IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received April 5, 2017, accepted May 2, 2017, date of publication June 1, 2017, date of current version July 24, 2017.

Digital Object Identifier 10.1109/ACCESS.2017.2710908

A Two-Way Network Beamforming Approach Based on Total Power Minimization With Symmetric Relay Beamforming Matrices

RAZGAR RAHIMI AND SHAHRAM SHAHBAZPANAHI, (Senior Member, IEEE)
Department of Electrical, Computer, and Software Engineering, University of Ontario Institute of Technology, Oshawa, ON L1H 7K4, Canada
Corresponding author: Shahram Shahbazpanahi (shahram.shahbazpanahi@uoit.ca)

ABSTRACT We study the total transmit power minimization problem for a two-way relay network under two constraints on the transceivers' received signal-to-noise-ratios. The network considered herein consists of multiple multi-antenna relay nodes and two single-antenna transceivers. Each relay transforms the vector of its received signals, by multiplying this vector with a complex beamforming matrix, thereby obtaining a new vector whose entries are transmitted over different antennas of that relay. Assuming the relay beamforming matrices and the transceivers' transmit powers as the design parameters, we first study the total power minimization problem under the assumption that the relay beamforming matrices are symmetric. Under such an assumption, we show that the total power minimization problem is amenable to a semi-closed-form solution, and thus, it can be solved efficiently. We then consider the case, where the relay beamforming matrices may not be symmetric and show that in this case, the total power minimization problem can be solved using a computationally prohibitive algorithm which involves a 2-D search over a grid in the space of the transceivers' transmit powers and semi-definite programming at each vertex of this grid. Our numerical results show that the symmetric assumption on the relay beamforming matrices incurs only insignificant loss, while this assumption allows us to significantly reduce the computational burden of solving the total power minimization problem.

INDEX TERMS MIMO relaying, two-way relay networks, bidirectional relay networks, total power minimization, symmetric relay beamforming, network beamforming, semi-definite programming.

I. INTRODUCTION

Recently, cooperative networks have been extensively investigated in the literature. In a two-way relay network, two transceivers exchange their information symbols with the help of a number of relays. The traditional relaying strategy for providing an interference-free communication between two transceivers establishes a connection in four time-slots. In the first time-slot the signal is transmitted in a one-way relaying scheme from one of the transceivers to the relays. In the second time-slot, each relay transmits a processed version of its received signal toward the other transceiver. Subsequently, in the next two time-slots, a similar communication link is established in the opposite direction. Using the so-called time division broadcast (TDBC) strategy, it is possible to reduce number of required time-slots from four to three [1], [2]. In the TDBC protocol, transceivers send their signals in two consecutive time-slots. Each relay node then broadcasts a signal which is somehow obtained from the first two signals received by that relay. The multiple access broadcast (MABC) scheme [3]–[6] is another two-way relaying approach, which reduces the number of the required time-slots to two time-slots. In the first time-slot, the two transceivers transmit their signals simultaneously toward the relays. In the second time-slot, each relay transmits a processed version of its received signal.

In two-way networks, increasing the number of relays can extend the coverage range and can improve the spectral efficiency and/or the connection reliability between the transceivers. It has been observed that equipping the relays with multiple antennas can provide similar benefits. Studies show that employing multiple multi-antenna relays can significantly boost the achievable advantages. The majority of the published results on two-way relay networks consider two-way relaying schemes with single antenna nodes, see [6] and references therein. The studies conducted on two-way networks with multi-antenna nodes mainly consider a two-way relay network including a single multi-antenna relay which assists the establishment of a link between two single-antenna transceivers [7], [8]. Compared to the volume of the results published on networks with single multi-antenna relay and on networks with multiple single-antenna relays, studies focusing on networks with multiple multi-antenna relays are scarce. In the sequel, we review some of the results on networks with multiple multi-antenna relays. Yilmaz *et al.* [9] study a multi-pair two-way relay network where all the transceiver pairs communicate via one multi-antenna relay. Xu and Hua [10], Roemer and Haardt [11], Wang and Tao [12], Zhang and Haardt [13], Zhang *et al.* [14], and Lee *et al.* [15] consider the case of two-way networks with a single multi-antenna relay and multi-antenna transceivers. Lee *et al.* [16] and Vaze and Heath [17] consider a two-way MIMO relay network with multiple relays where all nodes are equipped with multiple antennas. As an extension to the aforementioned two-way networks, a multi-hop two-way relay channel is investigated in [18], where all the network nodes are equipped with multiple antennas. Moreover, Kha *et al.* [19] consider a two-way relay network where multiple pairs of single-antenna transceivers communicate in a pairwise manner with the help of multiple multi-antenna relays.

The published results on two-way relay networks can be categorized in terms of their design objective and constraints. Khabbazibasmenj [7], Zhang *et al.* [14], Lee *et al.* [15], Lee *et al.* [16], and Vaze and Heath [17] aim to maximize the achievable sum-rate. A max-min fair criterion and weighted sum-rate are investigated in [19] and [10], respectively. The antenna selection problem based on max-min channel coefficients criterion in [8], the interference mitigation at the transceivers in [9], the diversity multiplexing tradeoff analysis of [18], the mean-square-error minimization approach of [12], and the energy efficiency maximization technique of [13] are other examples of studies conducted on the two-way relay networks. Alsharoa *et al.* [20] study the problem of distributed beamforming for a network consisting of multi-antenna relays and multi-anennata transceivers using a total power minimization approach. This approach leads to a particle swarm optimization based solution. Considered in [21] is a two-way relay network consisting of three multi-antenna nodes (two transceivers and one relay node). Focusing on optimal joint source precoding and relay beamforming optimization, the author derives the optimal structure of the source and relay precoding matrices via minimizing the mean squared error of the symbol estimates at the two transceivers. Based on this optimal structure, a new iterative algorithm is developed to jointly optimize the relay and source matrices.

Our investigation in this paper has two novel aspects. The first novel aspect is that we study two-way relay networks with *multiple multi-antenna relays* - a type of two-way relay network which has not been considered much in the literature. The second novel aspect of our investigation is that we use the total transmit power consumed in the entire network as our design objective. Majority of the results published on multi-antenna two-way relay networks deal with the power consumed in the relays as the design objective. The focus of this study is on a two-way relay network consisting of two single-antenna transceivers and multiple multi-antenna relays. Assuming an MABC relaying scheme, our goal is to jointly obtain the optimal relay beamforming matrices as well as the optimal transceivers' transmit powers which minimize the total transmit power under given signal-to-noise-ratio (SNR) constraints at the transceivers. To do so, two different types of beamforming matrices are considered. We first restrict the relay beamforming matrices to be symmetric, thereby rendering the end-to-end channel between the two transceivers reciprocal. Under such a symmetry condition, we show that the aforementioned total power minimization yields a semi-closed form solution. We then use the pioneer results of [14] to solve the total power minimization problem for the case with general beamforming matrices (without assuming that these matrices are symmetric). Our simulation results show that imposing symmetry condition on the relay beamforming matrices incurs negligible performance loss, in terms of the total transmit power, while allowing us to obtain the design parameters in a computationally efficient manner.

The main contributions of this paper are summarized below:

- We obtain jointly optimal relay beamforming matrices as well as the optimal transceiver transmit powers for a network with multiple multi-antenna relays such that the total transmit power is minimized under given SNR constraints at the transceivers. We also discuss the conditions for the problem to be feasible.
- We show that based the results of [14], the relay beamforming matrices have special structures which can be exploited to reduce the computational complexity of the problem. We use this structure to reduce the dimensionality of the power minimization problem.
- In order to guarantee the reciprocity of the end-to-end channel between the transceivers, we choose beamforming matrices to be symmetric. For this type of beamforming matrices, we prove that this power minimization problem has a unique semi-closed-form solution. That is, given a certain intermediate parameter, the symmetric beamforming matrices can be obtained in a closed-form. We prove that this parameter can be obtained using the efficient Newton-Raphson technique.
- Assuming no symmetry for the relay beamforming matrices, we show that relying on the results of [14], the total power minimization problem can be solved using a combination of a two-dimensional search and semi-definite programming.
- Using numerical examples, we compare the required power for maintaining the SNRs at the receiver front-end of the transceivers above given thresholds, for both schemes with general and symmetric relay beamforming matrices. Our numerical results show that the scheme

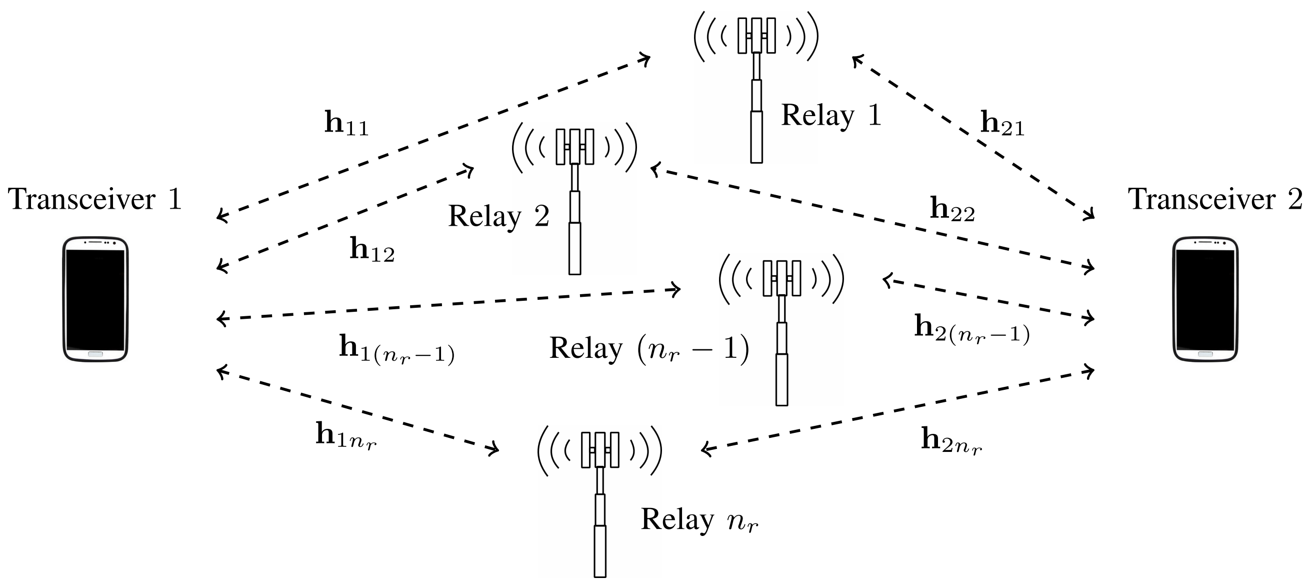

FIGURE 1. **A two-way relay network with multiple multi-antenna relays.**

with symmetric beamforming matrices keeps the SNRs above the required thresholds with a total network power which is very close to that for the scheme with optimal general beamforming matrices.

To the best of our knowledge these contributions are novel and they have not appeared in the literature for the problems considered in this paper. The problems of relay beamforming and transceivers' power allocation have been investigated in the literature for single-antenna multi-relay and multi-antenna single relay scenarios. However, the problem of total power minimization for multi-antenna multi-relay networks, where the beamforming matrices and transceivers power allocation need to be jointly considered, has not been studied, and this is exactly what this work aims to investigate. It is worth mentioning that the problem of total power minimization subject to quality of service constraints has been widely studied in the literature, see [6], [22]–[28]. The motivation behind total power minimization approaches is to find the minimal power consumption in the entire network while guaranteeing a certain quality of service at the receiver(s). Indeed, this approach aims to find the greenest design for the network.

Note that our work is different from those in [29] and [30], where *one-way* multi-antenna relaying schemes are studied. Golbon-Haghighi *et al.* [29] study a *one-way* network of multiple single-antenna source-destination pairs and a single multi-antenna relay, while the study in [30] considers a *one-way* network of multiple single-antenna source-destination pairs and multiple multi-antenna relays. Unlike [29] and [30], we are considering a *two-way* relay network of two single-antenna transceivers and multiple multi-antenna relays. Also in [29] and [30], the relay transmit power is minimized, assuming the source powers are fixed, while we herein minimize the total power consumed in the entire network assuming that the transceivers' transmit powers are to be optimally determined.

Note that the system model considered in this paper resembles the analog network coding for two-way relay networks introduced in [31]. The difference between this work and the work in [31] is that we consider a multi-relay scenario, while [31] studies a single-relay network.

Notation: Throughout this paper, we use small and capital boldface letters to denote vectors and matrices, respectively. The operators $(\cdot)^*$, $(\cdot)^T$, and $(\cdot)^H$ denote the complex conjugate, the transpose, and the Hermitian transpose, respectively. $[\mathbf{A}]_{ij}$ denotes the (i, j)-th entry of matrix $\mathbf{A}$. The operator $vec(\mathbf{A})$ is used for stacking the columns of $\mathbf{A}$ in one column vector $\mathbf{a}$. The operator $\otimes$ denotes the Kronecker product. $tr(\cdot)$ and $E\{\cdot\}$ denote the trace and statistical expectation operators, respectively. $\mathbf{A} \succeq 0$ means that $\mathbf{A}$ is a Hermitian positive semidefinite matrix. $\mathbf{I}_r$ is used to represent an $r \times r$ identity matrix. $\|\cdot\|$ and $|\cdot|$ denote the Euclidean norm of a vector and the absolute value of a complex scalar, respectively. $\mathcal{P}\{\mathbf{A}\}$ and $\lambda_{max}\{\mathbf{A}\}$ represent the normalized principal eigenvector and the principal eigenvalue of matrix $\mathbf{A}$, respectively.

II. SYSTEM MODEL

As shown in Fig. 1, the two-way relay network we consider consists of two single-antenna transceivers which wish to communicate with the help of n_r multi-antenna relays. The scenario we are considering can be used in cellular communication systems, where user devices can use only a single antenna due to their size and weight limitations and the base stations act as relays. Indeed, our scheme can be viewed as a distributed MIMO system used for connecting two single-antenna user devices. Equipping the relays (base stations) with multiple antennas allows local beamforming at the relays while distributed beamforming is materialized by all base stations collectively.

Each relay transforms the vector of its received signals by multiplying it with a complex "beamforming" matrix. We refer to such a scheme as transform-and-forward (TF) relaying protocol. To determine the relay beamforming matrices and the transceivers' transmit powers, we aim to minimize the total transmit power consumed in the entire network while SNRs at the receiver front-ends of the transceivers are kept higher than or equal to two given thresholds. Assuming that each relay node is equipped with M antennas, we consider the two time-slot MABC relaying scheme, where in the first time-slot, the two transceivers transmit their signals simultaneously and in the second time-slot, each relay forwards a linearly transformed version of its received signal vector to the two transceivers. We assume that no direct link exists between the transceivers, i.e., all data transmissions go through the relay nodes.

For $j \in \{1, 2\}$, let s_j denote the unit-power scalar information symbol transmitted by Transceiver j with transmission power p_j. Assuming frequency-flat fading transceiver-relay channels, the $M \times 1$ vector $\mathbf{x}_i$ of the received baseband signals at relay i in the first time-slot is given as

$$\mathbf{x}_i = \sqrt{p_1}\mathbf{h}_{1i}s_1 + \sqrt{p_2}\mathbf{h}_{2i}s_2 + \mathbf{n}_i, \quad \text{for } i \in \{1, \ldots, n_r\}. \tag{1}$$

Here, $\mathbf{n}_i$ is the $M \times 1$ received noise vector at the i-th relay, while $\mathbf{h}_{1i}$ and $\mathbf{h}_{2i}$ are the $M \times 1$ complex vectors of the coefficients corresponding to the channels between the i-th relay and Transceivers 1 and 2, respectively. Denoting the beamforming matrix of the i-th relay as an $M \times M$ complex matrix $\mathbf{A}_i$, the $M \times 1$ vector of the signal transmitted by the i-th relay is denoted by $\mathbf{t}_i$ and can be expressed as

$$\mathbf{t}_i = \mathbf{A}_i\mathbf{x}_i. \tag{2}$$

Assuming that the relay-transceiver channels are reciprocal for uplink and downlink transmissions, the received signals $y_1 = \sum_{i=1}^{n_r} \mathbf{h}_{1i}^T\mathbf{t}_i + \eta_1$ and $y_2 = \sum_{i=1}^{n_r} \mathbf{h}_{2i}^T\mathbf{t}_i + \eta_2$ at Transceivers 1 and 2 are written, respectively, as

$$y_1 = \sum_{i=1}^{n_r} \sqrt{p_1}\mathbf{h}_{1i}^T\mathbf{A}_i\mathbf{h}_{1i}s_1 + \sum_{i=1}^{n_r} \sqrt{p_2}\mathbf{h}_{1i}^T\mathbf{A}_i\mathbf{h}_{2i}s_2 + \sum_{i=1}^{n_r} \mathbf{h}_{1i}^T\mathbf{A}_i\mathbf{n}_i + \eta_1 \tag{3}$$

$$y_2 = \sum_{i=1}^{n_r} \sqrt{p_1}\mathbf{h}_{2i}^T\mathbf{A}_i\mathbf{h}_{1i}s_1 + \sum_{i=1}^{n_r} \sqrt{p_2}\mathbf{h}_{2i}^T\mathbf{A}_i\mathbf{h}_{2i}s_2 + \sum_{i=1}^{n_r} \mathbf{h}_{2i}^T\mathbf{A}_i\mathbf{n}_i + \eta_2 \tag{4}$$

where η_j is the received noise at Transceiver j, for $j \in \{1, 2\}$. Since the two transceivers know their own transmitted signals and assuming that they have the perfect knowledge of global channel state information (CSI), the first term in (3) and the second term in (4) (which are self-interference terms) can be subtracted from y_1 and y_2, respectively. The residual signals $\tilde{y}_1$ and $\tilde{y}_2$ are then given as

$$\tilde{y}_1 \triangleq \sum_{i=1}^{n_r} \sqrt{p_2}\mathbf{h}_{1i}^T\mathbf{A}_i\mathbf{h}_{2i}s_2 + \sum_{i=1}^{n_r} \mathbf{h}_{1i}^T\mathbf{A}_i\mathbf{n}_i + \eta_1 \tag{5}$$

$$\tilde{y}_2 \triangleq \sum_{i=1}^{n_r} \sqrt{p_1}\mathbf{h}_{2i}^T\mathbf{A}_i\mathbf{h}_{1i}s_1 + \sum_{i=1}^{n_r} \mathbf{h}_{2i}^T\mathbf{A}_i\mathbf{n}_i + \eta_2. \tag{6}$$

The noise processes at all nodes are assumed to be spatially white zero-mean complex Gaussian processes with variance σ^2. Therefore, we can write $E\{|\eta_1|^2\} = E\{|\eta_2|^2\} = \sigma^2$ and $E\{\mathbf{n}_i\mathbf{n}_i^H\} = \sigma^2\mathbf{I}_M$. Hence, using (5) and (6), we can express the SNRs at Transceivers 1 and 2 as

$$\mathrm{SNR}_1 = \frac{p_2\left|\sum_{i=1}^{n_r}\mathbf{h}_{1i}^T\mathbf{A}_i\mathbf{h}_{2i}\right|^2}{\sigma^2(1+\sum_{i=1}^{n_r}\|\mathbf{h}_{1i}^T\mathbf{A}_i\|^2)},$$

$$\mathrm{SNR}_2 = \frac{p_1\left|\sum_{i=1}^{n_r}\mathbf{h}_{2i}^T\mathbf{A}_i\mathbf{h}_{1i}\right|^2}{\sigma^2(1+\sum_{i=1}^{n_r}\|\mathbf{h}_{2i}^T\mathbf{A}_i\|^2)}. \tag{7}$$

The total transmit power P_T in the network is the summation of the transceivers' transmit powers and the transmit power of all the relays, that is $P_T = p_1 + p_2 + P_r$, where

$$P_r \triangleq p_1\sum_{i=1}^{n_r}\|\mathbf{A}_i\mathbf{h}_{1i}\|^2 + p_2\sum_{i=1}^{n_r}\|\mathbf{A}_i\mathbf{h}_{2i}\|^2 + \sigma^2\sum_{i=1}^{n_r} tr(\mathbf{A}_i\mathbf{A}_i^H) \tag{8}$$

is the total relay transmit power.

III. POWER MINIMIZATION

In the current study, we aim to find the beamforming matrices and the transceivers' transmit powers such that the total transmit power P_T is minimized, while the SNRs at Transceivers 1 and 2 are maintained above given thresholds γ_1 and γ_2, respectively. This power minimization problem can be expressed as[1 2 3]

$$\min_{p_1,p_2,\{\mathbf{A}_i\}_{i=1}^{n_r}} P_T \quad \text{subject to } \mathrm{SNR}_1 \geq \gamma_1, \quad \mathrm{SNR}_2 \geq \gamma_2. \tag{9}$$

[1]It is worth mentioning that a total power minimization approach has been widely considered as a design technique for relay networks, see for example [6], [20], [22]–[28]. The advantage of a total power minimization approach is to ensure the minimum amount of power is consumed in the entire network, thereby leading to the most power efficient design of the network.

[2]Note that the power consumption at each node is the sum of the node transmit power and the power consumed in the circuitry of the node. The latter power is the sum of the power consumption in the node circuitry, excluding the node power amplifier, which is constant, and the power consumed by the power amplifier and is a linear function of the node transmit power, see [32]. As such, minimizing the total transmit power will minimize the total power consumed in the network.

[3]Note that as shown in [14], the total power minimization problem in (9) can be used to solve a related problem, namely the weighted sum-rate maximization problem under a total power constraint. As shown in [14], the latter problem can be solved using a bisection type of algorithm along with an algorithm which solves the total power minimization problem. Interested readers are referred to [14] for more details on this approach.

Using (7) and (8), we can recast the optimization problem as

$$\min_{p_1,p_2,\{\mathbf{A}_i\}_{i=1}^{n_r}} \; p_1\left(1+\sum_{i=1}^{n_r}\|\mathbf{A}_i\mathbf{h}_{1i}\|^2\right)+p_2\left(1+\sum_{i=1}^{n_r}\|\mathbf{A}_i\mathbf{h}_{2i}\|^2\right) + \sigma^2\sum_{i=1}^{n_r} tr(\mathbf{A}_i\mathbf{A}_i^H)$$

$$\text{subject to} \quad \frac{p_2\left|\sum_{i=1}^{n_r}\mathbf{h}_{1i}^T\mathbf{A}_i\,\mathbf{h}_{2i}\right|^2}{\sigma^2\left(1+\sum_{i=1}^{n_r}\|\mathbf{h}_{1i}^T\mathbf{A}_i\|^2\right)} \geq \gamma_1, \quad \frac{p_1\left|\sum_{i=1}^{n_r}\mathbf{h}_{2i}^T\mathbf{A}_i\,\mathbf{h}_{1i}\right|^2}{\sigma^2\left(1+\sum_{i=1}^{n_r}\|\mathbf{h}_{2i}^T\mathbf{A}_i\|^2\right)} \geq \gamma_2. \tag{10}$$

We observe that at the optimum, the SNR inequality constraints in (10) are satisfied with equality, otherwise, if, at the optimum, any of these constraints is satisfied with inequality, then the corresponding optimal power can be reduced to satisfy this constraint with equality. This, in turn decreases the value of the objective function thereby contradicting the optimality. This observation implies that p_1 and p_2 can be respectively written as

$$p_1 = \frac{\sigma^2\gamma_2\left(1+\sum_{i=1}^{n_r}\|\mathbf{h}_{2i}^T\mathbf{A}_i\|^2\right)}{\left|\sum_{i=1}^{n_r}\mathbf{h}_{2i}^T\mathbf{A}_i\,\mathbf{h}_{1i}\right|^2}, \quad p_2 = \frac{\sigma^2\gamma_1\left(1+\sum_{i=1}^{n_r}\|\mathbf{h}_{1i}^T\mathbf{A}_i\|^2\right)}{\left|\sum_{i=1}^{n_r}\mathbf{h}_{1i}^T\mathbf{A}_i\,\mathbf{h}_{2i}\right|^2}. \tag{11}$$

Using (11), we rewrite (10) as the following unconstrained optimization problem:

$$\min_{\{\mathbf{A}_i\}_{i=1}^{n_r}} \frac{\gamma_2\sigma^2\left(1+\sum_{i=1}^{n_r}\|\mathbf{h}_{2i}^T\mathbf{A}_i\|^2\right)\left(1+\sum_{i=1}^{n_r}\|\mathbf{A}_i\mathbf{h}_{1i}\|^2\right)}{\left|\sum_{i=1}^{n_r}\mathbf{h}_{2i}^T\mathbf{A}_i\,\mathbf{h}_{1i}\right|^2} + \frac{\gamma_1\sigma^2\left(1+\sum_{i=1}^{n_r}\|\mathbf{h}_{1i}^T\mathbf{A}_i\|^2\right)\left(1+\sum_{i=1}^{n_r}\|\mathbf{A}_i\mathbf{h}_{2i}\|^2\right)}{\left|\sum_{i=1}^{n_r}\mathbf{h}_{1i}^T\mathbf{A}_i\,\mathbf{h}_{2i}\right|^2} + \sigma^2\sum_{i=1}^{n_r} tr(\mathbf{A}_i\mathbf{A}_i^H). \tag{12}$$

Let us denote the $M\times 2$ matrix that spans the vector space of $\mathbf{h}_{1i}$ and $\mathbf{h}_{2i}$ as $\mathbf{U}_i$, where $\mathbf{U}_i^H\mathbf{U}_i = \mathbf{I}_2$. Following [14, Th. 3.1], the optimal value of matrix $\mathbf{A}_i$ can be written, without any loss of optimality, as

$$\mathbf{A}_i = \mathbf{U}_i^*\mathbf{B}_i\mathbf{U}_i^H. \tag{13}$$

Here, $\mathbf{B}_i$ is a 2×2 complex matrix which can be viewed, as shown in the sequel, as the effective beamforming matrix of the i-th relay. In light of (13), the beamforming matrix $\mathbf{A}_i$ is a cascade of three operations. The first operation is a receive beamforming matrix $\mathbf{U}_i^H$, which filters out the components of the relay received noise vector that do not reside in the signal subspace defined as the space spanned by $\mathbf{h}_{1i}$ and $\mathbf{h}_{2i}$. The second operation is denoted with $\mathbf{B}_i$ which transforms the output vector of the relay receive beamformer into a new vector. The third operation is a transmit beamforming operation represented by matrix $\mathbf{U}_i^*$ which guarantees that the transformed vector is transmitted only into the signal subspace. The matrices $\{\mathbf{B}_i\}_{i=1}^{n_r}$ are now determined such that the total transmit power is minimized subject to SNR constraints. That is, instead of finding the optimal values $\{\mathbf{A}_i\}_{i=1}^{n_r}$, without loss of optimality, we can obtain the optimal values of $\{\mathbf{B}_i\}_{i=1}^{n_r}$.

Let us define $\mathbf{q}_{1i} \triangleq \mathbf{U}_i^H\mathbf{h}_{1i}$ and $\mathbf{q}_{2i} \triangleq \mathbf{U}_i^H\mathbf{h}_{2i}$ as the effective channel vectors between the i-th relay and Transceivers 1 and 2, respectively. Then, the unconstrained problem in (12) can be equivalently written as

$$\min_{\{\mathbf{B}_i\}_{i=1}^{n_r}} \frac{\gamma_2\sigma^2\left(1+\sum_{i=1}^{n_r}\|\mathbf{q}_{2i}^T\mathbf{B}_i\|^2\right)\left(1+\sum_{i=1}^{n_r}\|\mathbf{B}_i\mathbf{q}_{1i}\|^2\right)}{\left|\sum_{i=1}^{n_r}\mathbf{q}_{2i}^T\mathbf{B}_i\,\mathbf{q}_{1i}\right|^2} + \frac{\gamma_1\sigma^2\left(1+\sum_{i=1}^{n_r}\|\mathbf{q}_{1i}^T\mathbf{B}_i\|^2\right)\left(1+\sum_{i=1}^{n_r}\|\mathbf{B}_i\mathbf{q}_{2i}\|^2\right)}{\left|\sum_{i=1}^{n_r}\mathbf{q}_{1i}^T\mathbf{B}_i\,\mathbf{q}_{2i}\right|^2} + \sigma^2\sum_{i=1}^{n_r} tr(\mathbf{B}_i\mathbf{B}_i^H) \tag{14}$$

where the effective beamforming matrices $\{\mathbf{B}_i\}_{i=1}^{n_r}$ are now the optimization variables.

IV. POWER MINIMIZATION WITH SYMMETRIC BEAMFORMING MATRICES

A. SYMMETRIC RELAY BEAMFORMING MATRICES

To ensure the end-to-end reciprocity between the transceivers, we choose $\mathbf{A}_i$ to be a symmetric matrix, i.e., $\mathbf{A}_i = \mathbf{A}_i^T$. Indeed, from (3) and (4), the end-to-end gains are $\mathbf{h}_{1i}^T\mathbf{A}_i\mathbf{h}_{2i}$ and $\mathbf{h}_{2i}^T\mathbf{A}_i\mathbf{h}_{1i}$ which will be equal if we choose $\mathbf{A}_i = \mathbf{A}_i^T$. Assuming a symmetric[4] beamforming matrix $\mathbf{A}_i$, leads to a symmetric matrix $\mathbf{B}_i$, i.e., $\mathbf{B}_i = \mathbf{B}_i^T$. It is thus observed that in this case, for minimizing total power, the optimal scheme needs to determine $3n_r$ unknown complex parameters as each of the n_r matrices $\{\mathbf{B}_i\}_{i=1}^{n_r}$ has only three unknown complex parameters, which are to be optimally determined. Using the symmetric beamforming matrices assumption, the

[4] In the next section, we consider the case of non-symmetric beamforming matrices.

optimization problem (14) can be rewritten as

$$\min_{\{\mathbf{B}_i\}_{i=1}^{n_r}} \frac{\sigma^2(\gamma_1+\gamma_2)\left(1+\sum_{i=1}^{n_r}\|\mathbf{q}_{2i}^T\mathbf{B}_i\|^2\right)\left(1+\sum_{i=1}^{n_r}\|\mathbf{B}_i\mathbf{q}_{1i}\|^2\right)}{\left|\sum_{i=1}^{n_r}\mathbf{q}_{2i}^T\mathbf{B}_i\mathbf{q}_{1i}\right|^2} + \sigma^2\sum_{i=1}^{n_r} tr(\mathbf{B}_i\mathbf{B}_i^H)$$
$$\text{subject to } [\mathbf{B}_i]_{(1,2)} = [\mathbf{B}_i]_{(2,1)}, \quad \text{for } i = 1, 2, \ldots, n_r \quad (15)$$

where the last set of constraints guarantees that $\{\mathbf{B}_i\}_{i=1}^{n_r}$ are symmetric. Assuming that the beamforming matrices are symmetric renders the end-to-end channel over each relaying path reciprocal, i.e., $\mathbf{q}_{1i}^T\mathbf{B}_i\mathbf{q}_{2i} = \mathbf{q}_{2i}^T\mathbf{B}_i\mathbf{q}_{1i}$, and also leads to the following equalities $\|\mathbf{q}_{1i}^T\mathbf{B}_i\| = \|\mathbf{B}_i\mathbf{q}_{1i}\|$ and $\|\mathbf{q}_{2i}^T\mathbf{B}_i\| = \|\mathbf{B}_i\mathbf{q}_{2i}\|$, and thus, allows us to write the optimization problem (14) as in (15). The latter optimization, as we show in the sequel, is amenable to a computationally affordable solution, which is globally optimal under the assumption of symmetric beamforming matrices. We now observe that the matrices $\{\mathbf{B}_i\}_{i=1}^{n_r}$ remain unchanged for different values of γ_1 and γ_2 as long as $\gamma_1+\gamma_2$ does not change.[5] Hence, in (10), if we replace γ_2 with $\gamma_1+\gamma_2$ and then set γ_1 to 0, the optimal values of $\{\mathbf{A}_i\}_{i=1}^{n_r}$ (or equivalently the optimal values of $\{\mathbf{B}_i\}_{i=1}^{n_r}$) will not change. Note that in (10), replacing γ_1 with 0, means that p_2 will be equal to 0. Therefore, *as long as the optimal values of* $\{\mathbf{B}_i\}_{i=1}^{n_r}$ *are concerned*, we can solve the following optimization problem:

$$\min_{\tilde{p}_1,\{\mathbf{B}_i\}_{i=1}^{n_r}} \quad \tilde{p}_1\left(1+\sum_{i=1}^{n_r}\|\mathbf{B}_i\mathbf{q}_{1i}\|^2\right)+\sigma^2\sum_{i=1}^{n_r} tr(\mathbf{B}_i\mathbf{B}_i^H)$$
$$\text{subject to } \frac{\tilde{p}_1\left|\sum_{i=1}^{n_r}\mathbf{q}_{2i}^T\mathbf{B}_i\,\mathbf{q}_{1i}\right|^2}{\sigma^2\left(1+\sum_{i=1}^{n_r}\|\mathbf{q}_{2i}^T\mathbf{B}_i\|^2\right)} \geq \gamma_1+\gamma_2$$
$$[\mathbf{B}_i]_{(1,2)} = [\mathbf{B}_i]_{(2,1)}, \quad \text{for } i = 1, 2, \ldots, n_r. \quad (16)$$

Note that the optimal value for $\tilde{p}_1$ in (16) is not the same as the optimal value of p_1 in (10). In other words, the matrices $\{\mathbf{B}_i\}_{i=1}^{n_r}$ obtained by solving (10) are identical to the matrices $\{\mathbf{B}_i\}_{i=1}^{n_r}$ obtained by solving (16). However, the value of $\tilde{p}_1$ obtained by solving (16) is not the same as the value of p_1 obtained by solving (10). To obtain the optimal values of p_1 and p_2 in (10), once the optimal values of $\{\mathbf{B}_i\}_{i=1}^{n_r}$ in (16) are obtained, we can use (13) to obtain the corresponding optimal values of $\{\mathbf{A}_i\}_{i=1}^{n_r}$. The so-obtained $\mathbf{A}_i$'s can then be used in (11) to calculate the optimal values of p_1 and p_2. Indeed, by solving (16), we aim to find the optimal values of $\{\mathbf{B}_i\}_{i=1}^{n_r}$ and the transmit power of Transceiver 1 in a *one-way relay-assisted communication* scheme, where the received SNR at Transceiver 2 is at least equal to $\gamma_1+\gamma_2$. Using the following identities $tr(\mathbf{ABC}) = (vec(\mathbf{A}^T))^T\,(\mathbf{I}\otimes\mathbf{B})\,vec(\mathbf{C})$ and $tr(\mathbf{A}^T\mathbf{BCD}^T) = (vec(\mathbf{A}^T))^T\,(\mathbf{D}\otimes\mathbf{B})\,vec(\mathbf{C})$, defining $\mathbf{b}_i \triangleq (vec(\mathbf{B}_i^T))^*$ and $\mathbf{f}_i \triangleq vec(\mathbf{q}_{1i}\,\mathbf{q}_{2i}^T)$, and after some algebraic manipulation, we can rewrite the optimization problem in (16) as

$$\min_{\tilde{p}_1,\{\mathbf{b}_i\}_{i=1}^{n_r}} \quad \tilde{p}_1(1+\sum_{i=1}^{n_r}\mathbf{b}_i^H(\mathbf{I}_2\otimes\mathbf{q}_{1i}\mathbf{q}_{1i}^H)\mathbf{b}_i)+\sigma^2\sum_{i=1}^{n_r}\mathbf{b}_i^H\mathbf{b}_i$$
$$\text{subject to } \tilde{p}_1\left|\sum_{i=1}^{n_r}\mathbf{b}_i^H\mathbf{f}_i\right|^2 - \sigma^2(\gamma_1+\gamma_2)\left(1+\sum_{i=1}^{n_r}\mathbf{b}_i^H(\mathbf{q}_{2i}\mathbf{q}_{2i}^H\otimes\mathbf{I}_2)\mathbf{b}_i\right)\geq 0$$
$$[\mathbf{b}_i]_2 = [\mathbf{b}_i]_3, \quad \text{for } i = 1, 2, \ldots, n_r. \quad (17)$$

We now define $\mathbf{b} \triangleq [\mathbf{b}_1^T, \mathbf{b}_2^T, \ldots, \mathbf{b}_{n_r}^T]^T$ and $\mathbf{f} \triangleq [\mathbf{f}_1^T, \mathbf{f}_2^T, \ldots, \mathbf{f}_{n_r}^T]^T$, and hence, can write $\left|\sum_{i=1}^{n_r}\mathbf{b}_i^H\mathbf{f}_i\right|^2 = |\mathbf{b}^H\mathbf{f}|^2 = \mathbf{b}^H\mathbf{f}\,\mathbf{f}^H\mathbf{b}$. Doing so, we can express the optimization problem in (17) as

$$\min_{\tilde{p}_1}\ \tilde{p}_1 + \min_{\mathbf{b}}\ \mathbf{b}^H\left(\tilde{p}_1\mathbf{E}_0+\sigma^2\mathbf{I}_{4n_r}\right)\mathbf{b}$$
$$\text{s.t. } \mathbf{b}^H\left(\tilde{p}_1\mathbf{E}_1-\sigma^2(\gamma_1+\gamma_2)\mathbf{E}_2\right)\mathbf{b} \geq \sigma^2(\gamma_1+\gamma_2)$$
$$[\mathbf{b}]_{(i-1)n_r+2}=[\mathbf{b}]_{(i-1)n_r+3} \quad \text{for } i = 1, 2, \cdots, n_r \quad (18)$$

where $\mathbf{E}_0$, $\mathbf{E}_1$, and $\mathbf{E}_2$ are defined as

$$\mathbf{E}_0 \triangleq \text{blkdiag}\left(\{\mathbf{I}_2\otimes\mathbf{q}_{1i}\,\mathbf{q}_{1i}^H\}_{i=1}^{n_r}\right), \quad (19)$$
$$\mathbf{E}_1 \triangleq \mathbf{f}\,\mathbf{f}^H, \quad (20)$$
$$\mathbf{E}_2 \triangleq \text{blkdiag}\left(\{\mathbf{q}_{2i}\,\mathbf{q}_{2i}^H\otimes\mathbf{I}_2)\}_{i=1}^{n_r}\right). \quad (21)$$

Here blkdiag$(\cdot)$ stands for a block diagonal matrix. To solve (18), we can first fix $\tilde{p}_1$ and solve the minimization over $\mathbf{b}$. This value of $\mathbf{b}$ will be a function of $\tilde{p}_1$. We plug this value of $\mathbf{b}$ into the objective function of (18), thereby turning this function into a function of $\tilde{p}_1$ only. We then deal with solving a single-variable optimization problem. To further elaborate on this approach, we now focus on the inner minimization in (18).

B. INNER MINIMIZATION IN (18)

For *any given feasible value* of $\tilde{p}_1$, we rewrite this minimization as

$$\min_{\mathbf{b}}\ \mathbf{b}^H\left(\tilde{p}_1\mathbf{E}_0+\sigma^2\mathbf{I}_{4n_r}\right)\mathbf{b}$$
$$\text{s.t. } \mathbf{b}^H(\tilde{p}_1\,\mathbf{E}_1-\sigma^2(\gamma_1+\gamma_2)\mathbf{E}_2)\,\mathbf{b} \geq \sigma^2(\gamma_1+\gamma_2)$$
$$[\mathbf{b}]_{(i-1)n_r+2}=[\mathbf{b}]_{(i-1)n_r+3} \quad \text{for } i = 1, 2, \ldots, n_r. \quad (22)$$

Using the following definitions:

$$\mathbf{T} \triangleq \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{L} \triangleq \mathbf{I}_{n_r}\otimes\mathbf{T} \quad (23)$$

[5]Note that in case of single-antenna relays, each relay beamforming matrices shrinks to a scalar, and thus, the symmetric property of relay beamforming weights is automatically satisfied. The case of single-antenna relays which was studied in [6] and [26] has indeed inspired us to resort to symmetric beamforming matrices.

we can write $\mathbf{b}_i = \mathbf{T}\tilde{\mathbf{b}}_i$, where $\tilde{\mathbf{b}}_i = [\,[\mathbf{b}_i]_1\ [\mathbf{b}_i]_2\ [\mathbf{b}_i]_4\,]^T$ is the vector of the free parameters in $\mathbf{b}_i$. We can further write $\mathbf{b} = \mathbf{L}\tilde{\mathbf{b}}$, where $\tilde{\mathbf{b}} = [\tilde{\mathbf{b}}_1^T\ \tilde{\mathbf{b}}_2^T \ldots \tilde{\mathbf{b}}_{n_r}^T]^T$. These definitions enable us to rewrite (22) as

$$\min_{\tilde{\mathbf{b}}}.\ \tilde{\mathbf{b}}^H(\tilde{p}_1\tilde{\mathbf{E}}_0 + \sigma^2\mathbf{L}^H\mathbf{L})\tilde{\mathbf{b}}$$
$$\text{subject to}\ \ \tilde{\mathbf{b}}^H(\tilde{p}_1\tilde{\mathbf{E}}_1 - \sigma^2(\gamma_1+\gamma_2)\tilde{\mathbf{E}}_2)\tilde{\mathbf{b}} \geq \sigma^2(\gamma_1+\gamma_2) \quad (24)$$

where we further define: $\tilde{\mathbf{E}}_0 \triangleq \mathbf{L}^H\,\mathbf{E}_0\,\mathbf{L}$, $\tilde{\mathbf{E}}_1 \triangleq \mathbf{L}^H\,\mathbf{E}_1\,\mathbf{L}$, and $\tilde{\mathbf{E}}_2 \triangleq \mathbf{L}^H\,\mathbf{E}_2\,\mathbf{L}$.

We show in Appendix A that the problem in (24) is feasible if and only if

$$\tilde{p}_1 > \frac{\sigma^2(\gamma_1+\gamma_2)}{\mathbf{q}_1^H\mathbf{q}_1}. \quad (25)$$

We now aim to solve the minimization problem in (24) for any feasible value of $\tilde{p}_1$ which satisfies (25). We note that under the feasibility condition in (25), this problem is a quadratic programming problem. Based on the fact that for any feasible $\tilde{p}_1$ at the optimum, the inequality constraint in (24) is satisfied with equality, and thus, we can use the method of Lagrangian multipliers to solve (24). As a result, the solution to (24), denoted by $\tilde{\mathbf{b}}^{\text{opt}}(\tilde{p}_1)$, is obtained as[6]

$$\tilde{\mathbf{b}}^{\text{opt}}(\tilde{p}_1) = \alpha\mathbf{u}(\tilde{p}_1). \quad (26)$$

Here, $\mathbf{u}(\tilde{p}_1) = \mathcal{P}\{\mathbf{S}(\tilde{p}_1)\}$ is the normalized principal eigenvector of the matrix[7]

$$\mathbf{S}(\tilde{p}_1) = (\tilde{p}_1\tilde{\mathbf{E}}_0 + \sigma^2\mathbf{L}^H\mathbf{L})^{-1}(\tilde{p}_1\tilde{\mathbf{E}}_1 - \sigma^2(\gamma_1+\gamma_2)\tilde{\mathbf{E}}_2) \quad (27)$$

and α is a scalar factor which guarantees that the constraint in (24) is satisfied with equality and is given as

$$\alpha = \left(\frac{\sigma^2(\gamma_1+\gamma_2)}{\mathbf{u}^H(\tilde{p}_1)(\tilde{p}_1\,\tilde{\mathbf{E}}_1\ - \sigma^2(\gamma_1+\gamma_2)\tilde{\mathbf{E}}_2\,)\,\mathbf{u}(\tilde{p}_1)}\right)^{1/2}. \quad (28)$$

In the next subsection, we address the problem of optimally obtaining the parameter $\tilde{p}_1$.

C. OPTIMIZING $\tilde{p}_1$

We can now rewrite the main problem in (18) as

$$\min_{\tilde{p}_1}.\ \tilde{p}_1 + \frac{\sigma^2(\gamma_1+\gamma_2)}{\lambda(\tilde{p}_1)}\ \ \text{subject to}\ \ \tilde{p}_1 > \frac{\sigma^2(\gamma_1+\gamma_2)}{\mathbf{q}_1^H\mathbf{q}_1} \quad (29)$$

where $\lambda(\tilde{p}_1) = \lambda_{max}\{\mathbf{S}(\tilde{p}_1)\}$ represents the principal eigenvalue of the matrix $\mathbf{S}(\tilde{p}_1)$.

Lemma 1: The objective function in (29) has a unique extremum point in the interval $(\frac{\sigma^2(\gamma_1+\gamma_2)}{\mathbf{q}_1^H\mathbf{q}_1}, +\infty)$, *which is the global minimum of this objective function.*

Proof: See Appendix B. ∎

The unique solution to (29) can be obtained by equating the derivative of the objective function in (29) to zero. Denoting the objective function in (29) as $\psi(\tilde{p}_1)$, we show in Appendix C that derivative of $\psi(\tilde{p}_1)$ with respect to $\tilde{p}_1$ is given by

$$\begin{aligned} g(\tilde{p}_1) &\triangleq \frac{\partial\psi(\tilde{p}_1)}{\partial\tilde{p}_1} = 1 - \sigma^2(\gamma_1+\gamma_2)\frac{\frac{\partial}{\partial\tilde{p}_1}\lambda(\tilde{p}_1)}{\lambda^2(\tilde{p}_1)} \\ &= 1 - \sigma^2(\gamma_1+\gamma_2) \\ &\quad\times \frac{\tilde{p}_1^{-2} - \lambda(\tilde{p}_1)\tilde{\mathbf{f}}^H\mathbf{A}^{-1}(\tilde{p}_1)\tilde{\mathbf{E}}_0\mathbf{A}^{-1}(\tilde{p}_1)\tilde{\mathbf{f}}}{\lambda^2(\tilde{p}_1)\tilde{\mathbf{f}}^H\mathbf{A}^{-1}(\tilde{p}_1)(\tilde{p}_1\tilde{\mathbf{E}}_0+\sigma^2\mathbf{L}^T\mathbf{L})\mathbf{A}^{-1}(\tilde{p}_1)\tilde{\mathbf{f}}}. \end{aligned} \quad (30)$$

Here, the following definitions are used:

$$\mathbf{A}(\tilde{p}_1) \triangleq \sigma^2(\gamma_1+\gamma_2)\tilde{\mathbf{E}}_2 + \lambda(\tilde{p}_1)(\tilde{p}_1\tilde{\mathbf{E}}_0 + \sigma^2\mathbf{L}^T\,\mathbf{L}) \quad (31)$$
$$\tilde{\mathbf{f}} \triangleq \mathbf{L}^H\mathbf{f}, \quad (32)$$

and $\lambda(\tilde{p}_1)$ is the *largest eigenvalue* of the matrix $\mathbf{S}(\tilde{p}_1)$, and can be obtained, for any feasible value of $\tilde{p}_1$, as the *provably unique positive solution* to the following equation:

$$\tilde{p}_1\tilde{\mathbf{f}}^H\left(\sigma^2(\gamma_1+\gamma_2)\tilde{\mathbf{E}}_2 + \lambda(\tilde{p}_1)(\tilde{p}_1\tilde{\mathbf{E}}_0 + \sigma^2\mathbf{L}^T\mathbf{L})\right)^{-1}\tilde{\mathbf{f}} = 1. \quad (33)$$

This unique solution can be obtained using a simple Newton-Raphson method or a bisection method. Once $\lambda(\tilde{p}_1)$ is obtained, the corresponding value of $g(\tilde{p}_1)$ can be obtained and thus the equation $g(\tilde{p}_1) = 0$ can be solved using another bisection method, thereby the optimum value of $\tilde{p}_1$ can be obtained. Denoting the so-obtained optimal value of $\tilde{p}_1$ as $\tilde{p}_1^{\text{o}}$, we can use (26) to obtain $\tilde{\mathbf{b}}^{\text{opt}}(\tilde{p}_1^{\text{o}})$. The optimal value of $\mathbf{b}$ can then be calculated as $\mathbf{b}^{\text{opt}} = [\mathbf{b}_1^T\ \mathbf{b}_2^T \ldots \mathbf{b}_{n_r}^T]^T = \mathbf{L}\tilde{\mathbf{b}}^{\text{opt}}(\tilde{p}_1^{\text{o}})$. Reshaping $\mathbf{b}_i$ yields the optimal value of $\mathbf{B}_i$ and finally the optimal value of $\mathbf{A}_i$ can be obtained from $\mathbf{A}_i = \mathbf{U}_i^*\mathbf{B}_i\mathbf{U}_i^H$. One can then use the so-obtained $\mathbf{A}_i$ in (11) to obtain the transceivers' transmit powers in closed-forms.

The proposed technique is summarized as in Algorithm 1.

V. POWER MINIMIZATION WITH GENERAL BEAMFORMING MATRICES

In this section, we present the solution to the power minimization problem for the case when the beamforming matrices are not constrained to be symmetric. The solution to this case can then be used to evaluate the performance of the power minimization problem with symmetric beamforming matrices. To develop the solution to the case of general beamforming matrices, we rely on the pioneer results of [14], which considers a three-node two-way relay network and minimizes the transmit power consumed in a single multi-antenna relay subject to SNR constraints at two single-antenna transceivers. Note however that Zhang *et al.* [14] assume that the transceivers' transmit powers are fixed, while in our work, these powers are part of the design parameters. Nevertheless, the technique of [14] can

[6]Indeed, the optimization problem (24) is a quadratic programming problem and has a closed-form solution as in (26).

[7]From (23), we obtain

$$\mathbf{T}^T\mathbf{T} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{L}^T\mathbf{L} = \mathbf{I}_{n_r} \otimes \mathbf{T}^T\mathbf{T}.$$

Note that, $\mathbf{L}^T\mathbf{L}$ is a block diagonal matrix of full-rank matrices $\mathbf{T}^T\mathbf{T}$. Hence, $(\tilde{p}_1\tilde{\mathbf{E}}_0 + \sigma^2\mathbf{L}^T\,\mathbf{L})$ is a full-rank matrix and thus invertible.

Algorithm 1 Based on Bisection Method

1) Calculate $\mathbf{E}_0 = \text{blkdiag}\left(\{\mathbf{I}_2 \otimes \mathbf{q}_{1i}\mathbf{q}_{1i}^H\}_{i=1)}^{n_r}\right)$ and $\mathbf{E}_2 = \text{blkdiag}\left(\{\mathbf{q}_{2i}\,\mathbf{q}_{2i}^H \otimes \mathbf{I}_2\}_{i=1}^{n_r}\right)$ as well as $\mathbf{L} = \mathbf{I}_{n_r} \otimes \mathbf{T}$, where $\mathbf{T} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. Then, calculate $\tilde{\mathbf{E}}_0 = \mathbf{L}^H\,\mathbf{E}_0\,\mathbf{L}$, $\tilde{\mathbf{E}}_2 = \mathbf{L}^H\,\mathbf{E}_2\,\mathbf{L}$, and $\tilde{\mathbf{f}} = \mathbf{L}^H\mathbf{f}$ where the vector $\mathbf{f}$ is obtained as
$$\mathbf{f} = [(vec(\mathbf{q}_{11}\mathbf{q}_{21}^T))^T \quad (vec(\mathbf{q}_{12}\mathbf{q}_{22}^T))^T \quad \cdots \quad (vec(\mathbf{q}_{1n_r}\mathbf{q}_{2n_r}^T))^T]^T.$$
2) For any value of $z \in (\frac{\sigma^2(\gamma_1+\gamma_2)}{\mathbf{q}_1^H\mathbf{q}_1}, +\infty)$, define function $g(\cdot)$ as
$$g(z) = 1 - \sigma^2(\gamma_1+\gamma_2)\frac{z^{-2} - \lambda(z)\mathbf{u}^H(z)\tilde{\mathbf{E}}_0\mathbf{u}(z)}{\lambda^2(z)\mathbf{u}^H(z)(z\tilde{\mathbf{E}}_0 + \sigma^2\mathbf{L}^T\,\mathbf{L})\mathbf{u}(z)}.$$
Here, for any value of $z \in (\frac{\sigma^2(\gamma_1+\gamma_2)}{\mathbf{q}_1^H\mathbf{q}_1}, +\infty)$, the value of $\lambda(z)$ is obtained, using a bisection method, as the provably unique positive solution to the following non-linear equation:
$$z\,\mathbf{f}^H(\sigma^2(\gamma_1+\gamma_2)\tilde{\mathbf{E}}_2 + \lambda(z)(z\tilde{\mathbf{E}}_0 + \sigma^2\mathbf{L}^H\mathbf{L}))^{-1}\mathbf{f} - 1 = 0$$
and for any value of z, the $3n_r \times 1$ vector $\mathbf{u}(z)$ is obtained as
$$\mathbf{u}(z) = (\sigma^2(\gamma_1+\gamma_2)\tilde{\mathbf{E}}_2 + \lambda(z)(z\tilde{\mathbf{E}}_0 + \sigma^2\mathbf{L}^T\mathbf{L}))^{-1}\tilde{\mathbf{f}}$$
3) To solve $g(z) = 0$ in the interval $z \in (\frac{\sigma^2(\gamma_1+\gamma_2)}{\mathbf{q}_1^H\mathbf{q}_1}, +\infty)$, using a bisection method, choose z_l as
$$z_l = \frac{\sigma^2(\gamma_1+\gamma_2)}{\mathbf{q}_1^H\mathbf{q}_1} + \epsilon_1$$
where ϵ_1 is an arbitrarily small positive number such that $g(z_l) < 0$. Also, choose z_u large enough such that $g(z_u) > 0$.
4) Choose ϵ_2 to be an arbitrarily small positive number.
5) Choose $z = (z_l + z_u)/2$.
6) If $|g(z)| < \epsilon_2$, go to Step 7. If $g(z) < -\epsilon_2$, then $z_l = z$. If $g(z) > \epsilon_2$, then $z_u = z$. Go to Step 5.
7) Set $\tilde{p}_1^{\text{o}}$ equal to z and use a bisection technique to obtain the optimal value of λ, denoted as λ^{o}, as the unique positive solution to the following non-linear equation:
$$\tilde{p}_1^{\text{o}}\tilde{\mathbf{f}}^H(\sigma^2(\gamma_1+\gamma_2)\tilde{\mathbf{E}}_2 + \lambda(\tilde{p}_1^{\text{o}}\tilde{\mathbf{E}}_0 + \sigma^2\mathbf{L}^H\mathbf{L}))^{-1}\tilde{\mathbf{f}} - 1 = 0.$$
8) Calculate the total transmitter power, denoted as P_T, consumed in the entire network as
$$P_T = \tilde{p}_1^{\text{o}} + \frac{\sigma^2(\gamma_1+\gamma_2)}{\lambda^{\text{o}}}$$
9) Obtain $\tilde{\mathbf{b}}^{\text{opt}}(\tilde{p}_1) = [\tilde{\mathbf{b}}_1^T \quad \tilde{\mathbf{b}}_2^T \quad \cdots \quad \tilde{\mathbf{b}}_{n_r}^T]^T$ as
$$\tilde{\mathbf{b}}^{\text{opt}}(\tilde{p}_1) = \kappa\,\underbrace{(\sigma^2(\gamma_1+\gamma_2)\tilde{\mathbf{E}}_2 + \lambda^{\text{o}}(\tilde{p}_1^{\text{o}}\tilde{\mathbf{E}}_0 + \sigma^2\mathbf{L}^H\mathbf{L}))^{-1}\tilde{\mathbf{f}}}_{\mathbf{u}(\tilde{p}_1)}$$
where κ is obtained as
$$\kappa = \sqrt{\frac{\sigma^2(\gamma_1+\gamma_2)}{\lambda^{\text{o}}\mathbf{u}^H(\tilde{p}_1)(\tilde{p}_1\tilde{\mathbf{E}}_1 + \sigma^2\mathbf{L}^H\mathbf{L})\mathbf{u}(\tilde{p}_1)}}.$$
10) Calculate $\mathbf{b}^{\text{opt}} = [\mathbf{b}_1^T\ \mathbf{b}_2^T\ \ldots\ \mathbf{b}_{n_r}^T]^T = \mathbf{L}\tilde{\mathbf{b}}^{\text{opt}}(\tilde{p}_1^{\text{o}})$.
11) Reshape $\mathbf{b}_i$ to obtain the optimal value of the effective beamforming matrix $\mathbf{B}_i$ of the i-th relay, and finally, obtain the optimal value of the beamforming matrix of the i-th relay as $\mathbf{A}_i = \mathbf{U}_i^*\mathbf{B}_i\mathbf{U}_i^H$.
12) Use the so-obtained beamforming matrices to obtain the optimal values of the transceivers' transmit powers in closed-forms as:
$$p_1 = \frac{\sigma^2\gamma_2\left(1 + \sum_{i=1}^{n_r}\|\mathbf{h}_{2i}^T\mathbf{A}_i\|^2\right)}{\left|\sum_{i=1}^{n_r}\mathbf{h}_{2i}^T\mathbf{A}_i\,\mathbf{h}_{1i}\right|^2}, \qquad p_2 = \frac{\sigma^2\gamma_1\left(1 + \sum_{i=1}^{n_r}\|\mathbf{h}_{1i}^T\mathbf{A}_i\|^2\right)}{\left|\sum_{i=1}^{n_r}\mathbf{h}_{1i}^T\mathbf{A}_i\,\mathbf{h}_{2i}\right|^2}.$$

be combined with a two-dimensional search over the plane of (p_1, p_2) to find the optimal values of transceivers' transmit powers. In this section, we briefly review the technique of [14], while extending this technique to allow the optimization of transceivers' transmit powers.

Using (13), we can write the optimization problem (10) as

$$\min_{p_1,p_2,\{\mathbf{B}_i\}_{i=1}^{n_r}} \sum_{j=1}^{2} p_j \left(1+\sum_{i=1}^{n_r} \|\mathbf{B}_i \mathbf{q}_{ji}\|^2\right) + \sigma^2 \sum_{i=1}^{n_r} tr(\mathbf{B}_i\mathbf{B}_i^H)$$
$$\text{subject to} \quad \frac{p_j \left|\sum_{i=1}^{n_r} \mathbf{q}_{\bar{j}i}^T \mathbf{B}_i\, \mathbf{q}_{ji}\right|^2}{\sigma^2\left(1+\sum_{i=1}^{n_r}\|\mathbf{q}_{\bar{j}i}^T \mathbf{B}_i\|^2\right)} \geq \gamma_{\bar{j}}, \quad j \in \{1, 2\} \tag{34}$$

where $\mathbf{q}_{1i} = \mathbf{U}_i^H \mathbf{h}_{1i}$ and $\mathbf{q}_{2i} = \mathbf{U}_i^H \mathbf{h}_{2i}$ are defined as the effective channels between the i-th relay and Transceivers 1 and 2, respectively. Also we can rewrite the norms in problem (34) as $\|\mathbf{B}_i\, \mathbf{q}_{ji}\|^2 = \mathbf{b}_i^H(\mathbf{I}_2 \otimes (\mathbf{q}_{ji}^* \mathbf{q}_{ji}^T))\mathbf{b}_i$, $\|\mathbf{q}_{\bar{j}i}^T \mathbf{B}_i\|^2 = \mathbf{b}_i^H((\mathbf{q}_{\bar{j}i}^* \mathbf{q}_{\bar{j}i}^T)\otimes \mathbf{I}_2)\mathbf{b}_i$, and $tr(\mathbf{B}_i\mathbf{B}_i^H) = \mathbf{b}_i^H\mathbf{b}_i$, where we use the following definition: $\mathbf{b}_i \triangleq (vec(\mathbf{B}_i^T))^*$. Further, defining

$$\breve{\mathbf{f}} \triangleq [vec^T(\mathbf{q}_{21}\, \mathbf{q}_{11}^T) \cdots vec^T(\mathbf{q}_{2n_r}\, \mathbf{q}_{1n_r}^T)]^T$$

we can write

$$\left|\sum_{i=1}^{n_r} \mathbf{q}_{1i}^T \mathbf{B}_i\, \mathbf{q}_{2i}\right|^2 = \mathbf{b}^H \breve{\mathbf{f}}\breve{\mathbf{f}}^H\, \mathbf{b}.$$

The optimization problem (34) can now be recast as

$$\min_{p_1,p_2,\mathbf{b}} \; p_1 + p_2 + \mathbf{b}^H\,(p_1\mathbf{E}_0 + p_2\breve{\mathbf{E}}_0 + \sigma^2\mathbf{I}_{4n_r})\,\mathbf{b}$$
$$\text{subject to} \quad \mathbf{b}^H\,(p_2\,\breve{\mathbf{E}}_1 - \sigma^2\gamma_1\breve{\mathbf{E}}_2\,)\mathbf{b} \geq \sigma^2\gamma_1$$
$$\mathbf{b}^H\,(p_1\,\mathbf{E}_1 - \sigma^2\gamma_2\mathbf{E}_2\,)\mathbf{b} \geq \sigma^2\gamma_2 \tag{35}$$

where the following definitions are used:

$$\breve{\mathbf{E}}_0 \triangleq \text{blkdiag}\Big(\{\mathbf{I}_4 \otimes \mathbf{q}_{2i}\mathbf{q}_{2i}^H\}_{i=1}^{n_r}\Big),$$
$$\breve{\mathbf{E}}_1 \triangleq \breve{\mathbf{f}}\breve{\mathbf{f}}^H,$$
$$\breve{\mathbf{E}}_2 \triangleq \text{blkdiag}\Big(\{\mathbf{q}_{1i}\,\mathbf{q}_{1i}^H \otimes \mathbf{I}_4\}_{i=1}^{n_r}\Big).$$

The optimization problem (35) does not seem to be amenable to a closed-form solution. We can solve the problem by finding the optimal value for $\mathbf{b}$ for any given transceiver powers, p_1 and p_2, and then find the optimal values for p_1 and p_2 by finding those values of p_1 and p_2 which yield the smallest value for the objective function. For given values of p_1 and p_2, the minimization over $\mathbf{b}$ can be written as a quadratically constrained quadratic problem (QCQP). If the feasible region in (p_1, p_2) plane is quantized into a sufficiently fine grid, we can obtain the optimal value of $\mathbf{b}$ corresponding to each vertex of this grid. We then choose, as the solution to the problem, the values of p_1, p_2, and the corresponding value of $\mathbf{b}$, which lead to the minimum value of the objective function.

To solve the minimization over $\mathbf{b}$ for any given feasible pair of p_1 and p_2, we need to determine the set of feasible values of p_1 and p_2. One can see from the constraint in (35) that for those values of p_1 that make the matrix $(p_1\,\mathbf{E}_1 - \sigma^2\gamma_2\mathbf{E}_2)$ negative semi-definite, the problem becomes infeasible. Similar condition holds true for p_2 in matrix $(p_2\,\breve{\mathbf{E}}_1 - \sigma^2\gamma_2\breve{\mathbf{E}}_2)$. Hence, the infeasibility conditions can be written as

$$p_1\mathbf{E}_1 - \sigma^2\gamma_2\mathbf{E}_2 \preccurlyeq 0, \quad p_2\breve{\mathbf{E}}_1 - \sigma^2\gamma_1\breve{\mathbf{E}}_2 \preccurlyeq 0, \tag{36}$$

where the notation $\mathbf{Z} \preccurlyeq 0$ means that matrix $\mathbf{Z}$ is negative semi-definite. These conditions mean that the minimum values of p_1 and p_2 that make the problem feasible are those for which the largest eigenvalues of the matrices in (36) are greater than zero. It can be shown that the feasible values of p_1 and p_2 must satisfy

$$p_1 > \frac{\sigma^2\gamma_2}{\mathbf{q}_1^H\mathbf{q}_1}, \quad \text{and} \quad p_2 > \frac{\sigma^2\gamma_1}{\mathbf{q}_2^H\mathbf{q}_2}. \tag{37}$$

where $\mathbf{q}_1 \triangleq [\mathbf{q}_{11}^T, \mathbf{q}_{12}^T, \ldots, \mathbf{q}_{1n_r}^T]^T$, and $\mathbf{q}_2 \triangleq [\mathbf{q}_{21}^T, \mathbf{q}_{22}^T, \ldots, \mathbf{q}_{2n_r}^T]^T$. Hence, we need to start the exhaustive search over the values of p_1 and p_2 which satisfy (37). Let us consider the inner part of the minimization problem in (35) as

$$\min_{\mathbf{b}} \quad \mathbf{b}^H\,(p_1\mathbf{E}_0 + p_2\breve{\mathbf{E}}_0 + \sigma^2\mathbf{I}_{4n_r})\,\mathbf{b}$$
$$\text{subject to} \quad \mathbf{b}^H\,(p_2\breve{\mathbf{E}}_1 - \sigma^2\gamma_1\breve{\mathbf{E}}_2\,)\,\mathbf{b} \geq \sigma^2\gamma_1$$
$$\mathbf{b}^H\,(p_1\,\mathbf{E}_1 - \sigma^2\gamma_2\mathbf{E}_2\,)\,\mathbf{b} \geq \sigma^2\gamma_2 \tag{38}$$

Once a feasible pair of p_1 and p_2 is chosen, we can solve the minimization problem in (38), as explained in the sequel. Using the following definitions

$$\mathbf{G}_0 \triangleq (p_1\mathbf{E}_0 + p_2\breve{\mathbf{E}}_0 + \sigma^2\mathbf{I}_{4n_r})$$
$$\mathbf{G}_1 \triangleq \frac{p_1}{\sigma^2\gamma_2}\mathbf{E}_1 - \mathbf{E}_2, \quad \mathbf{G}_2 \triangleq \frac{p_2}{\sigma^2\gamma_1}\breve{\mathbf{E}}_1 - \breve{\mathbf{E}}_2$$

we can solve the problem using standard semi-definite program (SDP) tools [33]. Defining $\mathbf{X} \triangleq \mathbf{b}\mathbf{b}^H$, we can rewrite the problem in (38) as

$$\min_{\mathbf{X}} \; tr(\mathbf{G}_0\mathbf{X})$$
$$\text{s.t. } tr(\mathbf{G}_1\mathbf{X}) \geq 1, \; tr(\mathbf{G}_2\mathbf{X}) \geq 1, \; \text{rank}(\mathbf{X}) = 1, \mathbf{X} \succcurlyeq 0. \tag{39}$$

Due to the rank-one constraint, this problem is not convex but we can exploit a semi-definite relaxation (SDR) method to solve this problem [14]. *Interestingly enough, despite the relaxation, a rank-one solution to* (39) *exists and it can be extracted from the relaxed problem (for detailed procedure, refer to [34] and [35]).* This rank-one solution for $\mathbf{X}$ yields the optimal $\mathbf{b}$ for the problem in (38) for the chosen p_1 and p_2.

VI. REMARKS

The following remarks are in order.

Remark 1: In terms of computational complexity, the proposed symmetric beamforming technique involves finding the root of $g(\tilde{p}_1)$ using a simple bisection technique. In each iteration of this bisection technique, one has to find the unique positive root of (33) for a given value of $\tilde{p}_1$ using another simple bisection technique, thereby obtaining $\lambda(\tilde{p}_1)$. Both of these bisection methods converge very fast [36]. Considering that the number of iterations in these two bisection methods are insensitive to the problem size [36], the computational complexity of calculating $g(\tilde{p}_1)$ and $\lambda(\tilde{p}_1)$ is $\mathcal{O}(n_r)$. On the other hand, the general beamforming matrix based method involves solving an SDP problem at each vertex of the grid which covers the (p_1, p_2) plane. The computational complexity of solving an SDP problem at each of these vertices is $\mathcal{O}(n_r^4)$. Taking into account that the SDP problem has to be solved over all vertices, the computational complexity of the proposed algorithm is significantly lower than the SDP based solution. Indeed, the computational complexity of the combination of the SDP based technique and the exhaustive search method is prohibitively high, thereby justifying the use of the proposed method. In the next section, our numerical examples show that the performance loss caused by imposing symmetry on the relay beamforming matrices is negligible.

Remark 2: It is worth mentioning that the proposed scheme can be implemented in a distributed manner. To further explain this, the optimization problem (15) can be rewritten as

$$\min_{\mathbf{b}} \ \frac{\sigma^2(\gamma_1+\gamma_2)\left(1+\mathbf{b}^H\mathbf{E}_2\mathbf{b}\right)\left(1+\mathbf{b}^H\mathbf{E}_0\mathbf{b}\right)}{\mathbf{b}^H\mathbf{f}\mathbf{f}^H\mathbf{b}}+\sigma^2\mathbf{b}^H\mathbf{b}$$
$$\text{s.t.} \quad [\mathbf{b}]_{(i-1)n_r+2}=[\mathbf{b}]_{(i-1)n_r+3}, \quad \text{for } i=1,2,\ldots,n_r \tag{40}$$

or, equivalently, as

$$\min_{\tilde{\mathbf{b}}} \ \frac{\sigma^2(\gamma_1+\gamma_2)(1+\tilde{\mathbf{b}}^H\tilde{\mathbf{E}}_2\tilde{\mathbf{b}})(1+\tilde{\mathbf{b}}^H\tilde{\mathbf{E}}_0\tilde{\mathbf{b}})}{\tilde{\mathbf{b}}^H\tilde{\mathbf{f}}\tilde{\mathbf{f}}^H\tilde{\mathbf{b}}}+\sigma^2\tilde{\mathbf{b}}^H\mathbf{L}^H\mathbf{L}\tilde{\mathbf{b}}. \tag{41}$$

Differentiating the objective function of (41) with respect to $\tilde{p}_1$ and equating it to zero yields

$$\frac{1}{(\tilde{\mathbf{b}}^H\tilde{\mathbf{f}})}\tilde{\mathbf{f}}=\mathbf{Q}(\tilde{\mathbf{b}})\,\tilde{\mathbf{b}} \tag{42}$$

where the following definition is used:

$$\mathbf{Q}(\tilde{\mathbf{b}}) \triangleq \frac{\tilde{\mathbf{E}}_2}{(1+\tilde{\mathbf{b}}^H\,\tilde{\mathbf{E}}_2\,\tilde{\mathbf{b}})}+\frac{\tilde{\mathbf{E}}_0}{(1+\tilde{\mathbf{b}}^H\,\tilde{\mathbf{E}}_0\,\tilde{\mathbf{b}})} + \frac{1}{(\gamma_1+\gamma_2)}\frac{\mathbf{L}^H\mathbf{L}\,(\tilde{\mathbf{b}}^H\tilde{\mathbf{f}}\tilde{\mathbf{f}}^H\tilde{\mathbf{b}})}{(1+\tilde{\mathbf{b}}^H\,\tilde{\mathbf{E}}_2\,\tilde{\mathbf{b}})(1+\tilde{\mathbf{b}}^H\,\tilde{\mathbf{E}}_0\,\tilde{\mathbf{b}})}. \tag{43}$$

Further, defining $\mu_0 \triangleq (1+\tilde{\mathbf{b}}^H\,\tilde{\mathbf{E}}_0\,\tilde{\mathbf{b}})$, $\mu_2 \triangleq (1+\tilde{\mathbf{b}}^H\,\tilde{\mathbf{E}}_2\,\tilde{\mathbf{b}})$, and $\mu_1 \triangleq \tilde{\mathbf{b}}^H\tilde{\mathbf{f}}$, we can rewrite (42) as

$$\mu_0\,\mu_2\,\tilde{\mathbf{f}}=\left(\mu_0\tilde{\mathbf{E}}_2+\mu_2\tilde{\mathbf{E}}_0+\tfrac{|\mu_1|^2}{(\gamma_1+\gamma_2)}\mathbf{L}^H\mathbf{L}\right)\mu_1\tilde{\mathbf{b}}. \tag{44}$$

Since the matrix $\left(\mu_0\tilde{\mathbf{E}}_2+\mu_2\tilde{\mathbf{E}}_0+\tfrac{|\mu_1|^2}{(\gamma_1+\gamma_2)}\mathbf{L}^H\mathbf{L}\right)$ is invertible, we can obtain $\tilde{\mathbf{b}}$ as

$$\tilde{\mathbf{b}}=\frac{\mu_0\,\mu_2}{\mu_1}\left(\mu_0\tilde{\mathbf{E}}_2+\mu_2\tilde{\mathbf{E}}_0+\frac{|\mu_1|^2}{(\gamma_1+\gamma_2)}\mathbf{L}^H\mathbf{L}\right)^{-1}\tilde{\mathbf{f}}. \tag{45}$$

The fact that matrices $\tilde{\mathbf{E}}_0$, $\tilde{\mathbf{E}}_2$, and $\mathbf{L}^H\mathbf{L}$ are all block-diagonal matrices allows us to use (45) and write the optimal value of $\tilde{\mathbf{b}}_i$ for the i-th relay as

$$\tilde{\mathbf{b}}_i=\frac{\mu_0\mu_2}{\mu_1}\left(\mu_0(\tilde{\mathbf{E}}_2)_{(i)}+\mu_2(\tilde{\mathbf{E}}_0)_{(i)}+\frac{|\mu_1|^2}{(\gamma_1+\gamma_2)}(\mathbf{L}^H\mathbf{L})_{(i)}\right)^{-1}\tilde{\mathbf{f}}_i \tag{46}$$

where $(\tilde{\mathbf{E}}_2)_{(i)}$, $(\tilde{\mathbf{E}}_0)_{(i)}$, and $(\mathbf{L}^H\mathbf{L})_{(i)}$ are the i-th diagonal blocks of $\tilde{\mathbf{E}}_2$, $\tilde{\mathbf{E}}_0$, and $\mathbf{L}^H\mathbf{L}$, respectively. If one of the two transceivers broadcasts the three parameters μ_0, μ_1, and μ_2, the i-th relay can then use (46) to obtain its $\tilde{\mathbf{b}}_i$ vector from its local CSI. Indeed, the matrices $(\tilde{\mathbf{E}}_2)_{(i)}$, $(\tilde{\mathbf{E}}_0)_{(i)}$, and $(\mathbf{L}^H\mathbf{L})_{(i)}$ depend only on the local CSI of the i-th relay.

In terms of CSI acquisition, two scenarios can be implemented: 1) Due to the bidirectional nature of the communication, each transceiver (user device) can obtain all the channel coefficients through training, see for example [37]–[46]. Both transceivers can then obtain the parameter $\tilde{p}_1$ and consequently, calculate the vectorized version of the beamforming matrices as in (26), as well as find the transceivers' transmit powers from (11). One of the transceivers can then calculate the parameters μ_0, μ_1, and μ_2, and broadcast these parameters to all relays. Each relay can use these three parameters along with its local CSI as in (46) to obtain the vectorized version of its effective beamforming matrix. 2) In the second scenario, all relays (base stations) provide their CSI (which can be acquired using traditional training procedures) to one of the relays (main relay or main base station) through a back haul link (for example through an optical fiber link). The main relay can then use the global CSI to calculate the parameter $\tilde{p}_1$, and consequently, the vectorized version of the beamforming matrices as in (26), as well as the parameters μ_0, μ_1, and μ_2, and broadcast these parameters to other relays. Each relay can then use these three parameters along with its local CSI as in (46) to obtain the vectorized version of its effective beamforming matrix.

Remark 3: Note that the total power minimization approach utilized in this paper does not rely on individual per node power constraint. Adding such constraints can lead to the increase in the total power consumed in the entire network. As a result, it is recommended that the nodes hardware be designed to allow a relatively high amount of power consumption. Note also that it is reasonable to assume that the relay channel vectors are drawn from the same probability distribution function, and as a result, the long-term average transmit power of different relays will be the same. This is indeed what the numerical results of [6] showed for the case of two-way networks with multiple single-antenna relays.

Remark 4: It is also noteworthy that the relay beamforming matrices $\{\mathbf{A}_i\}_{i=1}^{n_r}$ can be written in terms of maximum ratio combining (MRC) and maximum ratio transmission (MRT) schemes. To show this, one can write $\mathbf{U}_i = [\mathbf{h}_{1i}\ \ \mathbf{h}_{2i}]\mathbf{W}_i$, where $\mathbf{W}_i$ is a 2×2 invertible matrix. As a result, using (13), the relay beamforming matrix can be written as $\mathbf{A}_i = [\mathbf{h}_{1i}^*\ \ \mathbf{h}_{2i}^*]\mathbf{W}_i^*\mathbf{B}_i\mathbf{W}_i^H[\mathbf{h}_{1i}\ \ \mathbf{h}_{2i}]^H$. Hence, the relay beamforming operation can be viewed as a cascade of an MRC operation, a multiplication of the MRC output with the matrix $\mathbf{W}_i^*\mathbf{B}_i\mathbf{W}_i^H$, and eventually an MRT scheme.

For very large M, i.e, in massive MIMO relaying schemes, where $\mathbf{h}_{1i}$ and $\mathbf{h}_{2i}$ are orthogonal, almost surely, one can easily show that at the optimum, matrix $\mathbf{C}_i \triangleq \mathbf{W}_i^*\mathbf{B}_i\mathbf{W}_i^H$ is anti-diagonal i.e., has zero diagonal entries. This means that self-interference will be zero. In this case, one still has to optimally obtain the two off-diagonal entries of matrix $\mathbf{C}_i$, To do so, one can show that we still need the same amount of CSI. The details of the derivations do not fit in the scope of this paper and we leave these details to future studies.

Remark 5: In this paper, we considered the network beamforming problem for a single-pair of transceivers. Designing network beamforming schemes to simultaneously establish communication between multiple pairs of transceivers in a peer-to-peer manner is yet another interesting problem. What we have done in this paper can be useful when considering a multi-pair sceneraio when the number of antennas at the relays is very large. Extending our result in this paper to a multi-pair scenario is possible but the details of such extension does not fit in the scope of this paper. We leave such an extension to our future work.

VII. SIMULATION RESULTS

In this section, we compare the performance of the proposed symmetric beamforming method, in terms of the total consumed power in the network, with the performance of the general beamforming method with no restriction on the beamforming matrices. We assume that the relays are randomly distributed between the two transceivers. Each transceiver-relay link is modeled as the product of three terms: a small-scale fading term (which is modeled as complex Gaussian random variables with zero mean and unit variance), a log-normal term with a standard deviation of 8 dB (which represents the shadowing effect), and a path loss component with a path loss exponent of 3.8. Also, the noise process in all nodes is assumed to be spatially white zero-mean Gaussian process with unit variance, i.e., $\sigma^2 = 1$.

Fig. 2 shows the average total transmit power, normalized to the noise power, versus equal SNR thresholds γ_1 for both the proposed symmetric beamforming method and the general beamforming technique in two different scenarios, i) $\gamma_2 = \gamma_1$ and ii) $\gamma_2 = \gamma_1/4$. As can be seen from this figure, in both scenarios, the total power required for satisfying the SNR constraints in the network with symmetric beamforming matrices is very close to the total power for the same network with general beamforming matrices, while the computational complexity of the symmetric beamforming method is significantly lower than that of the general beamforming method. As a result, assuming symmetric beamforming matrices offers computational saving with negligible performance loss, compared to the case when the beamforming matrices are not restricted to be symmetric. In the remainder of our simulation results, we focus on the proposed symmetric beamforming method.

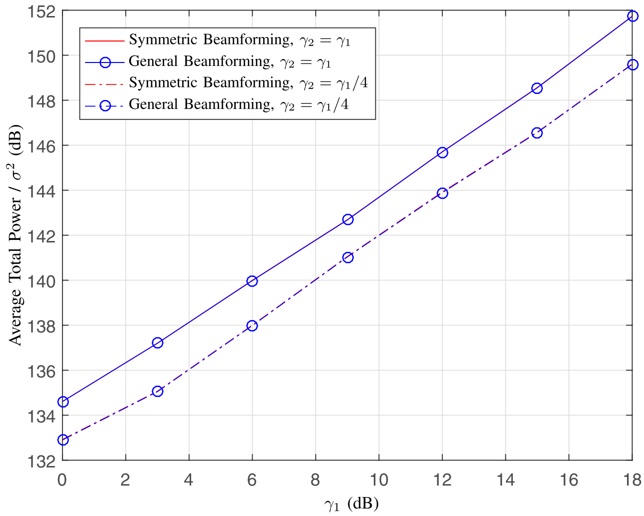

FIGURE 2. Average normalized total transmit power versus γ_1, for symmetric and general beamforming schemes, for $M = 4$ and $n_r = 4$.

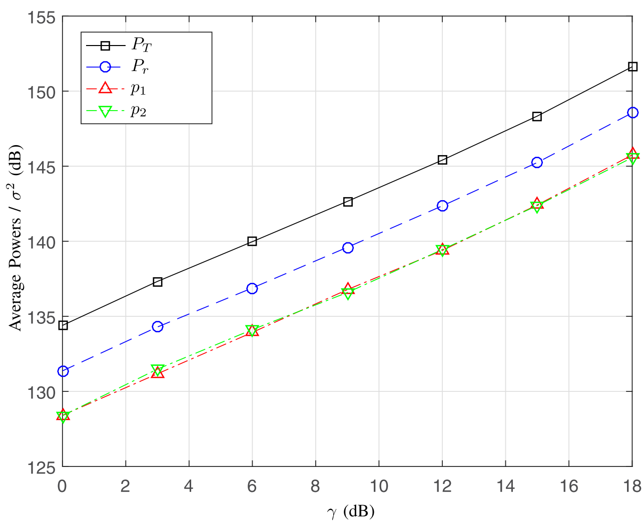

FIGURE 3. Average normalized total power, average normalized total relay power, and average normalized transceivers' transmit powers, versus $\gamma_1 = \gamma_2 = \gamma$, for $M = 4$ and $n_r = 4$.

Fig. 3 illustrates the average normalized values of the total consumed power in the network, the average normalized total transmit power of the relays, and the average normalized transceivers' transmit powers, versus equal SNR thresholds, i.e., $\gamma_1 = \gamma_2 \triangleq \gamma$, for a network consisting of two single-antenna transceivers and $n_r = 4$ relays each equipped with $M = 4$ antennas. As can be seen from this figure, the average total relay transmit power is 3 dB smaller than (i.e., half of) the average total transmit power consumed in the entire network. Although this figure shows averaged quantities, one can prove that *for any given set of channel realizations*, the total relay power is always half of the total transmit power consumed in the entire network, when $\gamma_1 = \gamma_2$. We can also observe from Fig. 3 that the average transmit power of each of the two transceivers are 6 dB lower than (or a quarter

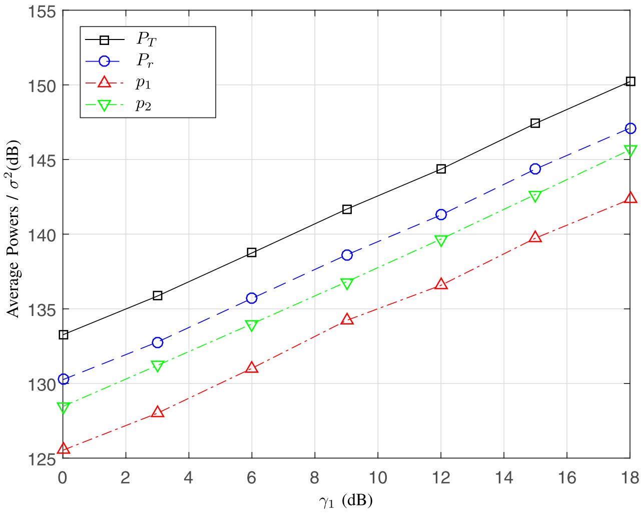

FIGURE 4. Average normalized total power, average normalized relay power, and average normalized transceivers' transmit powers, for non-equal SNR thresholds:$\gamma_2 = \gamma_1/2$, and for $M = 4$ and $n_r = 4$.

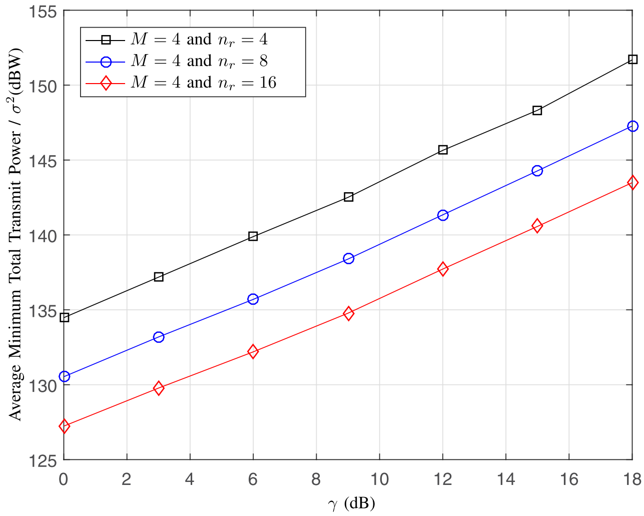

FIGURE 5. The average normalized total transmit power versus $\gamma_1 = \gamma_2 \triangleq \gamma$, for networks with different numbers of relays $n_r \in \{4, 8, 16\}$, and for $M = 4$.

of) the average total transmit power. Note however that this observation is correct only for average quantities and it may not hold true for a given channel realizations.

Fig. 4 shows the same quantities as in Fig. 3 for the case when we choose $\gamma_2 = \gamma_1/2$. As can be observed from this figure, the average total relay transmit power is about half of the average total transmit power consumed in the entire network. Note however that this observation is true for average quantities and may not hold for all channel realizations.

Note that in this paper, we did not consider per-node power constraints. Adding such constraints only shrinks the feasible set, and thus, increases the total power consumption. However, a guideline can be derived to choose the maximum average power consumption of each node. As shown in Fig. 3, under equal SNR thresholds, the power consumption of each of the two transceivers is 1/4 of the total power consumed in the entire network. Also, as the total relay power is half of the total transmit power, if the relay-transceiver channels are drawn from the same probability distribution function, then each relay node consumes, in average, $1/(2n_r)$ of the total transmit power.

Fig. 5 illustrates the normalized average minimum total transmit power for different number of relays each of which is equipped with $M = 4$ antennas. Fig. 6 illustrates the normalized average minimum total transmit power when $n_r = 4$ relays are equipped with 4, 8, and 16 antennas. As can be seen from Fig. 5, doubling number of relays, while keeping the number of antennas per relays unchanged, reduces the average minimum total transmit power by 2.98 to 3.94 dB over the depicted range of γ. Fig. 6 shows that doubling the number of antennas per relays, while keeping the number of relays unchanged, will reduce the minimum total transmit power by 2.91 to 3.13 dB over the chosen range of γ.

In Fig. 7, we plot the normalized average minimum total transmit power versus M, when the total number of the relay

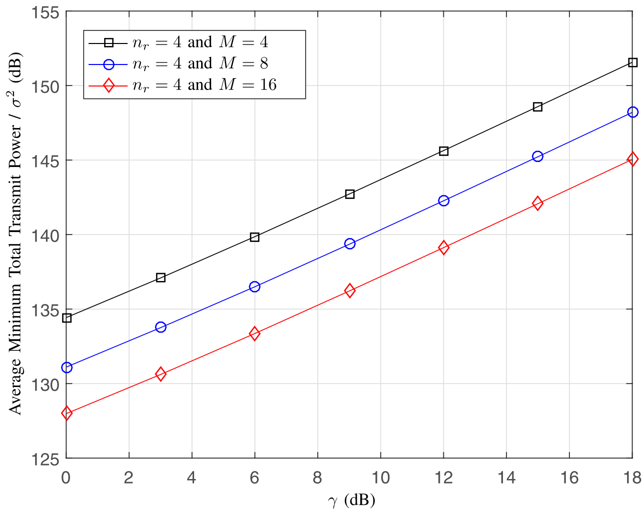

FIGURE 6. The normalized average minimum total transmit power, versus $\gamma_1 = \gamma_2 \triangleq \gamma$, for networks with $n_r = 4$, $M \in \{4, 8, 16\}$.

antennas employed in the network is constant ($Mn_r = 128$), for different values of γ. Interestingly, we observe that when $\gamma = 0$ dB is chosen, the minimum power will be achieved when $n_r = 16$ relays, each with $M = 8$ are used. As γ is increased to 10 dB, the minimum power can still be achieved when $n_r = 16$ relays, each with $M = 8$ are employed, Further increasing γ to 20 dB shows that the scenario with $n_r = 32$ relays, each with $M = 4$ antennas results in the minimum power consumption. In other words, when the SNR requirements are more stringent, the network should become "more distributed". This observation shows that there exists a trade-off between local beamforming at the relays and network beamforming distributed in the entire network. For low SNR requirements, local beamforming appears to be power-optimal while for high SNR requirements, network

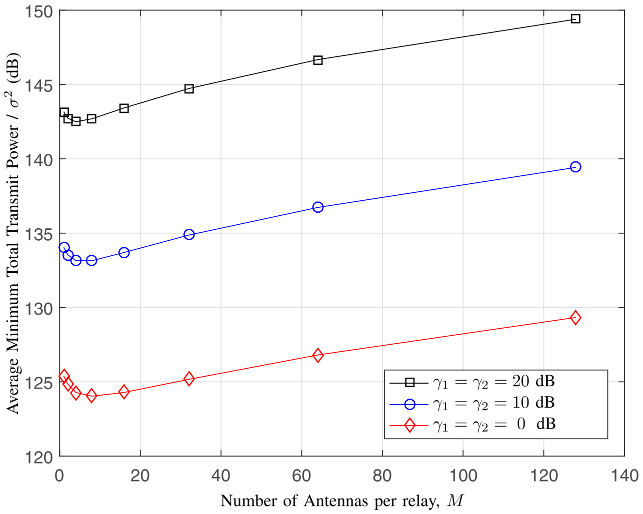

FIGURE 7. **The normalized average minimum total transmit power versus number of antennas per relay M, for $Mn_r = 128$ and for different values of γ.**

beamforming tends to be power-efficient. The theoretical justification/analysis of this trade-off is certainly an interesting research direction but it does not fit in the scope of this paper.

As shown in Fig. 5, for a *given number of antennas per relay*, increasing number of the relays consistently improves the performance of the proposed scheme. Also, Fig. 6 shows that *for a fixed number of relays*, increasing number of antennas per relay consistently improves the performance. However, *for a fixed number of total number of available antenna*, it appears from Fig. 7 that there exists an optimal number of antennas per relay, and thus an optimal number of relays, which lead to the best performance in terms of the total transmit power consumption. Finding the optimal number of relays and/or developing an optimal node selection strategy appears to be an interesting direction for future work on this topic.

VIII. CONCLUSIONS

In this paper, we studied the total transmit power minimization problem for a two-way relay network under two constraints on the transceivers' received signal-to-noise-ratios. The network we considered consists of multiple multi-antenna relay nodes and two single-antenna transceivers. Each relay transforms the vector of its received signals (by multiplying this vector with a complex ''beamforming'' matrix), thereby obtaining a new vector whose entries are transmitted over different antennas of that relay. Assuming the relay beamforming matrices and the transceivers' transceiver powers as the design parameters, we first considered the problem of total power minimization under the assumption that the relay beamforming matrices are symmetric. Under such an assumption, we showed that the total power minimization problem is amenable to a semi-closed-form solution, and thus, can be solved efficiently. We then considered the case where the relay beamforming matrices may not be symmetric and showed that in this case, the total power minimization problem can be solved using a computationally prohibitive algorithm which involves a two-dimensional search over a grid in the space of the transceivers' powers and semi-definite programming at each vertex of this grid. Our numerical results showed that the symmetric assumption on the relay beamforming matrices incurs only insignificant loss, while this assumption allows us to significantly reduce the computational burden of solving the total power minimization problem.

APPENDIX A
DERIVING THE FEASIBILITY CONDITION (25)

We observe from the constraint in (24) that for values of $\tilde{p}_1$ for which the matrix $(\tilde{p}_1\,\tilde{\mathbf{E}}_1 - \sigma^2(\gamma_1+\gamma_2)\tilde{\mathbf{E}}_2\,)$ is negative semi-definite, the problem becomes infeasible. Therefore, the infeasibility condition can be written as

$$\tilde{p}_1\,\mathbf{L}^H\mathbf{f}\mathbf{f}^H\mathbf{L} - \sigma^2(\gamma_1+\gamma_2)\mathbf{L}^H\mathbf{F}\,\mathbf{F}^H\mathbf{L} \preccurlyeq 0. \tag{47}$$

Here, we used the definition of matrix $\mathbf{E}_1$ in (19) along with the fact that matrix $\mathbf{E}_2$ in (21) can be written as $\mathbf{E}_2 = \mathbf{F}\,\mathbf{F}^H$, where the following definitions are used: $\mathbf{F} \triangleq \text{blkdiag}\{\mathbf{F}_1, \mathbf{F}_2, \ldots, \mathbf{F}_{n_r}\}$, $\mathbf{F}_i \triangleq [\,r_{1i}\mathbf{I}_2,\, r_{2i}\mathbf{I}_2\,]^T$, $r_{1i} \triangleq [\mathbf{q}_{2i}]_1$, and $r_{2i} \triangleq [\mathbf{q}_{2i}]_2$. Using these definitions, we can also write $\mathbf{f} = \mathbf{F}\mathbf{q}_1$, where $\mathbf{q}_1 \triangleq [\mathbf{q}_{11}^T, \mathbf{q}_{12}^T, \ldots, \mathbf{q}_{1n_r}^T]^T$. Hence, the infeasibility condition in (47) can be written as

$$\tilde{p}_1\,\mathbf{L}^H\mathbf{F}\,\mathbf{q}_1\,\mathbf{q}_1^H\,\mathbf{F}^H\mathbf{L} - \sigma^2(\gamma_1+\gamma_2)\mathbf{L}^H\mathbf{F}\,\mathbf{F}^H\mathbf{L} \preccurlyeq 0 \tag{48}$$

which is equivalent to the following condition on $\tilde{p}_1$:

$$\mathbf{L}^H\mathbf{F}\left(\tilde{p}_1\,\mathbf{q}_1\,\mathbf{q}_1^H - \sigma^2(\gamma_1+\gamma_2)\mathbf{I}_{2n_r}\right)\mathbf{F}^H\mathbf{L} \preccurlyeq 0. \tag{49}$$

We now argue that the condition in (49) is equivalent to the following condition:

$$\left(\tilde{p}_1\,\mathbf{q}_1\,\mathbf{q}_1^H - \sigma^2(\gamma_1+\gamma_2)\mathbf{I}_{2n_r}\right) \preccurlyeq 0. \tag{50}$$

It is obvious that if (50) holds true, then (49) also holds true. To show the reverse, we note that if (49) holds true, then for any $3n_r \times 1$ vector $\mathbf{z}$, we can write $\mathbf{z}^H\mathbf{L}^H\mathbf{F}\left(\tilde{p}_1\,\mathbf{q}_1\,\mathbf{q}_1^H - \sigma^2(\gamma_1+\gamma_2)\mathbf{I}_{2n_r}\right)\mathbf{F}^H\mathbf{L}\mathbf{z} < 0$. Since $\mathbf{F}^H\mathbf{L}$ is a fat matrix, the vector $\mathbf{F}^H\mathbf{L}\mathbf{z}$ can be any $2n_r \times 1$ vector. We hence conclude that the matrix $\tilde{p}_1\,\mathbf{q}_1\,\mathbf{q}_1^H - \sigma^2(\gamma_1+\gamma_2)\mathbf{I}_{2n_r}$ is negative semi-definite, i.e., (50) holds true. As a results, to find the feasible values of $\tilde{p}_1$, it is necessary and sufficient to find those values of $\tilde{p}_1$ which result in the largest eigenvalue of the matrix $\tilde{p}_1\,\mathbf{q}_1\,\mathbf{q}_1^H - \sigma^2(\gamma_1+\gamma_2)\mathbf{I}_{2n_r}$ being positive. The largest eigenvalue of this matrix is equal to $\tilde{p}_1\,\mathbf{q}_1^H\mathbf{q}_1 - \sigma^2(\gamma_1+\gamma_2)$. Hence, the problem in (24) is feasible if and only if

$$\tilde{p}_1 > \frac{\sigma^2(\gamma_1+\gamma_2)}{\mathbf{q}_1^H\mathbf{q}_1}. \tag{51}$$

The derivation of the feasibility condition is complete.

APPENDIX B
PROOF OF LEMMA 1

To prove that the objective function in (29), defined as $\psi(\tilde{p}_1) \triangleq \tilde{p}_1 + \frac{\sigma^2(\gamma_1+\gamma_2)}{\lambda(\tilde{p}_1)}$, has a unique extremum in the interval $(\frac{\sigma^2(\gamma_1+\gamma_2)}{\mathbf{q}_1^H\mathbf{q}_1}, +\infty)$, we first show that $\psi(\tilde{p}_1)$ approaches $+\infty$, either when $\tilde{p}_1 \rightarrow +\infty$ or when $\tilde{p}_1 \rightarrow \frac{\sigma^2(\gamma_1+\gamma_2)}{\mathbf{q}_1^H\mathbf{q}_1}$, and thus, $\psi(\tilde{p}_1)$ has at least one minimum in the interval $(\frac{\sigma^2(\gamma_1+\gamma_2)}{\mathbf{q}_1^H\mathbf{q}_1}, +\infty)$. We then show that this minimum is unique. Note that we can write

$$\begin{aligned}
&\lim_{\tilde{p}_1 \rightarrow \frac{\sigma^2(\gamma_1+\gamma_2)}{\mathbf{q}_1^H\mathbf{q}_1}} \mathbf{S}(\tilde{p}_1) \\
&= \left(\frac{\sigma^2(\gamma_1+\gamma_2)}{\mathbf{q}_1^H\mathbf{q}_1}\tilde{\mathbf{E}}_0 + \sigma^2\mathbf{L}^H\mathbf{L}\right)^{-1} \\
&\quad \times \left(\frac{\sigma^2(\gamma_1+\gamma_2)}{\mathbf{q}_1^H\mathbf{q}_1}\tilde{\mathbf{E}}_1 - \sigma^2(\gamma_1+\gamma_2)\tilde{\mathbf{E}}_2\right) \\
&= \left(\frac{\sigma^2(\gamma_1+\gamma_2)}{\|\mathbf{q}_1\|^2}\tilde{\mathbf{E}}_0 + \sigma^2\mathbf{L}^H\mathbf{L}\right)^{-1} \\
&\quad \times \left(\frac{\sigma^2(\gamma_1+\gamma_2)}{\|\mathbf{q}_1\|^2}\mathbf{L}^H\mathbf{F}\mathbf{q}_1\mathbf{q}_1^H\mathbf{F}^H\mathbf{L} - \sigma^2(\gamma_1+\gamma_2)\mathbf{L}^H\mathbf{F}\mathbf{F}^H\mathbf{L}\right) \\
&= \left(\frac{\sigma^2(\gamma_1+\gamma_2)}{\|\mathbf{q}_1\|^2}\tilde{\mathbf{E}}_0 + \sigma^2\mathbf{L}^H\mathbf{L}\right)^{-1} \\
&\quad \times \left(\sigma^2(\gamma_1+\gamma_2)\mathbf{L}^H\mathbf{F}\left(\frac{1}{\|\mathbf{q}_1\|^2}\mathbf{q}_1\mathbf{q}_1^H - \mathbf{I}_{2n_r}\right)\mathbf{F}^H\mathbf{L}\right)
\end{aligned} \tag{52}$$

It is obvious that the largest eigenvalue of matrix $\frac{1}{\mathbf{q}_1^H\mathbf{q}_1}\mathbf{q}_1\mathbf{q}_1^H - \mathbf{I}_{2n_r}$ in (52) is equal to zero. Hence, when $\tilde{p}_1 \rightarrow \frac{\sigma^2(\gamma_1+\gamma_2)}{\mathbf{q}_1^H\mathbf{q}_1}$, the largest eigenvalue of $\mathbf{S}(\tilde{p}_1)$, i.e., $\lambda(\tilde{p}_1)$ approaches 0, and thus, $\psi(\tilde{p}_1)$ approaches $+\infty$. It is also obvious that as $\tilde{p}_1$ approaches $+\infty$, the objective function $\psi(\tilde{p}_1)$ also approaches $+\infty$. Hence, $\psi(\tilde{p}_1)$ has at least one minimum in the interval $(\frac{\sigma^2(\gamma_1+\gamma_2)}{\mathbf{q}_1^H\mathbf{q}_1}, +\infty)$. We now prove that this minimum is the only extremum $\psi(\tilde{p}_1)$ can have. To this end, note that $\psi(\tilde{p}_1)$ is the sum of a monotonically increasing function (i.e., $\tilde{p}_1$) and the function $\frac{\sigma^2(\gamma_1+\gamma_2)}{\lambda(\tilde{p}_1)}$. If we can prove that $\lambda(\tilde{p}_1)$ is monotonically increasing with respect to $\tilde{p}_1$, we can then conclude that $\psi(\tilde{p}_1)$ has a unique minimum and the proof is then complete. We now prove that when $\tilde{p}_1 \in (\frac{\sigma^2(\gamma_1+\gamma_2)}{\mathbf{q}_1^H\mathbf{q}_1}, +\infty)$, $\lambda(\tilde{p}_1)$ is a monotonically increasing function of $\tilde{p}_1$. To prove this, in this interval, the derivative of $\lambda(\tilde{p}_1)$ with respect to $\tilde{p}_1$ is positive. The derivative of $\lambda(\tilde{p}_1)$ is obtained in Appendix C as (53)–(54), as shown at the bottom of this page, where, in the first inequality, we have used the fact that $\lambda(\tilde{p}_1)\tilde{\mathbf{f}}^H\mathbf{A}^{-1}(\tilde{p}_1)(\frac{\sigma^2}{\tilde{p}_1}\mathbf{L}^T\mathbf{L})\mathbf{A}^{-1}(\tilde{p}_1)\tilde{\mathbf{f}} > 0$, in the second inequality, we have used the fact that $\mathbf{E}_2$ is positive semi-definite, and in the last equality, we have used the fact that at optimum, as proven in Appendix C, $\tilde{\mathbf{f}}^H\mathbf{A}^{-1}(\tilde{p}_1)\tilde{\mathbf{f}} = \frac{1}{\tilde{p}_1}$ holds true. Hence, we conclude that $\frac{\partial\lambda(\tilde{p}_1)}{\partial\tilde{p}_1} > 0$ is positive, implying that $\lambda(\tilde{p}_1)$ is a monotonically increasing function of $\tilde{p}_1$. This completes the proof.

APPENDIX C

In what follows, we derive an expression for $\frac{\partial\lambda(\tilde{p}_1)}{\partial\tilde{p}_1}$. Since $\lambda(\tilde{p}_1)$ is the largest eigenvalue of the matrix $\mathbf{S}(\tilde{p}_1)$, we can write $\left(\mathbf{S}(\tilde{p}_1) - \lambda(\tilde{p}_1)\mathbf{I}_{3n_r}\right)\mathbf{u}(\tilde{p}_1) = \mathbf{0}$ which is equivalent to

$$\left((\tilde{p}_1\tilde{\mathbf{E}}_0 + \sigma^2\mathbf{L}^T\mathbf{L})^{-1}\left(\tilde{p}_1\tilde{\mathbf{E}}_1 - \sigma^2(\gamma_1+\gamma_2)\tilde{\mathbf{E}}_2\right) - \lambda(\tilde{p}_1)\mathbf{I}_{3n_r}\right) \times \mathbf{u}(\tilde{p}_1) = \mathbf{0} \tag{55}$$

where we use the definition of $\mathbf{S}(\tilde{p}_1)$ in (27). It follows from (55) that the matrix $\mathbf{S}(\tilde{p}_1) - \lambda(\tilde{p}_1)\mathbf{I}_{3n_r}$ has at least one zero eigenvalue. Multiplying (55) from left by $(\tilde{p}_1\tilde{\mathbf{E}}_0 + \sigma^2\mathbf{L}^T\mathbf{L})$, we arrive at

$$\left(\tilde{p}_1\tilde{\mathbf{E}}_1 - \sigma^2(\gamma_1+\gamma_2)\tilde{\mathbf{E}}_2 - \lambda(\tilde{p}_1)(\tilde{p}_1\tilde{\mathbf{E}}_0 + \sigma^2\mathbf{L}^T\mathbf{L})\right)\mathbf{u}(\tilde{p}_1) = \mathbf{0}. \tag{56}$$

Based on the fact that if $\tilde{p}_1 > \sigma^2(\gamma_1+\gamma_2)/\mathbf{q}_1^H\mathbf{q}_1$, then $\lambda(\tilde{p}_1) > 0$ holds true, and that the matrix $\mathbf{L}^T\mathbf{L}$ is full rank,

$$\begin{aligned}
\frac{\partial\lambda(\tilde{p}_1)}{\partial\tilde{p}_1} &= \frac{p_1^{-2} - \lambda(\tilde{p}_1)\tilde{\mathbf{f}}^H\mathbf{A}^{-1}(\tilde{p}_1)\tilde{\mathbf{E}}_0\mathbf{A}^{-1}(\tilde{p}_1)\tilde{\mathbf{f}}}{\tilde{\mathbf{f}}^H\mathbf{A}^{-1}(\tilde{p}_1)(\tilde{p}_1\tilde{\mathbf{E}}_0 + \sigma^2\mathbf{L}^T\mathbf{L})\mathbf{A}^{-1}(\tilde{p}_1)\tilde{\mathbf{f}}} \\
&> \frac{p_1^{-2} - \lambda(\tilde{p}_1)\tilde{\mathbf{f}}^H\mathbf{A}^{-1}(\tilde{p}_1)(\tilde{\mathbf{E}}_0 + \frac{\sigma^2}{\tilde{p}_1}\mathbf{L}^T\mathbf{L})\mathbf{A}^{-1}(\tilde{p}_1)\tilde{\mathbf{f}}}{\tilde{\mathbf{f}}^H\mathbf{A}^{-1}(\tilde{p}_1)(\tilde{p}_1\tilde{\mathbf{E}}_0 + \sigma^2\mathbf{L}^T\mathbf{L})\mathbf{A}^{-1}(\tilde{p}_1)\tilde{\mathbf{f}}}
\end{aligned} \tag{53}$$

$$\begin{aligned}
&= \frac{p_1^{-2}}{\tilde{\mathbf{f}}^H\mathbf{A}^{-1}(\tilde{p}_1)(\tilde{p}_1\tilde{\mathbf{E}}_0 + \sigma^2\mathbf{L}^T\mathbf{L})\mathbf{A}^{-1}(\tilde{p}_1)\tilde{\mathbf{f}}} - \frac{\lambda(\tilde{p}_1)}{\tilde{p}_1} = \frac{\lambda(\tilde{p}_1)}{\tilde{p}_1}\left(\frac{p_1^{-1}}{\tilde{\mathbf{f}}^H\mathbf{A}^{-1}(\tilde{p}_1)\lambda(\tilde{p}_1)(\tilde{p}_1\tilde{\mathbf{E}}_0 + \sigma^2\mathbf{L}^T\mathbf{L})\mathbf{A}^{-1}(\tilde{p}_1)\tilde{\mathbf{f}}} - 1\right) \\
&\geq \frac{\lambda(\tilde{p}_1)}{\tilde{p}_1} \times \left(\frac{p_1^{-1}}{\tilde{\mathbf{f}}^H\mathbf{A}^{-1}(\tilde{p}_1)\left(\sigma^2(\gamma_1+\gamma_2)\tilde{\mathbf{E}}_2 + \lambda(\tilde{p}_1)(\tilde{p}_1\tilde{\mathbf{E}}_0 + \sigma^2\mathbf{L}^T\mathbf{L})\right)\mathbf{A}^{-1}(\tilde{p}_1)\tilde{\mathbf{f}}} - 1\right) \\
&= \frac{\lambda(\tilde{p}_1)}{\tilde{p}_1}\left(\frac{p_1^{-1}}{\tilde{\mathbf{f}}^H\mathbf{A}^{-1}(\tilde{p}_1)\mathbf{A}(\tilde{p}_1)\mathbf{A}^{-1}(\tilde{p}_1)\tilde{\mathbf{f}}} - 1\right) = \frac{\lambda(\tilde{p}_1)}{\tilde{p}_1}\left(\frac{p_1^{-1}}{\tilde{\mathbf{f}}^H\mathbf{A}^{-1}(\tilde{p}_1)\tilde{\mathbf{f}}} - 1\right) = 0
\end{aligned} \tag{54}$$

we conclude that the matrix

$$\mathbf{A}(\tilde{p}_1) \triangleq \sigma^2(\gamma_1+\gamma_2)\tilde{\mathbf{E}}_2 + \lambda(\tilde{p}_1)(\tilde{p}_1\tilde{\mathbf{E}}_0 + \sigma^2\mathbf{L}^T\mathbf{L}) \tag{57}$$

is nonsingular, and hence, $\mathbf{A}^{-1}(\tilde{p}_1)$ exists. As a result, we can write (56) as

$$(\tilde{p}_1\mathbf{A}^{-1}(\tilde{p}_1)\tilde{\mathbf{E}}_1 - \mathbf{I}_{3n_r})\mathbf{u}(\tilde{p}_1) = \mathbf{0}. \tag{58}$$

In light of (58), we observe that the matrix $\tilde{p}_1\mathbf{A}^{-1}(\tilde{p}_1)\tilde{\mathbf{E}}_1 - \mathbf{I}_{3n_r}$ must have at least one zero eigenvalue. Defining $\tilde{\mathbf{f}} \triangleq \mathbf{L}^H\mathbf{f}$, we can write $\tilde{\mathbf{E}}_1 = \tilde{\mathbf{f}}\tilde{\mathbf{f}}^H$, which is a rank-one matrix. Hence, the matrix $\mathbf{A}^{-1}(\tilde{p}_1)\tilde{\mathbf{E}}_1$ is also rank-one. Therefore, all the eigenvalues of the matrix $\tilde{p}_1\mathbf{A}^{-1}(\tilde{p}_1)\tilde{\mathbf{E}}_1 - \mathbf{I}_{3n_r}$ are equal to -1, except the largest eigenvalue which is given by $\tilde{p}_1\tilde{\mathbf{f}}^H\mathbf{A}^{-1}(\tilde{p}_1)\tilde{\mathbf{f}} - 1$. Thus, equating this largest eigenvalue to 0 yields

$$\tilde{\mathbf{f}}^H\mathbf{A}^{-1}(\tilde{p}_1)\tilde{\mathbf{f}} = \frac{1}{\tilde{p}_1}. \tag{59}$$

Differentiating both sides of (59) with respect to $\tilde{p}_1$ yields

$$\tilde{\mathbf{f}}^H\mathbf{A}^{-1}(\tilde{p}_1)\frac{\partial\mathbf{A}(\tilde{p}_1)}{\partial\tilde{p}_1}\mathbf{A}^{-1}(\tilde{p}_1)\tilde{\mathbf{f}} = \frac{1}{\tilde{p}_1^2}. \tag{60}$$

We now use the fact that

$$\frac{\partial\mathbf{A}^{-1}(\tilde{p}_1)}{\partial\tilde{p}_1} = -\mathbf{A}^{-1}(\tilde{p}_1)\frac{\partial\mathbf{A}(\tilde{p}_1)}{\partial\tilde{p}_1}\mathbf{A}^{-1}(\tilde{p}_1)$$

and that

$$\frac{\partial\mathbf{A}(\tilde{p}_1)}{\partial\tilde{p}_1} = \frac{\partial\lambda(\tilde{p}_1)}{\partial\tilde{p}_1}(\tilde{p}_1\tilde{\mathbf{E}}_0 + \sigma^2\mathbf{L}^T\mathbf{L}) + \lambda(\tilde{p}_1)\tilde{\mathbf{E}}_0$$

to rewrite (60) as

$$\frac{\partial\lambda(\tilde{p}_1)}{\partial\tilde{p}_1}\tilde{\mathbf{f}}^H\mathbf{A}^{-1}(\tilde{p}_1)(\tilde{p}_1\tilde{\mathbf{E}}_0 + \sigma^2\mathbf{L}^T\mathbf{L})\mathbf{A}^{-1}(\tilde{p}_1)\tilde{\mathbf{f}} + \lambda(\tilde{p}_1)\tilde{\mathbf{f}}^H\mathbf{A}^{-1}(\tilde{p}_1)\tilde{\mathbf{E}}_0\mathbf{A}^{-1}(\tilde{p}_1)\tilde{\mathbf{f}} = \frac{1}{\tilde{p}_1^2}. \tag{61}$$

Therefore, we arrive at

$$\frac{\partial\lambda(\tilde{p}_1)}{\partial\tilde{p}_1} = \frac{\tilde{p}_1^{-2} - \lambda(\tilde{p}_1)\tilde{\mathbf{f}}^H\mathbf{A}^{-1}(\tilde{p}_1)\tilde{\mathbf{E}}_0\mathbf{A}^{-1}(\tilde{p}_1)\tilde{\mathbf{f}}}{\tilde{\mathbf{f}}^H\mathbf{A}^{-1}(\tilde{p}_1)(\tilde{p}_1\tilde{\mathbf{E}}_0 + \sigma^2\mathbf{L}^T\mathbf{L})\mathbf{A}^{-1}(\tilde{p}_1)\tilde{\mathbf{f}}}. \tag{62}$$

By substituting (62) in (30), we can write

$$g(\tilde{p}_1) = 1 - \sigma^2(\gamma_1+\gamma_2) \times \frac{\tilde{p}_1^{-2} - \lambda(\tilde{p}_1)\tilde{\mathbf{f}}^H\mathbf{A}^{-1}(\tilde{p}_1)\tilde{\mathbf{E}}_0\mathbf{A}^{-1}(\tilde{p}_1)\tilde{\mathbf{f}}}{\lambda^2(\tilde{p}_1)\tilde{\mathbf{f}}^H\mathbf{A}^{-1}(\tilde{p}_1)(\tilde{p}_1\tilde{\mathbf{E}}_0+\sigma^2\mathbf{L}^T\mathbf{L})\mathbf{A}^{-1}(\tilde{p}_1)\tilde{\mathbf{f}}} \tag{63}$$

Equating $g(\tilde{p}_1)$ to 0 does not yield a closed-form solution when $n_r > 1$, or when $M > 1$. However, the solution to the equation $g(\tilde{p}_1) = 0$ can be obtained using a simple Newton-Raphson method or a bisection technique. Note that in order to calculate $g(\tilde{p}_1)$ as in (63), one needs to calculate $\lambda(\tilde{p}_1)$ for each value of $\tilde{p}_1$. To calculate $\lambda(\tilde{p}_1)$, we plug $\mathbf{A}(\tilde{p}_1)$ from (57) into (59) and arrive at the following equality:

$$\tilde{p}_1\tilde{\mathbf{f}}^H\left(\sigma^2(\gamma_1+\gamma_2)\tilde{\mathbf{E}}_2 + \lambda(\tilde{p}_1)(\tilde{p}_1\tilde{\mathbf{E}}_0 + \sigma^2\mathbf{L}^T\mathbf{L})\right)^{-1}\tilde{\mathbf{f}} = 1. \tag{64}$$

which can be used to obtain $\lambda(\tilde{p}_1)$ for every feasible value of $\tilde{p}_1$. We now prove that (64) yields a unique value for $\lambda(\tilde{p}_1)$ for any given feasible value of $\tilde{p}_1$. To do so, we first observe that the function

$$\hbar(z) \triangleq \tilde{p}_1\tilde{\mathbf{f}}^H\left(\sigma^2(\gamma_1+\gamma_2)\tilde{\mathbf{E}}_2 + z(\tilde{p}_1\tilde{\mathbf{E}}_0 + \sigma^2\mathbf{L}^T\mathbf{L})\right)^{-1}\tilde{\mathbf{f}} \tag{65}$$

is monotonically decreasing in z, for any feasible value of $\tilde{p}_1$. We then observe that $\lim_{z\to+\infty}\hbar(z) = 0$. Hence, if we show that $\lim_{z\to 0}\hbar(z) > 1$ holds true for any feasible value of $\tilde{p}_1$, we can conclude that for any feasible value of $\tilde{p}_1$, the equation $\hbar(z) = 1$ has a unique solution, so does (64). To show that for any feasible value of $\tilde{p}_1$, we have $\lim_{z\to 0}\hbar(z) > 1$, we note that $\hbar(z) - 1 = \tilde{p}_1\tilde{\mathbf{f}}^H\left(\sigma^2(\gamma_1+\gamma_2)\tilde{\mathbf{E}}_2 + z(\tilde{p}_1\tilde{\mathbf{E}}_0 + \sigma^2\mathbf{L}^T\mathbf{L})\right)^{-1}\tilde{\mathbf{f}} - 1$ is the largest eigenvalue of the matrix $\mathbf{S}(\tilde{p}_1) - z\mathbf{I}_{3n_r}$. It is obvious that when $z \to 0$, the largest eigenvalue of matrix $\mathbf{S}(\tilde{p}_1) - z\mathbf{I}_{3n_r}$ approaches the largest eigenvalue of the matrix $\mathbf{S}(\tilde{p}_1) = (\tilde{p}_1\tilde{\mathbf{E}}_0 + \sigma^2\mathbf{L}^H\mathbf{L})^{-1}(\tilde{p}_1\tilde{\mathbf{E}}_1 - \sigma^2(\gamma_1+\gamma_2)\tilde{\mathbf{E}}_2)$. Note however that the largest eigenvalue of the matrix $\mathbf{S}(\tilde{p}_1)$ is larger than zero. Otherwise, if the largest eigenvalue of the matrix $\mathbf{S}(\tilde{p}_1)$ is smaller than, or equal to zero, as the matrix $\tilde{p}_1\tilde{\mathbf{E}}_0 + \sigma^2\mathbf{L}^T\mathbf{L}$ is positive definite, the matrix $(\tilde{p}_1\tilde{\mathbf{E}}_1 - \sigma^2(\gamma_1+\gamma_2)\tilde{\mathbf{E}}_2)$ will be non-positive definite for feasible values of $\tilde{p}_1$. This is indeed a contradiction, as for feasible values of $\tilde{p}_1$, i.e., when $\tilde{p}_1 > \sigma^2(\gamma_1+\gamma_2)/\mathbf{q}_1^H\mathbf{q}_1$, the matrix $(\tilde{p}_1\tilde{\mathbf{E}}_1 - \sigma^2(\gamma_1+\gamma_2)\tilde{\mathbf{E}}_2)$ may not be non-positive definite. Hence, the largest eigenvalue of the matrix $\mathbf{S}(\tilde{p}_1)$ is larger than zero, so is $\lim_{z\to 0}\hbar(z) - 1$. As a result, $\lim_{z\to 0}\hbar(z) > 1$, and hence, (64) has unique solution in terms of $\lambda(\tilde{p}_1)$.

REFERENCES

[1] Y. Wu, P. A. Chou, and S.-Y. Kung, "Information exchange in wireless networks with network coding and physical-layer broadcast," in *Proc. 39th Annu. Conf. Inf. Sciences Syst. (CISS)*, Mar. 2005, p. 1.

[2] M. Zaeri-Amirani, S. Shahbazpanahi, T. Mirfakhraie, and K. Ozdemir, "Performance tradeoffs in amplify-and-forward bidirectional network beamforming," *IEEE Trans. Signal Process.*, vol. 60, no. 8, pp. 4196–4209, Aug. 2012.

[3] I. Hammerstrom and A. Wittneben, "Power allocation schemes for amplify-and-forward MIMO-OFDM relay links," *IEEE Trans. Wireless Commun.*, vol. 6, no. 8, pp. 2798–2802, Aug. 2007.

[4] C. Yuen, W. H. Chin, Y. L. Guan, W. Chen, and T. Tee, "Bi-directional multi-antenna relay communications with wireless network coding," in *Proc. IEEE Veh. Technol. Conf., (VTC-Spring)*, Singapore, May 2008, pp. 1385–1388.

[5] T. Cui and J. Kliewer, "Memoryless relay strategies for two-way relay channels: Performance analysis and optimization," in *Proc. IEEE Int. Conf. Commun.*, Beijing, China, May 2008, pp. 1139–1143.

[6] 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.

[7] A. Khabbazibasmenj, F. Roemer, S. A. Vorobyov, and M. Haardt, "Sum-rate maximization in two-way AF MIMO relaying: Polynomial time solutions to a class of DC programming problems," *IEEE Trans. Signal Process.*, vol. 60, no. 10, pp. 5478–5493, Oct. 2012.

[8] M. Eslamifar, C. Yuen, W. H. Chin, and Y. L. Guan, "Max-min antenna selection for bi-directional multi-antenna relaying," in *Proc. IEEE Veh. Technol. Conf.*, May 2010, pp. 1–5.

[9] E. Yilmaz, R. Zakhour, D. Gesbert, and R. Knopp, "Multi-pair two-way relay channel with multiple antenna relay station," in *Proc. IEEE Int. Conf. Commun.*, May 2010, pp. 1–5.

[10] S. Xu and Y. Hua, "Optimal design of spatial source-and-relay matrices for a non-regenerative two-way MIMO relay system," *IEEE Trans. Wireless Commun.*, vol. 10, no. 5, pp. 1645–1655, May 2011.

[11] F. Roemer and M. Haardt, "Algebraic norm-maximizing (ANOMAX) transmit strategy for two-way relaying with MIMO amplify and forward relays," *IEEE Signal Process. Lett.*, vol. 16, no. 10, pp. 909–912, Oct. 2009.

[12] R. Wang and M. Tao, "Joint source and relay precoding designs for MIMO two-way relaying based on MSE criterion," *IEEE Trans. Signal Process.*, vol. 60, no. 3, pp. 1352–1365, Mar. 2012.

[13] J. Zhang and M. Haardt, "Energy efficient two-way non-regenerative relaying for relays with multiple antennas," *IEEE Signal Process. Lett.*, vol. 22, no. 8, pp. 1079–1083, Aug. 2015.

[14] R. Zhang, Y. C. Liang, C. C. Chai, and S. Cui, "Optimal beamforming for two-way multi-antenna relay channel with analogue network coding," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 5, pp. 699–712, Jun. 2009.

[15] K.-J. Lee, K. W. Lee, H. Sung, and I. Lee, "Sum-rate maximization for two-way MIMO amplify-and-forward relaying systems," in *Proc. IEEE Veh. Technol. Conf.*, Apr. 2009, pp. 1–5.

[16] K.-J. Lee, H. Sung, E. Park, and I. Lee, "Joint optimization for one and two-way MIMO AF multiple-relay systems," *IEEE Trans. Wireless Commun.*, vol. 9, no. 12, pp. 3671–3681, Dec. 2010.

[17] R. Vaze and R. W. Heath, "Capacity scaling for MIMO two-way relaying," in *Proc. IEEE Int. Symp. Inf. Theory, (ISIT)*, Jun. 2007, pp. 1451–1455.

[18] D. Gunduz, A. Goldsmith, and H. Poor, "MIMO two-way relay channel: Diversity-multiplexing tradeoff analysis," in *Proc. Asilomar Conf. Signals Syst. Comput. (ASILOMAR)*, Pacific Grove, CA, USA, Oct. 2008, pp. 1474–1478.

[19] H. H. Kha, H. D. Tuan, H. H. Nguyen, and H. H. M. Tam, "Joint design of user power allocation and relay beamforming in two-way MIMO relay networks," in *Proc. Int. Conf. Signal Process. Commun. Syst.*, Gold Coast, Australia, Dec. 2013, pp. 1–6.

[20] A. Alsharoa, H. Ghazzai, and M. S. Alouini, "Energy efficient design for MIMO two-way AF multiple relay networks," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2014, pp. 1007–1011.

[21] Y. Rong, "Joint source and relay optimization for two-way linear non-regenerative MIMO relay communications," *IEEE Trans. Signal Process.*, vol. 60, no. 12, pp. 6533–6546, Dec. 2012.

[22] L. P. Qian, Y. Wu, and Q. Chen, "Transmit power minimization for outage-constrained relay selection over Rayleigh-fading channels," *IEEE Commun. Lett.*, vol. 18, no. 8, pp. 1383–1386, Aug. 2014.

[23] T. Wang, Y. Fang, and L. Vandendorpe, "Power minimization for OFDM transmission with subcarrier-pair based opportunistic DF relaying," *IEEE Commun. Lett.*, vol. 17, no. 3, pp. 471–474, Mar. 2013.

[24] A. P. T. Lau and S. Cui, "Joint power minimization in wireless relay channels," *IEEE Trans. Wireless Commun.*, vol. 6, no. 8, pp. 2820–2824, Aug. 2007.

[25] D. H. N. Nguyen and H. H. Nguyen, "Power allocation in wireless multiuser multi-relay networks with distributed beamforming," *IET Commun.*, vol. 5, no. 14, pp. 2040–2051, Sep. 2011.

[26] S. Shahbazpanahi and M. Dong, "A semi-closed-form solution to optimal distributed beamforming for two-way relay networks," *IEEE Trans. Signal Process.*, vol. 60, no. 3, pp. 1511–1516, Mar. 2012.

[27] H. Chen, A. Gershman, and S. Shahbazpanahi, "Filter-and-forward distributed beamforming in relay networks with frequency selective fading," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1251–1262, Mar. 2010.

[28] S. Talwar, Y. Jing, and S. Shahbazpanahi, "Joint relay selection and power allocation for two-way relay networks," *IEEE Signal Process. Lett.*, vol. 18, no. 2, pp. 91–94, Feb. 2011.

[29] M. H. Golbon-Haghighi, M. Shirazi, B. Mahboobi, and M. Ardebilipour, "Optimal beamforming in wireless multiuser MIMO-relay networks," in *Proc. 21st Iranian Conf. Electr. Eng. (ICEE)*, May 2013, pp. 1–5.

[30] Y. Shi, J. Zhang, and K. B. Letaief, "Coordinated relay beamforming for amplify-and-forward two-hop interference networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2012, pp. 2408–2413.

[31] C. Thron and A. Aziz, "Algebraic method for optimal beamforming in two-way relay systems with analog network coding," in *Proc. IEEE Int. Symp. Signal Process. Inf. Technol. (ISSPIT)*, Dec. 2015, pp. 646–651.

[32] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-constrained modulation optimization," *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 2349–2360, Sep. 2005.

[33] M. Grant and S. Boyd. (Mar. 2014). *CVX: MATLAB Software for Disciplined Convex Programming, Version 2.1*. [Online]. Available: http://cvxr.com/cvx

[34] W. Ai, Y. Huang, and S. Zhang, "New results on Hermitian matrix rank-one decomposition," *Math. program., Ser. A*, vol. 128, nos. 1–2, pp. 253–283, Jun. 2011.

[35] Y. Ye and S. Zhang, "New results on quadratic minimization," *SIAM J. Optim.*, vol. 14, no. 1, pp. 245–267, 2003.

[36] S. Boyd, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[37] S. Zhang, F. Gao, C. Pei, and X. He, "Segment training based individual channel estimation in one-way relay network with power allocation," *IEEE Trans. Wireless Commun.*, vol. 12, no. 3, pp. 1300–1309, Mar. 2013.

[38] F. Gao, R. Zhang, and Y. C. Liang, "Channel estimation for OFDM modulated two-way relay networks," *IEEE Trans. Signal Process.*, vol. 57, no. 11, pp. 4443–4455, Nov. 2009.

[39] F. Gao, T. Cui, and A. Nallanathan, "Optimal training design for channel estimation in decode-and-forward relay networks with individual and total power constraints," *IEEE Trans. Signal Process.*, vol. 56, no. 12, pp. 5937–5949, Dec. 2008.

[40] F. Gao, T. Cui, and A. Nallanathan, "On channel estimation and optimal training design for amplify and forward relay networks," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 1907–1916, May 2008.

[41] F. Gao, B. Jiang, X. Gao, and X. D. Zhang, "Superimposed training based channel estimation for OFDM modulated amplify-and-forward relay networks," *IEEE Trans. Commun.*, vol. 59, no. 7, pp. 2029–2039, Jul. 2011.

[42] B. Jiang, F. Gao, X. Gao, and A. Nallanathan, "Channel estimation and training design for two-way relay networks with power allocation," *IEEE Trans. Wireless Commun.*, vol. 9, no. 6, pp. 2022–2032, Jun. 2010.

[43] G. Wang, F. Gao, Y.-C. Wu, and C. Tellambura, "Joint cfo and channel estimation for OFDM-based two-way relay networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 2, pp. 456–465, Feb. 2011.

[44] S. Zhang, F. Gao, and C. X. Pei, "Optimal training design for individual channel estimation in two-way relay networks," *IEEE Trans. Signal Process.*, vol. 60, no. 9, pp. 4987–4991, Sep. 2012.

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

[46] G. Wang, F. Gao, W. Chen, and C. Tellambura, "Channel estimation and training design for two-way relay networks in time-selective fading environments," *IEEE Trans. Commun.*, vol. 10, no. 8, pp. 2681–2691, Aug. 2011.

RAZGAR RAHIMI was born in Saqqez, Iran. He received the B.Sc. degree in electrical engineering from the Iran University of Science and Technology in 2005, and the M.Sc. degree in electrical engineering from Shahed University, Tehran, Iran, in 2010. He is currently pursuing the Ph.D. degree with the University of Ontario Institute of Technology, Oshawa, Canada. His recent research interests include signal processing, cognitive radio networks, and cooperative wireless communications.

SHAHRAM SHAHBAZPANAHI (M'02–SM'10) was born in Sanandaj, Iran. He received the B.Sc., M.Sc., and Ph.D. degrees in electrical engineering from the Sharif University of Technology, Tehran, Iran, in 1992, 1994, and 2001, respectively. From 1994 to 1996, he was a Faculty Member with the Department of Electrical Engineering, Razi University, Kermanshah, Iran. From 2001 to 2003, he was a Post-Doctoral Fellow with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada. From 2003 to 2004, he was a Visiting Researcher with the Department of Communication Systems, University of Duisburg-Essen, Duisburg, Germany. From 2004 to 2005, he was a Lecturer and an Adjunct Professor with the Department of Electrical and Computer Engineering, McMaster University. In 2005, he joined the Faculty of Engineering and Applied Science, University of Ontario Institute of Technology, Oshawa, ON, Canada, where he currently holds a full professor position. His research interests include statistical and array signal processing, space-time adaptive processing, detection and estimation, multi-antenna, multi-user, and cooperative communications; spread spectrum techniques, DSP programming, and hardware/real-time software design for telecommunication systems. He also served as an Elected Member of the Sensor Array and Multichannel Technical Committee of the IEEE Signal Processing Society. He has received several awards, including the Early Researcher Award from Ontario's Ministry of Research and Innovation, the NSERC Discovery Grant (three awards), the Research Excellence Award from the Faculty of Engineering and Applied Science, the University Of Ontario Institute Of Technology, and the Research Excellence Award, Early Stage, from the University of Ontario Institute of Technology. He has served as an Associate Editor of the IEEE Transactions on Signal Processing and the IEEE Signal Processing Letters. He served as a Senior Area Editor of the IEEE Signal Processing Letters.