b}\|_2^2 + \sigma_r^2)}} \mathbf{p}(\mathbf{r}) \mathbf{b}^H \mathbf{H}_{sr}^H, \quad (22)$$

which is the combination of an MRC equalizer and an equal-gain-transmission (EGT) equalizer. Let this scheme be called as *Equal power design*. Theoretically, it is only optimal for $\varepsilon^2 \leq \chi(1)$. However, we will later see in the simulation part that, for relatively large range of ε , this equal power design achieves a performance comparable to that of the POTDC-based algorithm. This can be explained by the expression of the received SNR in (6), whose value mainly depends on f° . Notice that by (5), f° is only related to the channel coefficient matrix $\tilde{\mathbf{H}}_{rd}$ rather than P_r or P_s , hence the gap between different approximations of f° obtained by difference methods is not so obvious for small ε . Therefore, as long as ε is not too large, the equal power design can be applied as a good alternative, where each antenna at the relay node uses the same power, which greatly relaxes the power amplifier design, and is exactly the objective of this work.

IV. POWER MINIMIZATION PROBLEM

In this section, we consider another criterion. Our goal is to minimize the maximum per-antenna power for a given SNR requirement, as formulated in problem \mathcal{P}_2 . We will reveal the relationship between problem \mathcal{P}_1 and problem \mathcal{P}_2 , showing that they reduce to the same problem, and thus can be solved by the same method.

By fixing the source BF vector \mathbf{b} and following the similar lines in Lemma 1, one can prove that the optimal \mathbf{W} in problem \mathcal{P}_2 is of rank one, given by $\mathbf{W} = \frac{\tilde{P}_r}{\|\mathbf{g}\|_2} \mathbf{w} \mathbf{g}^H$, with \mathbf{w} to be determined. Then problem \mathcal{P}_2 can be simplified into

$$\begin{aligned} \min_{\mathbf{w}, \mathbf{r}} \quad & P_r \\ \text{s.t.} \quad & \min_{\|\Delta \mathbf{H}\|_F \leq \varepsilon} P_r \left| \mathbf{r}^H (\tilde{\mathbf{H}}_{rd} + \Delta \mathbf{H}) \mathbf{w} \right|^2 \geq \bar{\gamma}_0, \\ & |w_i| \leq 1, \|\mathbf{r}\|_2 = 1. \end{aligned} \quad (23)$$

where $\bar{\gamma}_0 \triangleq \frac{\gamma_0 \sigma_s^2 (P_s \|\mathbf{g}\|_2^2 + \sigma_r^2)}{P_s \|\mathbf{g}\|_2^2 - \gamma_0 \sigma_r^2}$. Problem (23) can be transformed into

$$\begin{aligned} \min_{\mathbf{w}, \mathbf{r}} \quad & P_r \\ \text{s.t.} \quad & P_r \geq \frac{\bar{\gamma}_0}{\min_{\|\Delta \mathbf{H}\|_F \leq \varepsilon} \left| \mathbf{r}^H (\tilde{\mathbf{H}}_{rd} + \Delta \mathbf{H}) \mathbf{w} \right|^2}, \\ & |w_i| \leq 1, \|\mathbf{r}\|_2 = 1, \end{aligned} \quad (24)$$

If we treat P_r as a slack variable, then it is obvious that problem (24) is exactly equivalent to problem (3). This shows that the power minimization problem and the SNR maximization problem actually lead to the same problem. Letting

$P_r \max\{f^\circ, 0\}^2 = \bar{\gamma}_0$, we directly obtain the expression of P_r in (25). We summarize the above result in the following theorem.

Theorem 2: The optimal $(\mathbf{W}, \mathbf{b}, \mathbf{r})$ in problem \mathcal{P}_2 are given by

$$\begin{aligned} \mathbf{W} &= \frac{\mathbf{w}^\circ \mathbf{b}^H \mathbf{H}_{sr}^H}{\|\mathbf{H}_{sr} \mathbf{b}\|_2 \max\{f^\circ, 0\}} \sqrt{\frac{\gamma_0 \sigma_d^2}{P_s \|\mathbf{H}_{sr} \mathbf{b}\|_2^2 - \gamma_0 \sigma_r^2}}, \\ \mathbf{b} &= \nu \left(\mathbf{H}_{sr}^H \mathbf{H}_{sr} \right), \\ \mathbf{r} &= \frac{\tilde{\mathbf{H}}_{rd} \mathbf{w}^\circ}{\|\tilde{\mathbf{H}}_{rd} \mathbf{w}^\circ\|_2}, \end{aligned}$$

with the corresponding power given by

$$P_r = \frac{\gamma_0 \sigma_d^2 (P_s \|\mathbf{H}_{sr} \mathbf{b}\|_2^2 + \sigma_r^2)}{(P_s \|\mathbf{H}_{sr} \mathbf{b}\|_2^2 - \gamma_0 \sigma_r^2) \max\{f^\circ, 0\}^2}, \quad (25)$$

where \mathbf{w}° and f° are the optimal solution and the maximum value in (4), respectively.

Remark 3: According to Theorem 2, to ensure that P_r is finite, there must be $f^\circ > 0$ and $\gamma_0 < P_s \|\mathbf{H}_{sr} \mathbf{b}\|_2^2 / \sigma_r^2 = P_s \lambda_{\max}(\mathbf{H}_{sr}^H \mathbf{H}_{sr}) / \sigma_r^2$. Therefore, we claim that the necessary and sufficient *feasibility condition* of problem \mathcal{P}_2 is given by constraint (21) and $\gamma_0 < P_s \lambda_{\max}(\mathbf{H}_{sr}^H \mathbf{H}_{sr}) / \sigma_r^2$. Note that the first condition requires that the error bound should not be too large and the second condition means that the received SNR target at the destination cannot be larger than the received SNR at the relay.

V. COMPARISON WITH THE WORK IN [11]

When $\Delta \mathbf{H} = \mathbf{0}$, problem \mathcal{P}_1 and problem \mathcal{P}_2 were also discussed in [11]. The authors in [11] used an alternating optimization approach for the joint design of $(\mathbf{W}, \mathbf{b}, \mathbf{r})$. By fixing \mathbf{b} and \mathbf{r} , they obtained the optimal \mathbf{W} by solving an SDP problem, with complexity given by $\mathcal{O}(N^6)$ times the number of iterations (typically lies between 5 and 50 [29]). While fixing \mathbf{W} , the optimal \mathbf{b} and \mathbf{r} were given as the closed-form solution. This process is repeated until convergence. In contrast, in our work, we first determine the optimal \mathbf{b} as the closed-form solution and the optimal \mathbf{W} as a function of \mathbf{w} , and then propose Algorithm 1 for the joint solution of (\mathbf{w}, \mathbf{r}) , with complexity given by $\mathcal{O}(N \log_2 N)$ times the iteration number (as shown in Fig. 2). Once \mathbf{w} is obtained, \mathbf{W} can be directly determined by Lemma 1. Obviously, our work has much lower complexity and is much easier to implement since only the arithmetic operation rather than the advanced software package such as CVX is required. The average CPU time comparison is further illustrated in Fig. 3.

We will next show that our method and that in [11] have the same performance. Consider the κ th iteration in [11], by fixing $\mathbf{b}^{(\kappa)}$ and $\mathbf{r}^{(\kappa)}$, the optimal $\mathbf{W}^{(\kappa)}$ is obtained by the numerical results, which can be equivalently written as the closed-form expression in (22). That is,

$$\begin{aligned} \mathbf{W}^{(\kappa)} &= \sqrt{\frac{P_r}{\|\mathbf{H}_{sr} \mathbf{b}^{(\kappa)}\|_2^2 (P_s \|\mathbf{H}_{sr} \mathbf{b}^{(\kappa)}\|_2^2 + \sigma_r^2)}} \\ &\quad \times \mathbf{p} \left(\mathbf{r}^{(\kappa)} \right) \mathbf{b}^{(\kappa)H} \mathbf{H}_{sr}^H. \end{aligned}$$

Then the next step is to update $\mathbf{b}^{(\kappa+1)}$ and $\mathbf{r}^{(\kappa+1)}$. Substituting $\mathbf{W}^{(\kappa)}$ into problem \mathcal{P}_1 , it results in¹

$$\begin{aligned} &\max_{\|\mathbf{b}\|_2=1, \|\mathbf{r}\|_2=1} \\ &\frac{P_r P_s \|\mathbf{b}^{(\kappa)H} \mathbf{H}_{sr}^H \mathbf{H}_{sr} \mathbf{b}\|_2^2 |\mathbf{r}^H \tilde{\mathbf{H}}_{rd} \mathbf{p}(\mathbf{r}^{(\kappa)})|^2}{P_r \sigma_r^2 |\mathbf{r}^H \tilde{\mathbf{H}}_{rd} \mathbf{p}(\mathbf{r}^{(\kappa)})|^2 + \sigma_d^2 \|\mathbf{H}_{sr} \mathbf{b}^{(\kappa)}\|_2^2 (P_s \|\mathbf{H}_{sr} \mathbf{b}^{(\kappa)}\|_2^2 + \sigma_r^2)}. \end{aligned} \quad (26)$$

By (26), to optimize \mathbf{b} , one must maximize $\|\mathbf{b}^{(\kappa)H} \mathbf{H}_{sr}^H \mathbf{H}_{sr} \mathbf{b}\|_2^2$. Thus $\mathbf{b}^{(\kappa+1)}$ is updated as

$$\mathbf{b}^{(\kappa+1)} = \mathbf{H}_{sr}^H \mathbf{H}_{sr} \mathbf{b}^{(\kappa)} / \left\| \mathbf{H}_{sr}^H \mathbf{H}_{sr} \mathbf{b}^{(\kappa)} \right\|_2. \quad (27)$$

which achieves the optimal value as $\|\mathbf{b}^{(\kappa)H} \mathbf{H}_{sr}^H \mathbf{H}_{sr} \mathbf{b}\|_2^2$. On the other hand, to optimize \mathbf{r} , one must maximize $|\mathbf{r}^H \tilde{\mathbf{H}}_{rd} \mathbf{p}(\mathbf{r}^{(\kappa)})|$. Update

$$\mathbf{r}^{(\kappa+1)} = \frac{\tilde{\mathbf{H}}_{rd} \mathbf{p}(\mathbf{r}^{(\kappa)})}{\left\| \tilde{\mathbf{H}}_{rd} \mathbf{p}(\mathbf{r}^{(\kappa)}) \right\|_2}. \quad (28)$$

After that, repeat the above process in the next iteration. In summary, for each iteration, one only has to compute (27) and (28).

The iteration process will stop if it satisfies that

$$\left| \left\| \mathbf{b}^{(\kappa+1)H} \mathbf{H}_{sr}^H \mathbf{H}_{sr} \mathbf{b} \right\|_2^2 - \left\| \mathbf{b}^{(\kappa)H} \mathbf{H}_{sr}^H \mathbf{H}_{sr} \mathbf{b} \right\|_2^2 \right| \leq \epsilon, \quad (29)$$

$$\left| \left| \mathbf{r}^{(\kappa+1)H} \tilde{\mathbf{H}}_{rd} \mathbf{p}(\mathbf{r}^{(\kappa)}) \right| - \left| \mathbf{r}^{(\kappa)H} \tilde{\mathbf{H}}_{rd} \mathbf{p}(\mathbf{r}^{(\kappa-1)}) \right| \right| \leq \epsilon. \quad (30)$$

where ϵ is the solution accuracy.

From (29), it is easy to know that the optimal \mathbf{b} is given by $\arg \max_{\|\mathbf{b}\|_2=1} \|\mathbf{b}^H \mathbf{H}_{sr}^H \mathbf{H}_{sr} \mathbf{b}\|_2^2$, i.e., the principal eigenvector of $\mathbf{H}_{sr}^H \mathbf{H}_{sr}$. Meanwhile, (30) is equivalent to the stopping criterion in Algorithm 1 when $\epsilon = 0$. Then one can see that the method in [11] and that in our work have the same performance.

By the relationship between the SNR maximization problem \mathcal{P}_1 and power minimization problem \mathcal{P}_2 , it can be verified in the similar way that for problem \mathcal{P}_2 , the method in [11] has the same performance as that of our method.

VI. SIMULATIONS RESULTS

In this section, we provide numerical results to validate the proposed robust design in this paper. First, the convergence behavior and the required CPU time of Algorithm 1 is illustrated. Then the performance evaluation of our robust design is addressed. The parameters are set as $M_s = M_d = N$. The channel fading is modeled as Rayleigh fading, with each entries of \mathbf{H}_{sr} and $\tilde{\mathbf{H}}_{rd}$ satisfying $\mathcal{CN}(0, 1)$, and the noise variance parameter is set as $\sigma_d^2 = \sigma_r^2 = 1$. We set the power consumed at the source as 20 dBW and the given SNR target γ_0 as 15 dB.

¹We don't impose the relay per-antenna power constraints here when optimizing \mathbf{b} and \mathbf{r} for two reasons: First, the source and destination only have to maximize the received SNR from their own perspectives; more importantly, it can be proven that it results in $\mathbf{b}^{(\kappa+1)} = \mathbf{b}^{(\kappa)}$ by considering the power constraint.

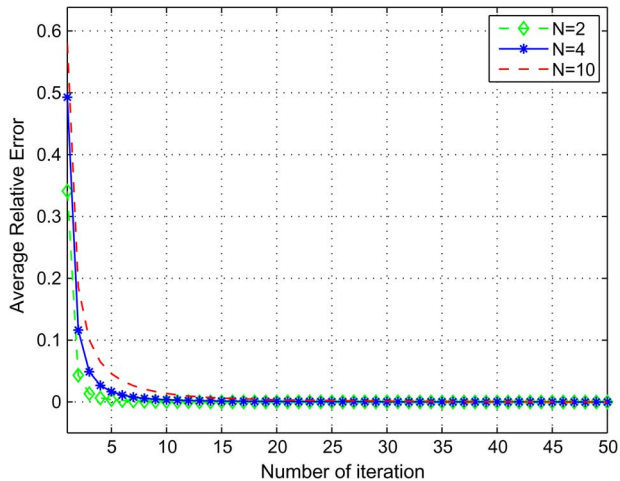


Fig. 2. Relative error to the optimal value of Algorithm 1 versus the number of iterations.

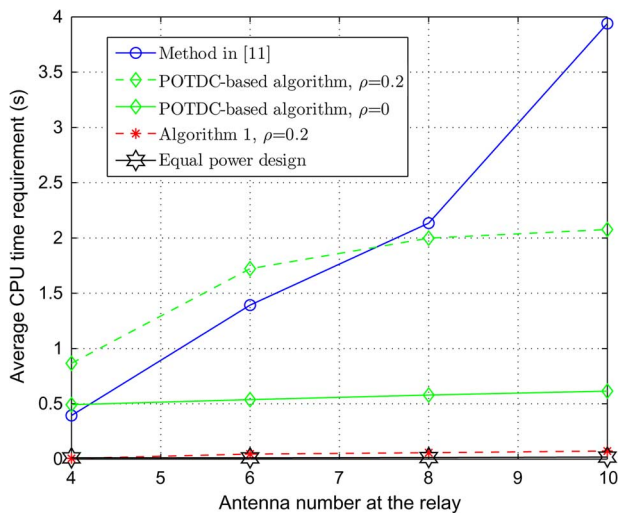


Fig. 3. Average CPU time comparison versus different N .

By Proposition 1, to make sure that each transmission is valid, the maximum value of ε can not exceed $\sqrt{\lambda_{\max}(\tilde{\mathbf{H}}_{rd}\tilde{\mathbf{H}}_{rd}^H)}$. Thereby we vary ε through the normalized parameter ρ , i.e., $\varepsilon^2 = \rho\lambda_{\max}(\tilde{\mathbf{H}}_{rd}\tilde{\mathbf{H}}_{rd}^H)$ with $\rho \in [0, 1)$. Then the larger the ρ is, the poorer the CSI quality will be. All results are averaged over 1000 channel realizations.

First, we study the convergence performance of Algorithm 1. We set $N \in \{2, 4, 10\}$. Fig. 2 shows the average iteration numbers to achieve some certain accuracies, which is defined as $(f_{AO} - \Phi(\mathbf{w}^{(\kappa)}, \mathbf{r}^{(\kappa)}))/f_{AO}$. It can be observed that Algorithm 1 converges quickly to the optimal value in about 5 to 20 iterations. Remembering that the complexity of Algorithm 1 is only $\mathcal{O}(N \log_2 N)$ in each iteration, we can claim that our proposed algorithm has quite low complexity.

By Theorem 1, we have expressed the optimal BF design as a function of \mathbf{w} . Hence in Fig. 3, we provide the CPU time comparison for the computation of \mathbf{w} by different methods: the POTDC-based algorithm in Section III-B, Algorithm 1 and the equal power design in remark 2. We also plot the time cost of the method in [11] as a benchmark. Notice that equal power

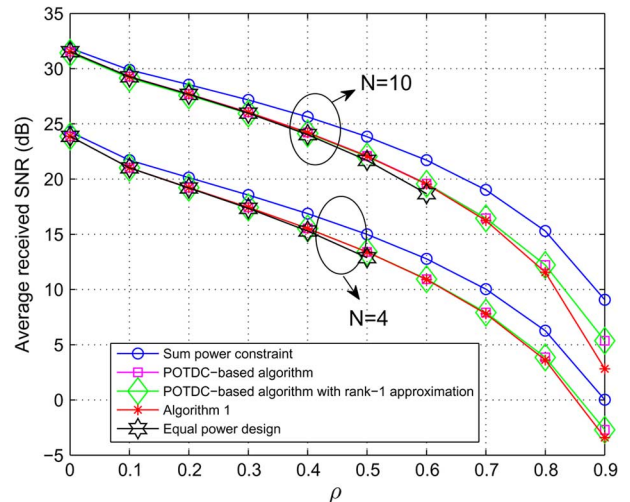


Fig. 4. Average worst case received SNR at the relay versus ρ .

design is just the special case of our proposed method when $\rho = 0$. It can be observed from Fig. 3 that the equal power design and our proposed robust method take up nearly the same time, which are much smaller than that in other methods. This is reasonable since Algorithm 1 adopts the analytical (or closed-form) solution in each iteration, while the other two methods only obtain the numerical results by solving the SDP problem.

We now presents some performance evaluation results to compare our robust BF design with other schemes: a) *Sum power constraint*: This is the robust design undersum power constraint [5], [15]. b) *POTDC-based algorithm*: This method solves problem (7) by the POTDC-based algorithm, and computes the corresponding received SNR by Theorem 1. c) *POTDC-based algorithm with rank-1 approximation*: This method computes the received SNR by replacing f° with f_{RA} . c) *Equal power design*: This method is also the nonrobust design, since it does not take the channel uncertainty error into account. The performance of the method in [11] is not illustrated here, as we have theoretically proved in Section V that it has the same performance as the *Equal power design*.

In Fig. 4, we address the relationship between the average worst-case received SNR and ρ . It can be seen that the system performance is deteriorated by the uncertainty error. The larger the ρ is, the smaller SNR value will be. Another observation is that the design under sum power constraints results in higher SNR than that in per-antenna power constraints, which is due to the flexibility of power distribution among antennas. Furthermore, for small to moderate ρ , equal power design has almost the same performance as the POTDC-based algorithm. Remembering that the equal power design quite simplifies the engineering designs, we claim that it is a competitive alternative for a simple relay node (inexpensive power amplifier). When ρ is large, the equal power design would lead to invalid transmission, while our proposed robust method still behaves well and preserves optimality to some extent.

In Fig. 5, we further plot the average worst-case received SNR versus the per-antenna power at the relay for $N = 10$. Simulations reveal that when $\rho = 0.2$, both the proposed robust method and the equal power design behaves as well as the

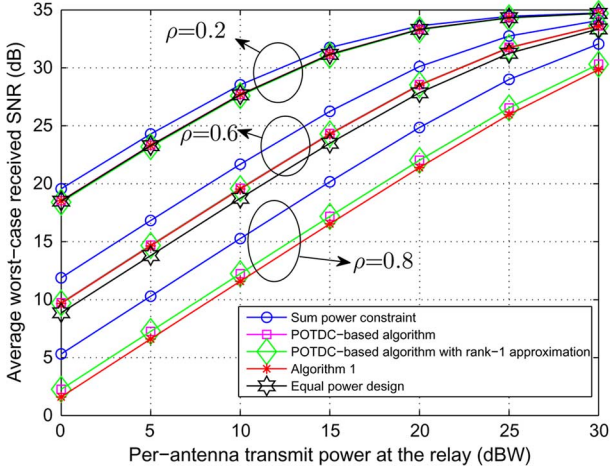


Fig. 5. Average worst case received SNR versus transmit power at each antenna of the relay when $N = 10$.

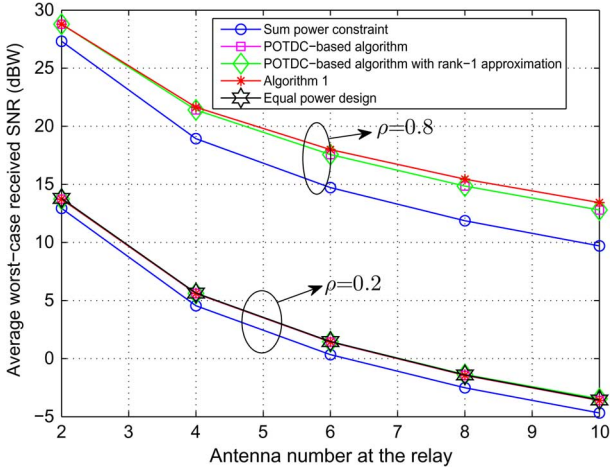


Fig. 6. Average per antenna power at the relay versus relay antenna number N .

POTDC-based algorithm. This gives us a hint that when the quality of the CSI is not too bad, equal power design is most simple way to achieve a pretty good performance. On the other hand, when ρ increases to 0.8, i.e., the CSI equality is very poor, equal power design could not be applied any more, while Algorithm 1 has a small performance gap to that of the POTDC-based algorithm. Considering the fact that the POTDC-based algorithm has much higher time cost than that of the two formers, it can be concluded that Equal power design (for small to moderate ρ) and Algorithm 1 (for large ρ) are more efficient for practical use.

Fig. 6 shows the average required per-antenna power with given γ_0 for different values of N . We can see that when ρ becomes larger, the required power usage at each antenna increases. This phenomenon reveals that when the channel CSI becomes poorer, the relay needs more power to guarantee the SNR target in the worst case, which is consistent with the intuition. The conclusion is similar from that we obtained in Fig. 4 and Fig. 5, which is due to the fact that the power minimization problem can be converted into the SNR maximization problem.

VII. CONCLUSION

In this paper, we consider an AF multi-antenna relay network with one source and one destination. Assuming that the relay only has imperfect CSI, we have derived a semi-closed-form solution for the joint optimal BF design. Then we propose a low-complexity iterative algorithm for obtaining the remaining unknown variable. We also indicate that both the SNR maximization problem and the power minimization problem can be ascribed into the same problem. Compared to the existing methods, our solution leads to drastic complexity reduction for solving joint source-relay-destination BF design.

APPENDIX I

PROOF OF LEMMA 1

Suppose that the singular value decomposition (SVD) of \mathbf{g} is

$$\mathbf{g} = \mathbf{U} \begin{bmatrix} \|\mathbf{g}\|_2 \\ \mathbf{0}_{N-1} \end{bmatrix} \triangleq \mathbf{U}\boldsymbol{\Sigma}, \quad (31)$$

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

$$\mathbf{W} = \mathbf{Y}\mathbf{U}^H \quad (32)$$

where $\mathbf{Y} \in \mathbb{C}^{N \times N}$ is a matrix to be determined. Upon substituting (32) and (31) into problem (1), we have

$$\begin{aligned} \max_{\mathbf{Y}, \mathbf{r}} \quad & \min_{\|\Delta \mathbf{H}\|_F \leq \epsilon} \frac{P_s \|\mathbf{r}^H (\tilde{\mathbf{H}}_{rd} + \Delta \mathbf{H}) \mathbf{Y} \boldsymbol{\Sigma}\|_2^2}{\sigma_r^2 \|\mathbf{r}^H (\tilde{\mathbf{H}}_{rd} + \Delta \mathbf{H}) \mathbf{Y}\|_2^2 + \sigma_d^2} \\ \text{s.t.} \quad & [\mathbf{Y}(P_s \boldsymbol{\Sigma} \boldsymbol{\Sigma}^H + \sigma_r^2) \mathbf{Y}^H]_{i,i} \leq P_r. \end{aligned} \quad (33)$$

We can further partition \mathbf{Y} as

$$\mathbf{Y} = [\tilde{\mathbf{w}} \quad \mathbf{Z}_y], \quad (34)$$

where $\tilde{\mathbf{w}} \triangleq [\tilde{w}_1, \dots, \tilde{w}_N]^T$ and $\mathbf{Z}_y \in \mathbb{C}^{N \times (N-1)}$.

Then we have

$$\mathbf{Y} \boldsymbol{\Sigma} = [\tilde{\mathbf{w}} \quad \mathbf{Z}_y] \begin{bmatrix} \|\mathbf{g}\|_2 \\ \mathbf{0} \end{bmatrix} = \|\mathbf{g}\|_2 \tilde{\mathbf{w}}. \quad (35)$$

Upon substituting (34) and (35) into (33), we have the received SNR at the destination as

SNR

$$\begin{aligned} &= \frac{P_s \|\mathbf{g}\|_2^2 |\mathbf{r}^H (\tilde{\mathbf{H}}_{rd} + \Delta \mathbf{H}) \tilde{\mathbf{w}}|^2}{\sigma_r^2 \|\mathbf{r}^H (\tilde{\mathbf{H}}_{rd} + \Delta \mathbf{H}) \mathbf{Y}\|_2^2 + \sigma_d^2} \\ &= \frac{P_s \|\mathbf{g}\|_2^2 |\mathbf{r}^H (\tilde{\mathbf{H}}_{rd} + \Delta \mathbf{H}) \tilde{\mathbf{w}}|^2}{\sigma_r^2 (|\mathbf{r}^H (\tilde{\mathbf{H}}_{rd} + \Delta \mathbf{H}) \tilde{\mathbf{w}}|_2^2 + \|\mathbf{r}^H (\tilde{\mathbf{H}}_{rd} + \Delta \mathbf{H}) \mathbf{Z}_y\|_2^2) + \sigma_d^2} \end{aligned}$$

and the per-antenna power becomes

$$\begin{aligned} &[\mathbf{Y}(P_s \boldsymbol{\Sigma} \boldsymbol{\Sigma}^H + \sigma_r^2) \mathbf{Y}^H]_{i,i} \\ &= [\tilde{\mathbf{w}}(P_s \|\mathbf{g}\|_2^2 + \sigma_r^2) \tilde{\mathbf{w}}^H]_{i,i} + \sigma_r^2 [\mathbf{Z}_y \mathbf{Z}_y^H]_{i,i} \\ &= (P_s \|\mathbf{g}\|_2^2 + \sigma_r^2) |\tilde{w}_i|^2 + \sigma_r^2 [\mathbf{Z}_y \mathbf{Z}_y^H]_{i,i}. \end{aligned}$$

Then problem (33) becomes

$$\begin{aligned} & \max_{\tilde{\mathbf{w}}, \mathbf{r}} \min_{\|\Delta \mathbf{H}\|_F \leq \varepsilon} \\ & \frac{P_s \|\mathbf{g}\|_2^2 |\mathbf{r}^H (\tilde{\mathbf{H}}_{rd} + \Delta \mathbf{H}) \tilde{\mathbf{w}}|^2}{\sigma_r^2 \left(|\mathbf{r}^H (\tilde{\mathbf{H}}_{rd} + \Delta \mathbf{H}) \tilde{\mathbf{w}}|_2^2 + \|\mathbf{r}^H (\tilde{\mathbf{H}}_{rd} + \Delta \mathbf{H}) \mathbf{Z}_y\|_2^2 \right) + \sigma_d^2}, \\ & \text{s.t. } (P_s \|\mathbf{g}\|_2^2 + \sigma_r^2) |\tilde{w}_i|^2 + \sigma_r^2 [\mathbf{Z}_y \mathbf{Z}_y^H]_{i,i} \leq P_r. \end{aligned} \quad (36)$$

For any feasible $\mathbf{Y} = [\tilde{\mathbf{w}} \ \mathbf{Z}_y]$ with $\mathbf{Z}_y \neq \mathbf{0}$, one can always find $\mathbf{Y}' = [\tilde{\mathbf{w}} \ \mathbf{0}]$ which can achieve a larger SNR. Thus we conclude that there must be $\mathbf{Z}_y = \mathbf{0}$. Denote $\mathbf{w} = \frac{1}{\tilde{P}_r} \tilde{\mathbf{w}}$. We can express problem (36) as

$$\begin{aligned} & \max_{\mathbf{w}, \mathbf{r}} \min_{\|\Delta \mathbf{H}\|_F \leq \varepsilon} \frac{\tilde{P}_r^2 P_s \|\mathbf{g}\|_2^2 |\mathbf{r}^H (\tilde{\mathbf{H}}_{rd} + \Delta \mathbf{H}) \mathbf{w}|^2}{\tilde{P}_r^2 \sigma_r^2 \left(|\mathbf{r}^H (\tilde{\mathbf{H}}_{rd} + \Delta \mathbf{H}) \mathbf{w}|_2^2 + \sigma_d^2 \right)}, \\ & \text{s.t. } |w_i| \leq 1. \end{aligned} \quad (37)$$

Then we can express \mathbf{W} as

$$\mathbf{W} = \tilde{P}_r \mathbf{w} (\mathbf{U})_1^H = \frac{\tilde{P}_r}{\|\mathbf{g}\|_2} \mathbf{w} \mathbf{g}^H,$$

where $(\mathbf{U})_1$ denotes the first column of \mathbf{U} . The proof is completed.

APPENDIX II

PROOF OF PROPOSITION 1

In this appendix, we will first show that $\Omega(k^\circ)$ is a decreasing function with respect to ε . Then we will give necessary and sufficient conditions for $\Omega(k^\circ) > 0$. To investigate the dynamic change of $\Omega(k^\circ)$ in terms of ε , we rewrite k° as $k^\circ = k^\circ(\varepsilon)$, making the dependence of k° on ε explicitly. For $\varepsilon^2 > \chi(N)$, we have $k^\circ = N$, resulting in $\Omega(k^\circ) = 0$. While for $\varepsilon = 0$, it is easy to show that $\Omega(k^\circ) > 0$. Then for any nonzero ε , the feasible region of ε is given by $(0, \sqrt{\chi(N)}]$, which can be decomposed into N subregions

$$(0, \sqrt{\chi(N)}] = (0, \sqrt{\chi(1)}] \cup \dots \cup (\sqrt{N-1}, \sqrt{\chi(N)}].$$

Denote $\Gamma_t \triangleq (\sqrt{t}, \sqrt{t+1}]$, for $0 \leq t \leq N-1$. For some $0 \leq t_1 \leq N-1$, consider some $\varepsilon_1 \in \Gamma_{t_1}$. Then for any $\varepsilon_2 > \varepsilon_1$, if $\varepsilon_2 \in \Gamma_{t_1}$, we have $k^\circ(\varepsilon_1) = k^\circ(\varepsilon_2) = t_1$. By (19), it can be easily verified that $\Omega(k^\circ(\varepsilon_1)) > \Omega(k^\circ(\varepsilon_2))$.

Now suppose that $\varepsilon_2 \in \Gamma_{t_1+1}$, for $0 \leq t_1+1 \leq N-1$, or i.e., $0 \leq t_1 \leq N-2$. Then $k^\circ(\varepsilon_2) = t_1+1$. In this case, we have

$$\begin{aligned} & \Omega(k^\circ(\varepsilon_1)) - \Omega(k^\circ(\varepsilon_2)) \\ & = \alpha_{\pi(t_1+1)} - \sqrt{(N-t_1) \left(\varepsilon_1^2 - \sum_{i \in \mathcal{S}(t_1)} \alpha_i^2 \right)} \\ & \quad + \sqrt{(N-t_1-1) \left(\varepsilon_2^2 - \sum_{i \in \mathcal{S}(t_1+1)} \alpha_i^2 \right)}. \end{aligned}$$

We aim to show that $\Omega(k^\circ(\varepsilon_1)) - \Omega(k^\circ(\varepsilon_2)) > 0$, which is equivalent to

$$\begin{aligned} & \alpha_{\pi(t_1+1)} + \sqrt{(N-t_1-1) \left(\varepsilon_2^2 - \sum_{i \in \mathcal{S}(t_1+1)} \alpha_i^2 \right)} \\ & > \sqrt{(N-t_1) \left(\varepsilon_1^2 - \sum_{i \in \mathcal{S}(t_1)} \alpha_i^2 \right)}. \end{aligned} \quad (38)$$

Square (38) on both sides, after some manipulations, we get

$$\begin{aligned} & (N-t_1)(\varepsilon_2^2 - \varepsilon_1^2) \\ & - \left(\sqrt{\varepsilon_2^2 - \sum_{i \in \mathcal{S}(t_1+1)} \alpha_i^2} - \alpha_{\pi(t_1+1)} \sqrt{N-t_1-1} \right)^2 > 0, \end{aligned}$$

or equivalently

$$(N-t_1)(\varepsilon_2^2 - \varepsilon_1^2) > (\sqrt{m} - \sqrt{n})^2, \quad (39)$$

where we have defined $m \triangleq \varepsilon_2^2 - \sum_{i \in \mathcal{S}(t_1+1)} \alpha_i^2$ and $n \triangleq (N-t_1-1)\alpha_{\pi(t_1+1)}^2$. Note that $m > n$ since $\varepsilon_2^2 > \chi(t_1+1)$. By the assumption $\varepsilon_1 \in \Gamma_{t_1}$, we have $\chi(t_1) < \varepsilon_1^2 \leq \chi(t_1+1)$, or i.e.,

$$\begin{aligned} \varepsilon_1^2 & \leq \sum_{i \in \mathcal{S}(t_1+1)} \alpha_i^2 + (N-t_1-1)\alpha_{\pi(t_1+1)}^2 \\ & = \varepsilon_2^2 - (m-n). \end{aligned} \quad (40)$$

By (40), if we can show that the following inequality holds, then (39) is proved.

$$(N-t_1)(m-n) > (\sqrt{m} - \sqrt{n})^2, \quad (41)$$

or equivalently,

$$(N-t_1)(\sqrt{m} + \sqrt{n}) \geq (\sqrt{m} - \sqrt{n}). \quad (42)$$

Notice that (42) always holds for $t_1 \leq N-2$. Following the similar lines, it can be verified that for any $\varepsilon_{i+1} \in \Gamma_{t_1+i}$, we have

$$\Omega(k^\circ(\varepsilon_{i+1})) < \dots < \Omega(k^\circ(\varepsilon_2)) < \Omega(k^\circ(\varepsilon_1)). \quad (43)$$

From (43), it can be concluded that $\Omega(k^\circ(\varepsilon))$ is a decreasing function with respect to ε . Hence the minimal value of $\Omega(k^\circ(\varepsilon))$ is given

$$\Omega(k^\circ(\sqrt{\chi(N)})) = \alpha_{\pi(N)} - \sqrt{\sum_{i=1}^N \alpha_i^2 - \sum_{i \in \mathcal{S}(N-1)} \alpha_i^2} = 0.$$

Hence for $\varepsilon^2 < \chi(N)$, we have $\Omega(k^\circ(\varepsilon)) > \Omega(k^\circ(\sqrt{\chi(N)})) = 0$, and vice versa.

Notice that $\chi(N) = \sum_{i=1}^N |\tilde{\mathbf{h}}_i^H \mathbf{r}|^2$. It is obvious that different \mathbf{r} leads to different value of $\chi(N)$. By the equality

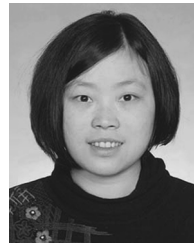
$$\max_{\|\mathbf{r}\|_2=1} \sum_{i=1}^N |\tilde{\mathbf{h}}_i^H \mathbf{r}|^2 = \max_{\|\mathbf{r}\|_2=1} \mathbf{r}^H \tilde{\mathbf{H}}_{rd} \tilde{\mathbf{H}}_{rd}^H \mathbf{r} = \lambda_{\max}(\tilde{\mathbf{H}}_{rd} \tilde{\mathbf{H}}_{rd}^H),$$

there always exists some \mathbf{r} , such that $\chi(N) = \sum_{i=1}^N |\tilde{\mathbf{h}}_i^H \mathbf{r}|^2 \leq \lambda_{\max}(\tilde{\mathbf{H}}_{rd} \tilde{\mathbf{H}}_{rd}^H)$. Hence $\varepsilon^2 < \chi(N) \leq \lambda_{\max}(\tilde{\mathbf{H}}_{rd} \tilde{\mathbf{H}}_{rd}^H)$ is the necessary and sufficient condition for $\Omega(k^\circ) > 0$. The proof is completed.

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. 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.
- [3] H. Wan, W. Chen, and X. Wang, "Joint source and relay design for MIMO relaying broadcast channels," *IEEE Commun. Lett.*, vol. 17, no. 2, pp. 345–348, Feb. 2013.
- [4] Z. Wang, W. Chen, F. Gao, and J. Li, "Capacity performance of efficient relay beamformings for dual-hop MIMO multi-relay networks with imperfect R-D CSI at relays," *IEEE Trans. Veh. Technol.*, vol. 60, no. 6, pp. 2608–2619, Jul. 2011.
- [5] B. Khoshnevis, Y. Wei, and R. Adve, "Grassmannian beamforming for MIMO amplify-and-forward relaying," *IEEE J. Sel. Area Commun.*, vol. 26, no. 8, pp. 1397–1407, Oct. 2008.
- [6] 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.
- [7] X. Zhang, Y. Xie, J. Li, and P. Stoica, "MIMO transmit beamforming under uniform elemental power constraint," *IEEE Trans. Signal Process.*, vol. 55, no. 11, pp. 5395–5406, Nov. 2007.
- [8] M. Vu, "MISO capacity with per-antenna power constraint," *IEEE Trans. Commun.*, vol. 59, no. 5, pp. 3336–3348, May 2011.
- [9] C. Xing, Z. Fei, Y. Zhou, Z. Pan, and H. Wang, "Transceiver Design with Matrix-Version Water-Filling Solutions under Mixed Power Constraints," [Online]. Available: <http://arxiv.org/pdf/1410.3673.pdf>
- [10] W. Yu T. Lan, "Transmitter optimization for the multi-antenna downlink with per-antenna power constraints," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2646–2660, Jun. 2007.
- [11] M. Dong, B. Liang, and Q. Xiao, "Unicast multi-antenna relay beamforming with per-antenna power control: Optimization and duality," *IEEE Trans. Signal Process.*, vol. 61, no. 23, pp. 6076–6091, Dec. 2013.
- [12] 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.
- [13] C. Xing, M. Xia, F. Gao, and Y. Wu, "Robust transceiver with Tomlinson-Harashima precoding for amplify-and-forward MIMO relaying systems," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 8, pp. 1370–1382, Sep. 2012.
- [14] C. Xing, S. Ma, Z. Fei, Y. Wu, and H. V. 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. 1196–1209, Mar. 2013.
- [15] 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.
- [16] H. Tang, W. Chen, J. Li, and H. Wan, "Achieving global optimality for joint source and relay beamforming design in two-hop relay channels," *IEEE Trans. Veh. Technol.*, vol. 63, no. 9, pp. 4422–4435, Nov. 2014.
- [17] 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.
- [18] H. Shen, J. Wang, W. Xu, Y. Rong, and C. Zhao, "A worst-case robust MMSE transceiver design for nonregenerative MIMO relaying," *IEEE Trans. Wireless Commun.*, vol. 13, no. 2, pp. 695–709, Feb. 2014.
- [19] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, Feb. 2014.
- [20] C. Jeong, B. Seo, S. R. Lee, H. Kim, and I. Kim, "Relay precoding for non-regenerative MIMO relay systems with partial CSI feedback," *IEEE Trans. Wireless Commun.*, vol. 11, no. 5, pp. 1698–1711, May 2012.
- [21] V. Havary-Nassab, S. Shahbazpanahi, and A. Grami, "Optimal distributed beamforming for two-way relay networks," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1238–1250, Mar. 2010.
- [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. 9, pp. 3272–3283, Sep. 2012.

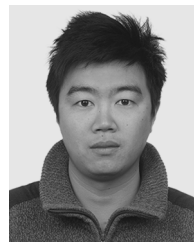
- [23] G. Zheng, K. K. Wong, A. Paulraj, and B. Ottersten, "Robust collaborative-relay beamforming," *IEEE Trans. Signal Process.*, vol. 57, no. 8, pp. 3130–3143, Aug. 2009.
- [24] M. Tao and R. Wang, "Robust relay beamforming for two-way relay networks," *IEEE Trans. Commun. Lett.*, vol. 16, no. 7, pp. 1052–1055, Jul. 2012.
- [25] Y. Huang, Q. Li, W. Ma, and S. Zhang, "Robust multicast beamforming for spectrum sharing-based cognitive radios," *IEEE Trans. Signal Process.*, vol. 60, no. 1, pp. 527–533, Jan. 2012.
- [26] A. Khabbazi-basmenj and S. A. Vorobyov, "Robust adaptive beamforming for general-rank signal model with positive semi-definite constraint via POTDC," *IEEE Trans. Signal Process.*, vol. 61, no. 23, pp. 6103–6117, Dec. 2013.
- [27] M. Grant and S. Boyd, "CVX' Users' Guide," 2009 [Online]. Available: <http://cvxr.com/cvx/doc/index.html>
- [28] N. D. Sidiropoulos, T. N. Davidson, and Z.-Q. Luo, "Transmit beamforming for physical-layer multicasting," *IEEE Trans. Signal Process.*, vol. 54, no. 6, pp. 2239–2251, Jun. 2006.
- [29] L. Vandenberghe and S. Boyd, "Semidefinite Programming," *SIAM Rev.*, pp. 49–95, 1996.
- [30] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.



Hongying Tang received the B.S. and M.S. degrees in Electronic engineering from the University of Electronic Science and Technology of China (UESTC), Chengdu, and Shanghai Jiao Tong University, Shanghai, China, in 2007 and 2009, respectively. She is currently pursuing her Ph.D. degree at Network Coding and Transmission Laboratory, Shanghai Jiao Tong University, Shanghai, China. Her current research interests include cooperative network and MIMO technique.



Wen Chen (M'03–SM'11) received the B.S. and M.S. degrees from Wuhan University, China, in 1990 and 1993, respectively, and the Ph.D. degree from the University of Electro-Communications, Tokyo, Japan, in 1999. He was a researcher of Japan Society for the Promotion of Sciences (JSPS) from 1999 through 2001. In 2001, he joined the University of Alberta, Canada, starting as a Postdoctoral Fellow with the Information Research Laboratory and continuing as a Research Associate in the Department of Electrical and Computer Engineering. Since 2006, he has been a Full Professor in the Department of Electronic Engineering, Shanghai Jiao Tong University, China, where he is also the Director of the Institute for Signal Processing and Systems. His interests cover network coding, cooperative communications, cognitive radio, and MIMO-OFDM systems. In this area, he has published 60 papers in IEEE journals and more than 100 papers in IEEE conferences.



Jun Li (M'09) received Ph. D degree in Electronic Engineering from Shanghai Jiao Tong University, Shanghai, P. R. China in 2009. From January 2009 to June 2009, he worked in the Department of Research and Innovation, Alcatel Lucent Shanghai Bell as a Research Scientist. From June 2009 to April 2012, he was a Postdoctoral Fellow at the School of Electrical Engineering and Telecommunications, the University of New South Wales, Australia. From April 2012 to now, he is a Research Fellow at the School of Electrical Engineering, the University of Sydney, Australia. He served as the Technical Program Committee member for several international conferences such as APCC2009, APCC2010, VTC2011 (Spring), ICC2011, TENCON2012, APCC2013, VTC2014 (Fall), and ICC2014. His research interests include network information theory, channel coding theory, wireless network coding and cooperative communications.