

Sum-Rate Maximization With Minimum Power Consumption for MIMO DF Two-Way Relaying— Part II: Network Optimization

Jie Gao, *Student Member, IEEE*, Sergiy A. Vorobyov, *Senior Member, IEEE*, Hai Jiang, *Member, IEEE*, Jianshu Zhang, *Student Member, IEEE*, and Martin Haardt, *Senior Member, IEEE*

Abstract—In Part II of this two-part paper, a sum-rate-maximizing power allocation with minimum power consumption is derived for multiple-input multiple-output (MIMO) decode-and-forward (DF) two-way relaying (TWR) in a network optimization scenario. In this scenario, the relay and the source nodes jointly optimize their power allocation strategies to achieve network optimality. Unlike the relay optimization scenario considered in Part I, which features low complexity but does not achieve network optimality, the network-level optimal power allocation can be achieved in the network optimization scenario at the cost of higher complexity. The network optimization problem is considered in two cases each with several subcases. It is shown that the considered problem, which is originally nonconvex, can be transferred into different convex problems for all but two subcases. For the remaining two subcases, one for each case, it is proved that the optimal strategies for the source nodes and the relay must satisfy certain properties. Based on these properties, an algorithm is proposed for finding the optimal solution. The effect of asymmetry in the power limits, number of antennas, and channels is also considered. Such asymmetry is shown to have a negative effect on both the achievable sum-rate and the power allocation efficiency in MIMO DF TWR. Simulation results demonstrate the performance of the proposed algorithm and the effect of asymmetry in the system.

Index Terms—MIMO DF two-way relaying, joint power allocation, sum-rate maximization with minimum power consumption, asymmetry.

I. INTRODUCTION

TWO-WAY RELAYING (TWR) is a promising protocol featuring high spectral efficiency [1]. Optimizing transmit strategies such as power allocation of the participating nodes

Manuscript received August 18, 2012; revised December 14, 2012; accepted January 31, 2013. Date of publication May 16, 2013; date of current version June 24, 2013. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Shuguang Cui. This work was supported in parts by the Natural Science and Engineering Research Council (NSERC) of Canada and the Carl Zeiss Award, Germany. This work was partly funded by the European Union FP7-ICT project EMPhAtiC under grant agreement no. 318362.

J. Gao and H. Jiang are with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6G 2V4, Canada (e-mail: jgao3@ualberta.ca; hai1@ualberta.ca).

S. A. Vorobyov is with the Department of Signal Processing and Acoustics, Aalto University, FI-00076 AALTO, Finland, on leave from the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6G 2V4, Canada (e-mail: svor@ieee.org).

J. Zhang and M. Haardt are with the Communication Research Laboratory, Ilmenau University of Technology, Ilmenau 98693, Germany (e-mail: jianshu.zhang@tu-ilmenau.de; martin.haardt@tu-ilmenau.de).

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

Digital Object Identifier 10.1109/TSP.2013.2263501

in a TWR helps to maximize the spectral efficiency in terms of sum-rate [2]–[7]. As shown in Part I of this two-part paper [2], achieving the maximum sum-rate in TWR does not necessarily demand the consumption of all the available power at all participating nodes. As a result, it is of interest to find the power allocation which minimizes the power consumption of the participating nodes among all power allocations that achieve the maximum sum-rate in TWR. For brevity, this objective of optimizing the power allocation at the participating nodes is called the sum-rate maximization with minimum power consumption. In Part I of this two-part paper, the multiple-input multiple-output (MIMO) decode-and-forward (DF) TWR is investigated in the *relay optimization scenario*, in which the relay optimizes its own power allocation to achieve sum-rate maximization with minimum power consumption given the power allocation of the source nodes. The solution of the relay optimization scenario derived in Part I gives the optimal power allocation of the relay in a MIMO DF TWR system in the case when there is no coordination between the relay and the source nodes. Although this power allocation is in general sub-optimal on the network level, it is a viable and preferable solution for power allocation when the considered MIMO DF TWR system has limitation on the computational capability of finding the power allocation strategy. If the participating nodes have sufficient computational capability, a better performance than that in the relay optimization scenario can be achieved. In such a case, the relay and the source nodes can jointly optimize their power allocation strategies for sum-rate maximization with minimum total power consumption.

Joint optimization of transmit strategies of the relay and source nodes for MIMO TWR has been studied in [4]–[7]. Transmit strategies for maximizing the weighted sum-rate of a TWR system are studied in [4], in which the optimal solution is found through alternative optimization over the transmit strategies of the relay and source nodes. In [5], a low-complexity sub-optimal design of relay and source node transmit strategies is derived for either sum-rate maximization or power consumption minimization under quality-of-service requirements. The joint source node and relay precoding design for minimizing the mean squared error in a MIMO TWR system is studied in [6]. The optimal solution is found through an alternative optimization of several sub-problems obtained from the original nonconvex problem. The authors in [7] solve the robust joint source and relay optimization problem for a MIMO TWR system with imperfect channel state information. Deriving the optimal solution for the joint optimization problem generally requires alternative optimization over the transmit

strategies of the relay and the source nodes, which leads to high complexity [4], [6], [7]. All the above works consider MIMO amplify-and-forward (AF) TWR.

Considering the fact that DF TWR may achieve better performance than AF TWR, especially at low signal-to-noise ratio (SNR) [8], and the fact that DF TWR has the flexibility of performing separate power allocation/precoding for relaying the communication on each direction, it is of interest to study the problem of joint optimization over the power allocation strategies of the relay and the source nodes for MIMO DF TWR. If we further consider the power efficiency, the problem becomes more complicated. Part II of this two-part paper studies the problem of sum-rate maximization with minimum power consumption for MIMO DF TWR when the relay and the source nodes jointly optimize their power allocations. This scenario is referred to as *network optimization scenario*. The objective of this part is to find the jointly optimal power allocation of the relay and the source nodes while reducing the complexity of finding the optimal solution.¹ The contributions of Part II are as follows.

First, we show that the considered network optimization problem is nonconvex. Based on the comparison of the maximum achievable sum-rates of the multiple-access (MA) and broadcasting (BC) phases, the network optimization problem is considered for the case that the maximum achievable sum-rate of the MA phase is larger than or equal to that of the BC phase and the case that the maximum achievable sum-rate of the MA phase is less than that of the BC phase, respectively. In each case, we show that the original problem can be transferred, under certain conditions, into equivalent convex problem(s) which can be solved with low complexity. Accordingly, the above two cases are further analyzed in terms of subcases. For the subcases in which the original problem can be transferred into equivalent convex problems, the problem of sum-rate maximization and the problem of power consumption minimization are decoupled so that the sum-rate in one of the MA and BC phase is maximized while the power consumption in the other phase is minimized. The complexity of finding the optimal solution of the network optimization problem in the above subcases is therefore low.

Second, for the remaining two subcases in which the original problem cannot be transferred to a convex form, we prove properties that the optimal solution must satisfy. Based on these properties, we propose an algorithm for finding the jointly optimal power allocation for the relay and the source nodes. While the proposed algorithm finds the optimal solution iteratively, the optimization problems that the relay and the source nodes need to solve in each iteration are convex and simple. As a result, the complexity of the proposed algorithm for finding the optimal solution of the nonconvex joint optimization problem is acceptable in these two subcases.

Third, we demonstrate the effect of asymmetry on MIMO DF TWR in the network optimization scenario. Similar to the relay optimization scenario, we show that asymmetry in power limits, number of antennas, and channels can lead to performance degradation in both the achievable sum-rate and the power allocation efficiency. Specifically, we show that the optimal power allocation in both of the aforementioned two subcases in

which the original problem cannot be transformed to a convex form is not as efficient as that in other subcases. Then, it is shown through analysis and simulation that the asymmetry in the power limits, number of antennas, and channels leads to a larger occurrence probability of the two inefficient subcases. As a result, we show that the asymmetry leads to performance degradation in the MIMO DF TWR system.

The rest of the paper is organized as follows. Section II gives the system model considered in this work. The network optimization problem is studied in Section III. Simulation results are shown in Section IV and Section V concludes the paper. The Appendix provides proofs for the lemmas and theorems.

II. SYSTEM MODEL

A TWR with two source nodes and one relay is considered, where source node i ($i = 1, 2$) and the relay have n_i and n_r antennas, respectively. The information symbol vector and the precoding matrix of source node i are denoted as s_i and \mathbf{W}_i , respectively, where s_i is a complex Gaussian vector and the elements of s_i , $\forall i$ are independent and identically distributed with zero mean and unit variance. The channels from source node i to the relay and from the relay to source node i are denoted as \mathbf{H}_{ir} and \mathbf{H}_{ri} , respectively. It is assumed that source node i knows \mathbf{H}_{ri} and the relay knows \mathbf{H}_{ir} , $\forall i$ as receiver side channel knowledge. It is also assumed that the relay knows \mathbf{H}_{ri} , $\forall i$ by using either channel reciprocity or channel feedback. For example, if the system works in the time-division duplex mode, \mathbf{H}_{ri} , $\forall i$ are known at the relay due to channel reciprocity. Otherwise, when the system works in the frequency-division duplex mode, the relay needs feedback from the source nodes to obtain \mathbf{H}_{ri} , $\forall i$.

In the MA phase, source node i transmits the signal $\mathbf{W}_i s_i$ to the relay. The sum-rate of the MA phase, denoted as $R^{\text{ma}}(\mathbf{D})$, is bounded by [10]

$$R^{\text{ma}}(\mathbf{D}) = \log \left| \mathbf{I} + (\mathbf{H}_{1r} \mathbf{D}_1 \mathbf{H}_{1r}^H + \mathbf{H}_{2r} \mathbf{D}_2 \mathbf{H}_{2r}^H) (\sigma_r^2)^{-1} \right| \quad (1)$$

where \mathbf{I} denotes the identity matrix, the $(\cdot)^H$ stands for the conjugate transpose, $\mathbf{D}_i \triangleq \mathbf{W}_i \mathbf{W}_i^H$, $\forall i$, $\mathbf{D} \triangleq [\mathbf{D}_1, \mathbf{D}_2]$, and $\sigma_r^2 \mathbf{I}$ is the noise covariance matrix at the relay.

The relay decodes s_1 and s_2 from the received signal, performs precoding for each of them, and then forwards the superposition of the precoded information symbols to the source nodes in the BC phase. Note that the exclusive-OR (XOR) based network coding is adopted at the relay in some works (for example [11]). While XOR based network coding may achieve better performance in terms of sum-rate than the symbol-level superposition, it relies largely on the symmetry of the traffic from the two source nodes. The asymmetry in the traffic in the two directions can lead to significant degradation in the performance of XOR based network coding in TWR [12], [13]. As the general case of TWR is considered and there is no guarantee of traffic symmetry, the approach of symbol-level superposition is adopted here at the relay as it is considered in [1].

With the receiver side channel knowledge and the knowledge of the relay precoder, each source node is able to subtract its self-interference from the received signal. Denote \mathbf{T}_{ri} as the relay precoding matrix for relaying the signal from source node

¹Some preliminary results were presented at a conference [9].

j to source node i .² Let $\mathbf{B}_i \triangleq \mathbf{T}_{r_i} \mathbf{T}_{r_i}^H, \forall i$ and $\mathbf{B} \triangleq [\mathbf{B}_1, \mathbf{B}_2]$. Then the information rate for the communication from the relay to source node i , denoted as $\hat{R}_{r_i}(\mathbf{B}_i)$, is expressed as

$$\hat{R}_{r_i}(\mathbf{B}_i) = \log \left| \mathbf{I} + (\mathbf{H}_{r_i} \mathbf{B}_i \mathbf{H}_{r_i}^H) (\sigma_i^2)^{-1} \right| \quad (2)$$

where $\sigma_i^2 \mathbf{I}$ is the noise covariance matrix at source node i . The sum-rate of the BC phase, denoted as $R^{\text{bc}}(\mathbf{B})$, is defined as

$$R^{\text{bc}}(\mathbf{B}) = \hat{R}_{r_1}(\mathbf{B}_1) + \hat{R}_{r_2}(\mathbf{B}_2). \quad (3)$$

The end-to-end information rate from source node j to source node i , denoted as $R_{ji}(\mathbf{B}_i, \mathbf{D}_j)$, is given by

$$R_{ji}(\mathbf{B}_i, \mathbf{D}_j) = \frac{1}{2} \min\{\hat{R}_{r_i}(\mathbf{B}_i), \bar{R}_{j_r}(\mathbf{D}_j)\} \quad (4)$$

where

$$\bar{R}_{j_r}(\mathbf{D}_j) = \log \left| \mathbf{I} + (\mathbf{H}_{j_r} \mathbf{D}_j \mathbf{H}_{j_r}^H) (\sigma_r^2)^{-1} \right|. \quad (5)$$

Then the sum-rate for communication over both MA and BC phases for the considered DF TWR can be written as [1]

$$R^{\text{tw}}(\mathbf{B}, \mathbf{D}) = \frac{1}{2} \min\{R^{\text{ma}}(\mathbf{D}), R(\mathbf{B}, \mathbf{D})\} \quad (6)$$

where

$$R(\mathbf{B}, \mathbf{D}) = \min\{\hat{R}_{r_1}(\mathbf{B}_1), \bar{R}_{2r}(\mathbf{D}_2)\} + \min\{\hat{R}_{r_2}(\mathbf{B}_2), \bar{R}_{1r}(\mathbf{D}_1)\}. \quad (7)$$

Denote the singular value decomposition (SVD) of \mathbf{H}_{r_i} as $\mathbf{H}_{r_i} = \mathbf{U}_{r_i} \mathbf{\Omega}_{r_i} \mathbf{V}_{r_i}^H$. We assume that the rank of \mathbf{H}_{r_i} is r_{r_i} ($r_{r_i} \leq \min(n_i, n_r)$) and the first r_{r_i} diagonal elements of $\mathbf{\Omega}_{r_i}$, denoted as $\omega_{r_i}(1), \dots, \omega_{r_i}(r_{r_i})$, are non-zero. Since the source nodes can subtract their self-interference in the BC phase and the relay has channel knowledge of $\mathbf{H}_{r_i}, \forall i$, the power allocation of the relay for relaying the signal in either direction should be based on waterfilling regardless of how the relay distributes its power between relaying the signals in the two directions. The actual water-levels used by the relay for relaying the signal from source node j to source node i are denoted as $1/\lambda_i, \forall i$. With water-level $1/\lambda_i$, \mathbf{B}_i can be given as $\mathbf{B}_i = \mathbf{V}_{r_i} \mathbf{P}_{r_i}(\lambda_i) \mathbf{V}_{r_i}^H$, where $\mathbf{P}_{r_i}(\lambda_i) = \text{diag}((\frac{1}{\lambda_i} - \frac{1}{\alpha_i(1)})^+, \dots, (\frac{1}{\lambda_i} - \frac{1}{\alpha_i(r_{r_i})})^+, 0, \dots, 0)$ in which $\text{diag}(\cdot)$ stands for making a diagonal matrix using the given elements, $(\cdot)^+$ denotes the projection to the positive orthant, $\alpha_i(k) \triangleq |\omega_{r_i}(k)|^2 / \sigma_i^2$, and there are $(n_r - r_{r_i})$ zeros on the main diagonal of $\mathbf{P}_{r_i}(\lambda_i)$.³ It holds that

$$\hat{R}_{r_1}(\mathbf{B}_1) = \sum_{k \in \mathcal{I}_1} \log \left(1 + \left(\frac{1}{\lambda_1} \alpha_1(k) - 1 \right)^+ \right) \quad (8a)$$

$$\hat{R}_{r_2}(\mathbf{B}_2) = \sum_{k \in \mathcal{I}_2} \log \left(1 + \left(\frac{1}{\lambda_2} \alpha_2(k) - 1 \right)^+ \right) \quad (8b)$$

²It is assumed as default throughout the paper that the user indices i and j satisfy $i \neq j$.

³Details on waterfilling based solution of power allocation can be found, for example, in Section III.A in [14].

where $\mathcal{I}_i = \{1, \dots, r_{r_i}\}$. Therefore, the rate $\hat{R}_{r_i}(\mathbf{B}_i)$ obtained using water-level $1/\lambda_i$ is alternatively denoted as $\hat{R}_{r_i}(\lambda_i)$.

From (6), it can be seen that the maximization of the sum-rate using minimum total power potentially involves balancing between $R^{\text{ma}}(\mathbf{D})$ and $R(\mathbf{B}, \mathbf{D})$ and between $\hat{R}_{r_i}(\mathbf{B}_i)$ and $\bar{R}_{j_r}(\mathbf{D}_j), \forall i$. However, it is not explicit how such rate balancing affects the power allocation of the relay and the source nodes. In order to adjust the above rates through power allocation, we introduce the relative water-levels. Same as in Part I, $1/\mu_1(\mathbf{D}_1)$, $1/\mu_2(\mathbf{D}_2)$, and $1/\mu_{\text{ma}}(\mathbf{D})$ are defined as

$$\sum_{k \in \mathcal{I}_2} \log \left(1 + \left(\frac{1}{\mu_1(\mathbf{D}_1)} \alpha_2(k) - 1 \right)^+ \right) = \bar{R}_{1r}(\mathbf{D}_1) \quad (9a)$$

$$\sum_{k \in \mathcal{I}_1} \log \left(1 + \left(\frac{1}{\mu_2(\mathbf{D}_2)} \alpha_1(k) - 1 \right)^+ \right) = \bar{R}_{2r}(\mathbf{D}_2) \quad (9b)$$

$$\sum_i \sum_{k \in \mathcal{I}_i} \log \left(1 + \left(\frac{1}{\mu_{\text{ma}}(\mathbf{D})} \alpha_i(k) - 1 \right)^+ \right) = R^{\text{ma}}(\mathbf{D}). \quad (9c)$$

Given the above definition, if waterfilling is performed on $\omega_{r_j}(k)$'s, $\forall k \in \mathcal{I}_j$ using the water-level $1/\mu_i(\mathbf{D}_i)$, then the information rate of the transmission from the relay to source node j using the resulting power allocation achieves precisely $\bar{R}_{j_r}(\mathbf{D}_i)$. If waterfilling is performed on $\omega_{r_i}(k)$'s, $\forall k \in \mathcal{I}_i, \forall i$ using the water-level $1/\mu_{\text{ma}}(\mathbf{D})$, then the sum-rate of the transmission from the relay to the two source nodes using the resulting power allocation achieves precisely $R^{\text{ma}}(\mathbf{D})$. For brevity, $\mu_1(\mathbf{D}_1)$, $\mu_2(\mathbf{D}_2)$, and $\mu_{\text{ma}}(\mathbf{D})$ are denoted hereafter as μ_1 , μ_2 and μ_{ma} , respectively. The same markers/superscripts on \mathbf{D}_i and/or \mathbf{D} are used on μ_i and/or μ_{ma} to represent the connection. For example, $\mu_i(\mathbf{D}_i^0)$ and $\mu_{\text{ma}}(\tilde{\mathbf{D}})$ are briefly denoted as μ_i^0 and $\tilde{\mu}_{\text{ma}}$, respectively.

For the network optimization scenario considered here, the relay and the source nodes jointly maximize the sum-rate in (6) with minimum total transmission power in the network.⁴ Similar to the relay optimization scenario, the relay needs to know \mathbf{W}_1 and \mathbf{W}_2 while both source nodes need to know \mathbf{T}_{r_1} and \mathbf{T}_{r_2} . In the network optimization scenario, it is preferable that the TWR is able to operate in a centralized mode in which the relay can serve as a central node that carries out the computations. If the system works in a decentralized mode, it may lead to high overhead because of the information exchange during the iterative optimization process.

Given the above system model, we next solve the network optimization problem.

III. NETWORK OPTIMIZATION

In the network optimization scenario, the relay and the source nodes jointly optimize their power allocation to achieve sum-rate maximization with minimum total power consumption in the system for the MIMO DF TWR. Compared to the optimal solution of the relay optimization problem in Part I, the optimal solution of the network optimization problem achieves larger sum-rate and/or less power consumption at the cost of higher computational complexity.

⁴The term 'sum-rate' by default means $R^{\text{tw}}(\mathbf{B}, \mathbf{D})$ when we do not specify it to be the sum-rate of the BC or MA phase.

The sum-rate maximization part can be formulated as the following optimization problem⁵

$$\max_{\{\mathbf{B}, \mathbf{D}\}} R^{\text{tw}}(\mathbf{B}, \mathbf{D}) \quad (10a)$$

$$\text{s.t.} \quad \text{Tr}\{\mathbf{D}_i\} \leq P_i^{\text{max}}, \forall i \quad (10b)$$

$$\text{Tr}\{\mathbf{B}_1 + \mathbf{B}_2\} \leq P_r^{\text{max}} \quad (10c)$$

where $\text{Tr}\{\cdot\}$ denotes the trace, and P_i^{max} and P_r^{max} are the power limits for source node i and the relay, respectively. The above problem is a convex problem which can be rewritten into the standard form by introducing variables t, t_1, t_2 as follows

$$\max_{\{t, t_1, t_2, \mathbf{B}, \mathbf{D}\}} t \quad (11a)$$

$$\text{s.t.} \quad t \leq R^{\text{ma}}(\mathbf{D}), \quad t \leq t_1 + t_2 \quad (11b)$$

$$t_i \leq \hat{R}_{rj}(\mathbf{B}_j), \quad t_i \leq \bar{R}_{ir}(\mathbf{D}_i), \forall i \quad (11c)$$

$$\text{Tr}\{\mathbf{D}_i\} \leq P_i^{\text{max}}, \forall i, \quad \text{Tr}\{\mathbf{B}_1 + \mathbf{B}_2\} \leq P_r^{\text{max}}. \quad (11d)$$

If transmission power minimization is also taken into account, the following constraints become necessary

$$\hat{R}_{ri}(\mathbf{B}_i) \leq \bar{R}_{jr}(\mathbf{D}_j), \forall i \quad (12a)$$

$$R^{\text{ma}}(\mathbf{D}) = R(\mathbf{B}, \mathbf{D}). \quad (12b)$$

The reason why the above constraints are necessary if transmission power minimization also needs to be taken into account is as follows. Given the fact that $R^{\text{ma}}(\mathbf{D}) < \bar{R}_{1r}(\mathbf{D}_1) + \bar{R}_{2r}(\mathbf{D}_2)$ whenever $\text{Tr}\{\mathbf{D}_1\} + \text{Tr}\{\mathbf{D}_2\} > 0$, it can be seen that the power consumption of the relay can be reduced by reducing $\text{Tr}\{\mathbf{B}_i\}$ without decreasing the sum-rate $R^{\text{tw}}(\mathbf{B}, \mathbf{D})$ in (6) if $\hat{R}_{ri}(\mathbf{B}_i) > \bar{R}_{jr}(\mathbf{D}_j)$. Therefore, the constraint (12a) is necessary. Subject to (12a), $R^{\text{tw}}(\mathbf{B}, \mathbf{D})$ in (6) can be written as $\min\{R^{\text{ma}}(\mathbf{D}), \hat{R}_{r1}(\mathbf{B}_1) + \hat{R}_{r2}(\mathbf{B}_2)\}/2$. Using the fact that $R^{\text{ma}}(\mathbf{D}) < \hat{R}_{r1}(\mathbf{B}_1) + \hat{R}_{r2}(\mathbf{B}_2)$ when $\hat{R}_{r1}(\mathbf{B}_1) = \bar{R}_{1r}(\mathbf{D}_1)$ and $\hat{R}_{r2}(\mathbf{B}_2) = \bar{R}_{2r}(\mathbf{D}_2)$, it can be shown that the power consumption of at least one source node can be reduced without decreasing $R^{\text{tw}}(\mathbf{B}, \mathbf{D})$ if $R^{\text{ma}}(\mathbf{D}) > R(\mathbf{B}, \mathbf{D})$ while the power consumption of the relay can be reduced without decreasing $R^{\text{tw}}(\mathbf{B}, \mathbf{D})$ if $R^{\text{ma}}(\mathbf{D}) < R(\mathbf{B}, \mathbf{D})$. Thus, the constraint (12b) is also necessary.

Considering the constraints (12a) and (12b), the problem of finding the optimal power allocation becomes nonconvex. Relating (9a)–(9c) with (8a)–(8b), the above two constraints (12a) and (12b) can be rewritten as

$$\lambda_i \geq \mu_j, \forall i \quad (13a)$$

$$\begin{aligned} & \sum_i \sum_{k \in \mathcal{I}_i} \log \left(1 + \left(\frac{1}{\lambda_i} \alpha_i(k) - 1 \right)^+ \right) \\ & = \sum_i \sum_{k \in \mathcal{I}_i} \log \left(1 + \left(\frac{1}{\mu_{\text{ma}}} \alpha_i(k) - 1 \right)^+ \right). \end{aligned} \quad (13b)$$

It should be noted that the constraints (12a) and (12b), or equivalently (13a) and (13b), are not sufficient in general. Due to the intrinsic complexity of the considered problem, it is too

⁵The positive semi-definite constraints $\mathbf{D}_i \succeq 0, \forall i$ and $\mathbf{B}_i \succeq 0, \forall i$ are assumed as default and omitted for brevity in all formulations of optimization problems in this paper.

complicated to formulate a general sufficient and necessary condition for optimality of the original problem of sum-rate maximization with minimum power consumption. Instead, we will show sufficient and necessary optimality conditions for the equivalent problems in the subcases in which the original problem can be transferred into equivalent convex problem(s). For other subcases, we will develop important properties based on the above necessary conditions (12), or equivalently (13), which can significantly reduce the computational complexity of searching for the optimal solution.

The following lemma that applies for all subcases is introduced for subsequent analysis.

Lemma 1: Given \mathbf{D}_1 and \mathbf{D}_2 with $P_1^{\text{max}} \geq \text{Tr}\{\mathbf{D}_1\} > 0$ and $P_2^{\text{max}} \geq \text{Tr}\{\mathbf{D}_2\} > 0$, if $1/\mu_i > 1/\mu_{\text{ma}} > 1/\mu_j$, then the following two results hold true: 1) $1/\mu_{\text{ma}}(\hat{\mathbf{D}}) \leq 1/\mu_j$ where $\hat{\mathbf{D}} = [\hat{\mathbf{D}}_1, \hat{\mathbf{D}}_2]$ with $\hat{\mathbf{D}}_i = \mathbf{0}$ and $\hat{\mathbf{D}}_j = \mathbf{D}_j$, 2) there exists $t \in [0, 1)$ such that with $\hat{\mathbf{D}}_i = t\mathbf{D}_i$ and $\hat{\mathbf{D}}_j = \mathbf{D}_j$, we have $1/\mu_i(\hat{\mathbf{D}}_i) > 1/\mu_{\text{ma}}(\hat{\mathbf{D}}) = 1/\mu_j$ where $\hat{\mathbf{D}} = [\hat{\mathbf{D}}_1, \hat{\mathbf{D}}_2]$.

Proof: See Appendix A. ■

Lemma 1 relates the source nodes transmit strategy \mathbf{D} with the relative water-levels $1/\mu_1, 1/\mu_2$, and $1/\mu_{\text{ma}}$. It shows a range that the relative water-level $1/\mu_{\text{ma}}$ can reach by fixing \mathbf{D}_j and changing \mathbf{D}_i given that $1/\mu_i > 1/\mu_{\text{ma}} > 1/\mu_j$.

Lemma 2: The optimal solution of the network optimization problem has the following property

$$\lambda_j = \mu_i > \mu_{\text{ma}} \quad \text{if } \lambda_i < \lambda_j \text{ or } \mu_i > \mu_{\text{ma}}. \quad (14)$$

Proof: See Appendix B. ■

Lemma 2 develops a property of the optimal solution that follows from the constraints (13a) and (13b). This property is needed for future analysis.

In the scenario of network optimization, the three nodes aim at finding the optimal matrices \mathbf{D} and \mathbf{B} that minimize $\text{Tr}\{\mathbf{D}_1\} + \text{Tr}\{\mathbf{D}_2\} + \text{Tr}\{\mathbf{B}_1 + \mathbf{B}_2\}$ among all \mathbf{D} and \mathbf{B} that achieve the maximum of the objective function in (10). Considering the fact that the optimal \mathbf{B} and \mathbf{D} depend on each other, solving this problem generally involves alternative optimization of \mathbf{B} and \mathbf{D} . It is, however, of interest to avoid such alternative process when it is possible due to its high complexity. Next we use an initial power allocation⁶ to classify the problem of finding the optimal \mathbf{B} and \mathbf{D} for network optimization into two cases, each with several subcases.

Consider the following initial power allocation of the source nodes and the relay, which decides the maximum achievable sum-rates of the MA and BC phases, respectively. The initial power allocation of the source nodes is the solution to the following problem

$$\max_{\mathbf{D}} R^{\text{ma}}(\mathbf{D}) \quad (15a)$$

$$\text{s.t.} \quad \text{Tr}\{\mathbf{D}_i\} \leq P_i^{\text{max}}, \forall i \quad (15b)$$

which is a power allocation problem on multiple-access channels studied in [14], while the initial power allocation of the relay is to allocate P_r^{max} on $\alpha_i(k)$'s, $\forall k \in \mathcal{I}_i, \forall i$ based on the waterfilling procedure. Denote the optimal solution of (15) as $\mathbf{D}^0 = [\mathbf{D}_1^0, \mathbf{D}_2^0]$ and the water level corresponding to the relay's initial power allocation as $1/\lambda^0$. The case when $R^{\text{ma}}(\mathbf{D}^0) \geq$

⁶Note that the initial power allocation is not the solution to the considered problem and it is only used for enabling classification.

$\hat{R}_{r1}(\lambda^0) + \hat{R}_{r2}(\lambda^0)$, i.e., when the maximum achievable sum-rate of the MA phase is larger than or equal to that of the BC phase, is denoted as Case I and the case when $R^{\text{ma}}(\mathbf{D}^0) < \hat{R}_{r1}(\lambda^0) + \hat{R}_{r2}(\lambda^0)$, i.e., when the maximum achievable sum-rate of the MA phase is less than that of the BC phase, is denoted as Case II. We next study the problem of maximizing $R^{\text{tw}}(\mathbf{B}, \mathbf{D})$ with minimum power consumption and find the optimal power allocation for Cases I and II, respectively, in the following subsections.

A. Finding the Optimal Solution in Case I, i.e., $R^{\text{ma}}(\mathbf{D}^0) \geq \hat{R}_{r1}(\lambda^0) + \hat{R}_{r2}(\lambda^0)$

Since $R^{\text{ma}}(\mathbf{D}^0) \geq \hat{R}_{r1}(\lambda^0) + \hat{R}_{r2}(\lambda^0)$, it can be inferred that $1/\lambda_0 \leq 1/\mu_{\text{ma}}^0$. In this case, the sum-rate $R^{\text{tw}}(\mathbf{B}, \mathbf{D})$ in (6) is upper-bounded by the sum-rate $\hat{R}_{r1}(\lambda^0) + \hat{R}_{r2}(\lambda^0)$. The following two subcases should be considered separately.

Subcase I-1: The following convex optimization problem is feasible

$$\min_{\mathbf{D}} \text{Tr}\{\mathbf{D}_1\} + \text{Tr}\{\mathbf{D}_2\} \quad (16a)$$

$$\text{s.t. } R^{\text{ma}}(\mathbf{D}) \geq \hat{R}_{r1}(\lambda^0) + \hat{R}_{r2}(\lambda^0) \quad (16b)$$

$$\bar{R}_{1r}(\mathbf{D}_1) \geq \hat{R}_{r2}(\lambda^0) \quad (16c)$$

$$\bar{R}_{2r}(\mathbf{D}_2) \geq \hat{R}_{r1}(\lambda^0) \quad (16d)$$

$$\text{Tr}(\mathbf{D}_i) \leq P_i^{\text{max}}, \forall i. \quad (16e)$$

In this subcase, the maximum sum-rate $R^{\text{tw}}(\mathbf{B}, \mathbf{D})$ can achieve $\hat{R}_{r1}(\lambda^0) + \hat{R}_{r2}(\lambda^0)$. In order to achieve this maximum sum-rate, it is necessary that $\lambda_1 = \lambda_2 = \lambda^0$. Therefore, the relay should use up all available power P_r^{max} at optimality, and the optimal $\mathbf{B}_i, \forall i$ are equal to $\mathbf{V}_{ri} \mathbf{P}_{ri}(\lambda^0) \mathbf{V}_{ri}^H, \forall i$ where $\mathbf{P}_{ri}(\lambda_i)$ is given in Section II. As a result, the original problem simplifies to finding the optimal \mathbf{D}_1 and \mathbf{D}_2 such that $R^{\text{tw}}(\mathbf{B}, \mathbf{D})$ achieves $\hat{R}_{r1}(\lambda^0) + \hat{R}_{r2}(\lambda^0)$ with minimum power consumption. Using (6) and (7), it can be shown that a sufficient and necessary condition for \mathbf{D} to be optimal in this subcase is that \mathbf{D} is the optimal solution to the convex optimization problem (16). Denoting the optimal solution to the problem (16) as $\mathbf{D}^* = [\mathbf{D}_1^*, \mathbf{D}_2^*]$, the total power consumption in this subcase is $P_r^{\text{max}} + \text{Tr}\{\mathbf{D}_1^*\} + \text{Tr}\{\mathbf{D}_2^*\}$.

It can be seen that the optimal solution of \mathbf{B} and \mathbf{D} in Subcase I-1 satisfies the general constraint (13a), or equivalently (12a), for the original problem since the constraints (16c) and (16d) are considered in the problem (16). It can also be shown that the above optimal solution in Subcase I-1 also satisfies the general constraint (13b), or equivalently (12b), for the original problem as stated in the following theorem.

Theorem 1: The optimal solution in Subcase I-1 satisfies $\mu_{\text{ma}}^* = \lambda^0$, and thereby satisfies (13b) given that $\lambda_1 = \lambda_2 = \lambda^0$ at optimality.

Proof: See Appendix C.

Considering the constraints (16b)–(16e), it can be seen that the problem (16) is feasible if and only if the following problem

$$\max_{\mathbf{D}} \bar{R}_{jr}(\mathbf{D}_j) \quad (17a)$$

$$\text{s.t. } \bar{R}_{ir}(\mathbf{D}_i) \geq \hat{R}_{rj}(\lambda^0) \quad (17b)$$

$$R^{\text{ma}}(\mathbf{D}) \geq \hat{R}_{r1}(\lambda^0) + \hat{R}_{r2}(\lambda^0) \quad (17c)$$

$$\text{Tr}(\mathbf{D}_1) \leq P_1^{\text{max}} \quad (17d)$$

$$\text{Tr}(\mathbf{D}_2) \leq P_2^{\text{max}} \quad (17e)$$

is feasible and its optimal solution, denoted as \mathbf{D}^* , satisfies $\bar{R}_{jr}(\mathbf{D}_j^*) \geq \hat{R}_{ri}(\lambda^0), \forall j$.⁷ However, it is possible that $\bar{R}_{jr}(\mathbf{D}_j^*) < \hat{R}_{ri}(\lambda^0)$ for some i and j . It is also possible that the problem (17) is not even feasible. In either of the above two situations the problem (16) is infeasible. This leads to the second subcase of Case I.

Subcase I-2: The problem (16) is infeasible.

Unlike Subcase I-1, the maximum sum-rate $R^{\text{tw}}(\mathbf{B}, \mathbf{D})$ in this subcase cannot achieve $\hat{R}_{r1}(\lambda^0) + \hat{R}_{r2}(\lambda^0)$. As mentioned above, there are two possible situations when the problem (16) is infeasible: (i) $\bar{R}_{jr}(\mathbf{D}_j^*) < \hat{R}_{ri}(\lambda^0), \forall j$ and (ii) the problem (17) is infeasible for at least one specific value of i and the corresponding j . Using Lemma 1 in Part I of this two-part paper [2] and the fact that $R^{\text{ma}}(\mathbf{D}^0) \geq \hat{R}_{r1}(\lambda^0) + \hat{R}_{r2}(\lambda^0)$ for Case I, it can be shown that if the problem (17) is infeasible for specific values of i and j , then it is feasible (but $\bar{R}_{jr}(\mathbf{D}_j^*) < \hat{R}_{ri}(\lambda^0)$) when the values of i and j are switched. Therefore, the problem (16) is infeasible if and only if there exists at least one specific value of j in $\{1, 2\}$ such that problem (17) is feasible but $\bar{R}_{jr}(\mathbf{D}_j^*) < \hat{R}_{ri}(\lambda^0)$. Denote this specific value of j as l and denote the corresponding i as \bar{l} . It infers, based on the definitions (9a)–(9c), that $1/\mu_l < 1/\lambda^0$ whenever $1/\mu_{\text{ma}} \geq 1/\lambda^0$ and $1/\mu_{\bar{l}} \geq 1/\lambda^0$. As a result, whenever $1/\mu_{\text{ma}} \geq 1/\lambda^0$, or equivalently, $R^{\text{ma}}(\mathbf{D}) \geq \hat{R}_{r1}(\lambda^0) + \hat{R}_{r2}(\lambda^0)$, the sum-rate $R^{\text{tw}}(\mathbf{B}, \mathbf{D})$ is bounded by $\hat{R}_{r1}(\lambda_1) + \hat{R}_{r2}(\lambda_2)$ according to (6), which is less than $\hat{R}_{r1}(\lambda^0) + \hat{R}_{r2}(\lambda^0)$ when $1/\mu_l < 1/\lambda^0$ (according to the constraint (13a) and Lemma 2 in Part I [2]). Moreover, whenever $1/\mu_{\text{ma}} < 1/\lambda^0$, or equivalently, $R^{\text{ma}}(\mathbf{D}) < \hat{R}_{r1}(\lambda^0) + \hat{R}_{r2}(\lambda^0)$, the sum-rate $R^{\text{tw}}(\mathbf{B}, \mathbf{D})$ is bounded by $R^{\text{ma}}(\mathbf{D})$ according to (6), which is also less than $\hat{R}_{r1}(\lambda^0) + \hat{R}_{r2}(\lambda^0)$. Therefore, the maximum sum-rate $R^{\text{tw}}(\mathbf{B}, \mathbf{D})$ in this subcase always cannot achieve $\hat{R}_{r1}(\lambda^0) + \hat{R}_{r2}(\lambda^0)$.

With the above denotation of l and \bar{l} , the following theorem characterizes the optimal solution in this subcase.

Theorem 2: Denote the optimal \mathbf{D}_i in Subcase I-2 as $\mathbf{D}_i^*, \forall i$ and the optimal λ_i as $\lambda_i^*, \forall i$. The optimal strategies for the source nodes and the relay satisfy the following properties:

1. $\min_i \{1/\mu_i^*\} < 1/\mu_{\text{ma}}^* < 1/\lambda^0$;
2. The relay uses full power P_r^{max} ;
3. \mathbf{D}^* maximizes $\min_i \{1/\mu_i\}$ among all \mathbf{D} 's that satisfy

$$R^{\text{ma}}(\mathbf{D}) \geq R^{\text{ma}}(\mathbf{D}^*) \quad (18a)$$

$$\text{Tr}(\mathbf{D}_i) \leq P_i^{\text{max}}, \forall i \quad (18b)$$

4. $1/\mu_l^* < 1/\mu_{\bar{l}}^*$.

Proof: See Appendix D.

While the original problem cannot be simplified into an equivalent form in this subcase, the properties in Theorem 2 help to significantly reduce the complexity of searching for the optimal solution by narrowing down the set of qualifying power allocations. Denote \mathbf{D}_l^1 as the \mathbf{D}_l that maximizes $\bar{R}_{lr}(\mathbf{D}_l)$ subject to the constraints $\mu_l \geq \mu_{\text{ma}}$ and $\text{Tr}\{\mathbf{D}_l\} \leq P_l^{\text{max}}$ and denote μ_l^1 as the corresponding μ_l . According to Theorem 2, if $\hat{R}_{r\bar{l}}(\lambda_{\bar{l}}^\dagger) + \hat{R}_{rl}(\lambda_l^\dagger) \leq R^{\text{ma}}(\mathbf{D}^\dagger)$, where

$$\lambda_l^\dagger = \mu_l^1 \quad (19a)$$

$$\text{Tr}\{\mathbf{P}_{r\bar{l}}(\lambda_{\bar{l}}^\dagger)\} + \text{Tr}\{\mathbf{P}_{rl}(\lambda_l^\dagger)\} = P_r^{\text{max}} \quad (19b)$$

⁷Note that if $\bar{R}_{jr}(\mathbf{D}_j^*) \geq \hat{R}_{ri}(\lambda^0)$ for $i = 1, j = 2$ in (17) then it also holds that $\bar{R}_{jr}(\mathbf{D}_j^*) \geq \hat{R}_{ri}(\lambda^0)$ for $i = 2, j = 1$ and vice versa.

and \mathbf{D}^\dagger is the optimal solution of the following problem

$$\max_{\mathbf{D}} R^{\text{ma}}(\mathbf{D}) \quad (20a)$$

$$\text{s.t. } \bar{R}_{lr}(\mathbf{D}_l) \geq \bar{R}_{lr}(\mathbf{D}_l^\dagger) \quad (20b)$$

$$\text{Tr}(\mathbf{D}_i) \leq P_i^{\text{max}}, \forall i \quad (20c)$$

then the maximum achievable sum-rate in Subcase I-2 is $\hat{R}_{r\bar{l}}(\lambda_{\bar{l}}^\dagger) + \hat{R}_{rl}(\lambda_l^\dagger)$ (according to Lemma 2 in Part I [2]). The optimal $\mathbf{B}_i, \forall i$ in this subcase is given by $\mathbf{B}_i = \mathbf{V}_{ri} \mathbf{P}_{ri}(\lambda_i^\dagger) \mathbf{V}_{ri}^H$, and the optimal \mathbf{D} is the solution to the following power minimization problem

$$\min_{\mathbf{D}} \text{Tr}\{\mathbf{D}_1\} + \text{Tr}\{\mathbf{D}_2\} \quad (21a)$$

$$\text{s.t. } R^{\text{ma}}(\mathbf{D}) \geq \sum_i \hat{R}_{ri}(\lambda_i^\dagger) \quad (21b)$$

$$\bar{R}_{lr}(\mathbf{D}_l) \geq \hat{R}_{rl}(\lambda_l^\dagger) \quad (21c)$$

$$\bar{R}_{r\bar{l}}(\mathbf{D}_{\bar{l}}) \geq \hat{R}_{r\bar{l}}(\lambda_{\bar{l}}^\dagger) \quad (21d)$$

$$\text{Tr}(\mathbf{D}_i) \leq P_i^{\text{max}}, \forall i. \quad (21e)$$

If $\hat{R}_{r\bar{l}}(\lambda_{\bar{l}}^\dagger) + \hat{R}_{rl}(\lambda_l^\dagger) > R^{\text{ma}}(\mathbf{D}^\dagger)$, denote the objective $R^{\text{tw}}(\mathbf{B}, \mathbf{D})$ as R^{obj} . According to Theorem 2, the optimal solution can be found by maximizing R^{obj} so that it can be achieved by both $R^{\text{ma}}(\mathbf{D})$ and $\sum_i \hat{R}_{ri}(\lambda_i)$ subject to the following two constraints: 1) $1/\lambda_{\bar{l}} = 1/\tilde{\mu}_l$ which is obtained according to Lemma 2 and Properties 1 and 4 of Theorem 2; 2) $1/\lambda_l$ is obtained by waterfilling the remaining power on $\alpha_l(k), \forall k \in \mathcal{I}_l$ (Property 2 of Theorem 2), where $1/\tilde{\mu}_l = 1/\mu_l(\mathbf{D})$ is the optimal value of the objective function in the following optimization problem (Property 3 of Theorem 2)

$$\max_{\mathbf{D}} \frac{1}{\mu_l} \quad (22a)$$

$$\text{s.t. } R^{\text{ma}}(\mathbf{D}) \geq R^{\text{obj}} \quad (22b)$$

$$\text{Tr}(\mathbf{D}_i) \leq P_i^{\text{max}}, \forall i. \quad (22c)$$

Here $\tilde{\mathbf{D}}$ denotes the optimal solution of (22) for the given R^{obj} . Since maximizing $1/\mu_l$ is equivalent to maximizing $\bar{R}_{lr}(\mathbf{D}_l)$, the objective function of the above problem can be substituted by $\bar{R}_{lr}(\mathbf{D}_l)$ and $1/\tilde{\mu}_l$ can be obtained from the optimal value of $\bar{R}_{lr}(\mathbf{D}_l)$ in the above problem using (9a) or (9b). As mentioned at the beginning of Subcase I-2, the optimal $R^{\text{tw}}(\mathbf{B}, \mathbf{D})$ is less than $\sum_i \hat{R}_{ri}(\lambda^0)$. Therefore, starting from the point by setting $R^{\text{obj}} = \sum_i \hat{R}_{ri}(\lambda^0)$, we can adjust R^{obj} as follows, in order to achieve the optimal $R^{\text{tw}}(\mathbf{B}, \mathbf{D})$. We first solve the following problem given R^{obj}

$$\max_{\mathbf{D}} \bar{R}_{lr}(\mathbf{D}_l) \quad (23a)$$

$$\text{s.t. } R^{\text{ma}}(\mathbf{D}) \geq R^{\text{obj}} \quad (23b)$$

$$\text{Tr}(\mathbf{D}_i) \leq P_i^{\text{max}}, \forall i \quad (23c)$$

to get the optimal $\tilde{\mathbf{D}}$ for the given R^{obj} and obtain the resulting $1/\tilde{\mu}_l = 1/\mu_l(\tilde{\mathbf{D}})$. Then, we set $1/\lambda_{\bar{l}} = 1/\tilde{\mu}_l$ and allocate all the remaining power on $\alpha_l(k)$'s, $\forall k \in \mathcal{I}_l$. If the resulting

$\sum_i \hat{R}_{ri}(\lambda_i)$ is less than R^{obj} , it infers that R^{obj} should be decreased in (23) and the above process should be repeated. On the other hand, if the resulting $\sum_i \hat{R}_{ri}(\lambda_i)$ is larger than R^{obj} , then R^{obj} should be increased in (23) and the above process should be repeated. The optimal solution is found when the resulting $\sum_i \hat{R}_{ri}(\lambda_i)$ is equal to R^{obj} . With an appropriate step size of increasing/decreasing R^{obj} , R^{obj} in the above procedure converges to the optimal $R^{\text{tw}}(\mathbf{B}, \mathbf{D})$.

After obtaining the optimal R^{obj} , $1/\tilde{\mu}_l$, and $\lambda_{\bar{l}}$, the source nodes solve the problem of power minimization, which is

$$\min_{\mathbf{D}} \text{Tr}\{\mathbf{D}_1\} + \text{Tr}\{\mathbf{D}_2\} \quad (24a)$$

$$\text{s.t. } R^{\text{ma}}(\mathbf{D}) \geq R^{\text{obj}} \quad (24b)$$

$$\bar{R}_{lr}(\mathbf{D}_l) \geq \hat{R}_{rl}(\lambda_l) \quad (24c)$$

$$\bar{R}_{r\bar{l}}(\mathbf{D}_{\bar{l}}) \geq \hat{R}_{r\bar{l}}(\lambda_{\bar{l}}) \quad (24d)$$

$$\text{Tr}(\mathbf{D}_i) \leq P_i^{\text{max}}, \forall i. \quad (24e)$$

However, it can be shown that if $\bar{R}_{lr}(\tilde{\mathbf{D}}_l)$ is not the maximum that $\bar{R}_{lr}(\mathbf{D}_l)$ can achieve subject to the constraint (22c) (without the constraint (22b)), then \mathbf{B} and \mathbf{D} remain the same after solving the above problem.

Using Property 2 of Theorem 2, it can be seen from (21) and (24) that the minimization of total power consumption becomes the minimization of the source node power consumption in Subcase I-2 since the relay always needs to consume all its available power for achieving optimality.

The complete procedure of finding the optimal solution in Case I is summarized in the algorithm in Table I. The algorithm finds the optimal solution either in one shot (Steps 1 and 2) or through a bisection search for the optimal R^{obj} (Steps 3 to 5). Denoting $\Delta \triangleq R^{\text{max}} - R^{\text{min}}$, the worst case number of iterations in the bisection search is $\log(\Delta/\epsilon)$. Within each iteration, a convex problem, i.e., problem (23), is solved followed by a simple waterfilling procedure which has linear complexity for the given R^{obj} . Therefore, the complexity of the proposed algorithm is low.

Subcases I-1 and I-2 cover all possible situations for Case I that $R^{\text{ma}}(\mathbf{D}^0) \geq \hat{R}_{r1}(\lambda^0) + \hat{R}_{r2}(\lambda^0)$.

B. Finding the Optimal Solution in Case II, i.e., $R^{\text{ma}}(\mathbf{D}^0) < \hat{R}_{r1}(\lambda^0) + \hat{R}_{r2}(\lambda^0)$

Since $R^{\text{ma}}(\mathbf{D}^0) < \hat{R}_{r1}(\lambda^0) + \hat{R}_{r2}(\lambda^0)$, it can be seen using (8a), (8b), and (9c) that $1/\lambda_0 > 1/\mu_{\text{ma}}^0$. The following four subcases are possible.

Subcase II-1: $1/\mu_{\text{ma}}^0 \leq \min\{1/\mu_1^0, 1/\mu_2^0\}$. In this subcase, the maximum $R^{\text{tw}}(\mathbf{B}, \mathbf{D})$ is bounded by $R^{\text{ma}}(\mathbf{D}^0)$. The optimal \mathbf{D} is \mathbf{D}^0 , and consequently both source nodes use all their available power at optimality. It can be seen that a sufficient and necessary condition for \mathbf{B} to be optimal in this subcase is that \mathbf{B} is the optimal solution to the following convex optimization problem

$$\min_{\mathbf{B}} \text{Tr}\{\mathbf{B}_1 + \mathbf{B}_2\} \quad (25a)$$

$$\text{s.t. } \hat{R}_{r1}(\mathbf{B}_1) + \hat{R}_{r2}(\mathbf{B}_2) \geq R^{\text{ma}}(\mathbf{D}^0). \quad (25b)$$

TABLE I
ALGORITHM FOR FINDING THE OPTIMAL SOLUTION FOR CASE I

1. Check if the problem (16) is feasible. If yes, find the optimal \mathbf{D} from the problem (16). The optimal \mathbf{B} is given by $\mathbf{B}_i = \mathbf{V}_{ri} \mathbf{P}_{ri}(\lambda_i^0) \mathbf{V}_{ri}^H, \forall i$. Otherwise, specify l and \bar{l} so that the problem (17) is feasible but $\hat{R}_{lr}(\mathbf{D}_l^0) < \hat{R}_{r\bar{l}}(\lambda^0)$ and proceed to Step 2.
2. Obtain \mathbf{D}_l^0 and μ_l^\dagger . Calculate $\lambda_i^\dagger, \forall i$ using (19). Check if $\sum_i \hat{R}_{ri}(\lambda_i^\dagger) \leq R^{\text{ma}}(\mathbf{D}^\dagger)$. If yes, the optimal \mathbf{B} is given by $\mathbf{B}_i = \mathbf{V}_{ri} \mathbf{P}_{ri}(\lambda_i^\dagger) \mathbf{V}_{ri}^H, \forall i$. Find the optimal \mathbf{D} from (21). Otherwise, proceed to Step 3.
3. Set $R^{\text{max}} = \sum_i \hat{R}_{ri}(\lambda_i^0)$ and $R^{\text{min}} = 0$. Initialize $R^{\text{obj}} = R^{\text{max}}$ and proceed to Step 4.
4. Solve the problem (23) and obtain $\bar{\mathbf{D}}$ and $1/\bar{\mu}_l$. Set $1/\lambda_{\bar{l}} = 1/\bar{\mu}_l$. Allocate all the remaining power on $\alpha_l(k)$'s, $\forall k \in \mathcal{I}_l$ using waterfilling and obtain $1/\lambda_l$. Check if $ \sum_i \hat{R}_{ri}(\lambda_i) - R^{\text{ma}}(\bar{\mathbf{D}}) < \epsilon$, where ϵ is the positive tolerance. If yes, proceed to Step 6 with R^{obj} and $\lambda_i, \forall i$. Otherwise, proceed to Step 5.
5. If $R^{\text{ma}}(\bar{\mathbf{D}}) - \sum_i \hat{R}_{ri}(\lambda_i) > \epsilon$, set $R^{\text{max}} = R^{\text{obj}}$. If $\sum_i \hat{R}_{ri}(\lambda_i) - R^{\text{ma}}(\bar{\mathbf{D}}) > \epsilon$, set $R^{\text{min}} = R^{\text{obj}}$. Let $R^{\text{obj}} = (R^{\text{max}} + R^{\text{min}})/2$ and go back to Step 4.
6. Solve the power minimization problem (24). Output \mathbf{D} and $\mathbf{B}_i = \mathbf{V}_{ri} \mathbf{P}_{ri}(\lambda_i) \mathbf{V}_{ri}^H, \forall i$.

The solution of (25) can be found in closed-form and it is given by $\mathbf{B}_i = \mathbf{V}_{ri} \mathbf{P}_{ri}(\mu_{\text{ma}}^0) \mathbf{V}_{ri}^H, \forall i$.

Subcase II-2: there exist l and \bar{l} such that $1/\mu_l^0 \leq 1/\mu_{\text{ma}}^0 < 1/\mu_{\bar{l}}^0 \leq 1/\lambda^0$.⁸ In this subcase, the maximum achievable $R^{\text{tw}}(\mathbf{B}, \mathbf{D})$ is also $R^{\text{ma}}(\mathbf{D}^0)$. Therefore, the optimal \mathbf{D} is \mathbf{D}^0 and both source nodes use all their available power at optimality. It can be shown that a sufficient and necessary condition for \mathbf{B} to be optimal in this subcase is that \mathbf{B} is the optimal solution to the following convex optimization problem

$$\min_{\mathbf{B}} \text{Tr}\{\mathbf{B}_1 + \mathbf{B}_2\} \quad (26a)$$

$$\text{s.t. } \hat{R}_{r1}(\mathbf{B}_1) + \hat{R}_{r2}(\mathbf{B}_2) \geq R^{\text{ma}}(\mathbf{D}^0) \quad (26b)$$

$$\hat{R}_{r\bar{l}}(\mathbf{B}_{\bar{l}}) = \bar{R}_{lr}(\mathbf{D}_l^0). \quad (26c)$$

The solution of (26) can also be expressed in closed-form. The optimal $\mathbf{B}_{\bar{l}}$ is given by $\mathbf{B}_{\bar{l}} = \mathbf{V}_{r\bar{l}} \mathbf{P}_{r\bar{l}}(\mu_{\bar{l}}^0) \mathbf{V}_{r\bar{l}}^H$ and the optimal \mathbf{B}_l is given by $\mathbf{B}_l = \mathbf{V}_{rl} \mathbf{P}_{rl}(\lambda_l) \mathbf{V}_{rl}^H$, where λ_l satisfies $\hat{R}_{rl}(\lambda_l) = R^{\text{ma}}(\mathbf{D}^0) - \bar{R}_{lr}(\mathbf{D}_l^0)$.

Subcase II-3: there exist l and \bar{l} such that $1/\mu_l^0 \leq 1/\mu_{\text{ma}}^0 < 1/\lambda^0 < 1/\mu_{\bar{l}}^0$ and there exists λ_l such that

$$\hat{R}_{rl}(\lambda_l) \geq R^{\text{ma}}(\mathbf{D}^0) - \bar{R}_{lr}(\mathbf{D}_l^0) \quad (27a)$$

$$\text{Tr}\{\mathbf{P}_{rl}(\lambda_l)\} \leq P_r^{\text{max}} - \text{Tr}\{\mathbf{P}_{r\bar{l}}(\mu_{\bar{l}}^0)\}. \quad (27b)$$

The optimal solutions of \mathbf{B} and \mathbf{D} in this subcase are the same as those given in Subcase II-2.

In the above three subcases, the maximum achievable $R^{\text{tw}}(\mathbf{B}, \mathbf{D})$ is $R^{\text{ma}}(\mathbf{D}^0)$. Therefore, the original problem of maximizing $R^{\text{tw}}(\mathbf{B}, \mathbf{D})$ with minimum total power consumption in the network simplifies to the problem that the relay uses minimum power consumption to achieve the BC phase sum-rate $\hat{R}_{r1}(\mathbf{B}_1) + \hat{R}_{r2}(\mathbf{B}_2)$ that is equal to $R^{\text{ma}}(\mathbf{D}^0)$.

Subcase II-4: there exist l and \bar{l} such that $1/\mu_l^0 \leq 1/\mu_{\text{ma}}^0 < 1/\lambda^0 < 1/\mu_{\bar{l}}^0$ and there is no λ_l that satisfies the conditions

⁸For the consistency of denotation, the constrained indices $l \in \{1, 2\}$ and $\bar{l} \in \{1, 2\} \setminus \{l\}$ are also used here in Case II. However, it should be noted that they are not determined by the same constraint as in Case I.

TABLE II
SUMMARY OF THE OVERALL ALGORITHM FOR NETWORK OPTIMIZATION

1. Initial power allocation. The source nodes solve the MA sum-rate maximization problem (15) and obtain $\mathbf{D}^0, \hat{R}_{lr}(\mathbf{D}_l^0), \forall i$, and $R^{\text{ma}}(\mathbf{D}^0)$. The relay obtains λ^0 and $\hat{R}_{ri}(\lambda^0), \forall i$.
2. Determining the cases. Check if $R^{\text{ma}}(\mathbf{D}^0) \geq \sum_i \hat{R}_{ri}(\lambda^0)$. If yes, proceed to Step 3. Otherwise, proceed to Step 4.
3. Case I. Determine the subcase based on $\mu_1^0, \mu_2^0, \mu_{\text{ma}}^0$, and λ^0 . For Subcase I-1, the relay's optimal strategy is $\mathbf{B}_i = \mathbf{V}_{ri} \mathbf{P}_{ri}(\lambda^0) \mathbf{V}_{ri}^H$ while the source nodes solve problem (16) for transmission power minimization. For Subcase I-2, use Steps 2 to 6 of the algorithm in Table I for deriving the optimal strategies for both the source nodes and the relay.
4. Case II. Determine the subcase based on $\mu_1^0, \mu_2^0, \mu_{\text{ma}}^0$, and λ^0 . For Subcases II-1, II-2, and II-3, the optimal strategy for source i is \mathbf{D}_i^0 and the relay minimizes its transmission power via solving the problems (25) or (26). For Subcase II-4, substitute $R^{\text{max}} = \sum_i \hat{R}_{ri}(\lambda^0)$ in Step 3 of the algorithm in Table I by $R^{\text{max}} = R^{\text{ma}}(\mathbf{D}^0)$ and use Steps 2 to 6 of the algorithm in Table I for finding the optimal strategies for both the source nodes and the relay.

in (27). In this subcase, the maximum $R(\mathbf{B}, \mathbf{D})$ cannot achieve $R^{\text{ma}}(\mathbf{D}^0)$ although $R^{\text{ma}}(\mathbf{D}^0) < \hat{R}_{r1}(\lambda^0) + \hat{R}_{r2}(\lambda^0)$.

Theorem 3: Denote the optimal \mathbf{D}_i as \mathbf{D}_i^* , $\forall i$ and the optimal λ_i as λ_i^* , $\forall i$. In Subcase II-4, the optimal strategies for the source nodes and the relay satisfy the following properties:

1. $\min_i \{1/\mu_i^*\} < 1/\mu_{\text{ma}}^* < 1/\mu_{\text{ma}}^0$;
2. Properties 2–4 in Theorem 2 also apply for Subcase II-4.

Proof: See Appendix E.

According to Theorem 3, the original problem of maximizing $R^{\text{tw}}(\mathbf{B}, \mathbf{D})$ with minimum total power consumption becomes the problem of finding the maximum achievable $R^{\text{tw}}(\mathbf{B}, \mathbf{D})$ with the relay using all its available power and the source nodes using minimum power. From Theorem 3, it can be seen that the optimal solutions in the Subcases I-2 and II-4 share very similar properties. There is also an intuitive way to understand the similarity. Although Subcases I-2 and II-4 are classified to opposite cases according to the initial power allocation, it is the same for both of them that $R(\mathbf{B}, \mathbf{D})$ cannot achieve $R^{\text{ma}}(\mathbf{D}^0)$. As a result, the relay needs to use as much power as possible and the source nodes need to decrease $R^{\text{ma}}(\mathbf{D})$ from $R^{\text{ma}}(\mathbf{D}^0)$ until the maximum $R(\mathbf{B}, \mathbf{D})$ can achieve $R^{\text{ma}}(\mathbf{D})$. This similarity leads to the common properties of the above two subcases. Moreover, due to this similarity between Theorems 2 and 3, Steps 2 to 6 of the algorithm in Table I can be used to derive the optimal solution in Subcase II-4 if the part of $R^{\text{max}} = \sum_i \hat{R}_{ri}(\lambda^0)$ in Step 3 is substituted by $R^{\text{max}} = R^{\text{ma}}(\mathbf{D}^0)$.

Concluding Cases I and II, the complete procedure of deriving the optimal solution to the problem of sum-rate maximization with minimum total transmission power for the scenario of network optimization is summarized in Table II.

C. Discussion: Efficiency and the Effect of Asymmetry

In the previous two subsections, we have found the solution of the network optimization problem for different subcases. Given the solution, the subcases can now be compared and related to each other for more insights.

The solutions found in all subcases are *optimal* in the sense that they achieve the maximum achievable sum-rate with the minimum possible total power consumption. However, the optimal solutions in different subcases may not be equally good from another viewpoint which is power efficiency at the relay

and the source nodes. Specifically, although the power allocation of the source nodes and the relay jointly maximizes the sum-rate of the TWR over the MA and BC phases at optimality, the power allocation of these nodes may not be optimal in their individual phase of transmission, which is MA phase for the source nodes and BC phase for the relay. In fact, the power allocations in the two phases have to compromise with each other in order to achieve optimality over two phases. It is so because of the rate balancing constraints (12a) and (12b). It infers that there is a cost of coordinating the relay and source nodes to achieve optimality over two phases. This cost can be very different depending on the specific subcase. In order to show the difference in this cost, we use the metric *efficiency* defined next. A given power allocation of the relay (source nodes) is considered as *efficient* if it maximizes the BC (MA) phase sum-rate with the actual power consumption of this power allocation. For example, if the power allocation of the relay consumes the power of $P_r \leq P_r^{\max}$ at optimality and achieves sum-rate R^{bc} in the BC phase, then this power allocation is efficient if R^{bc} is the maximum achievable sum-rate in the BC phase with power consumption P_r . It is inefficient otherwise. It can be shown that the chance that the optimal power allocation is efficient for both the relay and the source nodes is small (this only happens in Subcase II-1 and possibly Subcases I-1). Therefore, a joint power allocation of the relay and source nodes is considered to be inefficient if it is inefficient for both the relay and the source nodes, and it is considered to be efficient otherwise. The following conclusions can be drawn for the scenario of network optimization.

First, it can be shown that the optimal power allocation is efficient in Subcase I-1 and generally inefficient in Subcase I-2. Specifically, the optimal power allocation of the relay is always efficient in Subcase I-1 while the optimal power allocation of the source nodes can be either efficient or inefficient. In contrast, the optimal power allocation of the relay is always inefficient in Subcase I-2 while the optimal power allocation of the source nodes is also inefficient in general. For Case II, the optimal power allocation is efficient in Subcases II-1, II-2, and II-3 and generally inefficient in Subcase II-4. Specifically, the optimal power allocation of the source nodes is efficient in Subcases II-1, II-2, and II-3 and generally inefficient in Subcase II-4 while the optimal power allocation of the relay is efficient in Subcase II-1 and inefficient in Subcases II-2, II-3, and II-4.

Second, the optimal power allocation in Subcase I-1 achieves $\hat{R}_{r1}(\lambda^0) + \hat{R}_{r2}(\lambda^0)$. In this subcase, the relay uses its full power and achieves the maximum achievable BC phase sum-rate. The source nodes minimize their power consumption while achieving the maximum sum-rate and in general they do not use up all their available power at optimality. Unlike Subcase I-1, both source nodes may use up their available power in Subcase I-2 while the achieved sum-rate is smaller than $\hat{R}_{r1}(\lambda^0) + \hat{R}_{r2}(\lambda^0)$. Similarly, the optimal power allocation in Subcases II-1, II-2, and II-3 achieves $R^{\text{ma}}(\mathbf{D}^0)$ with the source nodes using their full power while the relay does not necessarily use up its available power. In contrast, the optimal power allocation in Subcase II-4 consumes all the available power of the relay while the achieved sum-rate is smaller than $R^{\text{ma}}(\mathbf{D}^0)$. Therefore, it can be seen that for Subcase I-1 and Subcases II-1, II-2, and II-3, in which the optimal power allocation is efficient, either the maximum possible sum-rate of the MA phase or that of the BC phase can be achieved at optimality. Moreover,

the source nodes and the relay generally do not both use up their available power. In Subcases I-2 and II-4, in which the optimal power allocation is inefficient, the achieved sum-rate is however smaller than the maximum possible sum-rate of the MA phase and also smaller than that of the BC phase, while it is possible that all nodes use up their available power.

Third, it can be shown for Case I that the difference between $\max_i\{1/\mu_i^0\}$ and $\min_i\{1/\mu_i^0\}$ increases in general as the subcase changes from Subcase I-1 to Subcase I-2. Similar result can be observed in Case II. As the subcase changes from Subcase II-1, via Subcases II-2 and II-3, to Subcase II-4, the difference between $\max_i\{1/\mu_i^0\}$ and $\min_i\{1/\mu_i^0\}$ increases.

Last, from the definitions of $\mu_i^0, \forall i$, it can be seen that large difference between $\max_i\{1/\mu_i^0\}$ and $\min_i\{1/\mu_i^0\}$ can be, and most likely is, a result of asymmetry in the power limits, number of antennas, and/or channels at the two source nodes. It will also be shown in detail later in the simulations that such asymmetry can increase the occurrence of the two inefficient subcases, i.e., Subcases I-2 and II-4. In contrary, if the two source nodes have the same available power, same number of antennas, and same channel matrices, then $1/\mu_1^0 = 1/\mu_2^0 > 1/\mu_{\text{ma}}^0$. As a result, only Subcases I-1 and II-1 are possible, in which the optimal power allocation is efficient. Combining this fact with the observations in the above three paragraphs, it can be seen that the asymmetry in the power limits, number of antennas, and/or channels at the two source nodes can lead to a degradation in the power allocation efficiency for the considered scenario of network optimization. As efficiency reveals the cost of coordination between the relay and source nodes required to achieve optimality over the two phases in the network optimization scenario, it can be seen that such cost is low with source node symmetry and high otherwise.

IV. SIMULATIONS

In this section, we provide simulation examples for some results presented earlier and demonstrate the proposed algorithm for network optimization in Table I. The general setup is as follows. The elements of the channels \mathbf{H}_{ri} and $\mathbf{H}_{ir}, \forall i$ are generated from complex Gaussian distribution with zero mean and unit variance. The noise powers $\sigma_i^2, \forall i$ and σ_r^2 are set to 1. The rates $R^{\text{ma}}(\mathbf{D})$, $\bar{R}_{ir}(\mathbf{B}_i)$, and $\hat{R}_{ri}(\mathbf{D}_i)$ are briefly denoted as R^{ma} , \bar{R}_{ir} , and \hat{R}_{ri} , respectively, in all figures.

Example 1: The Process of Finding the Optimal Solution for Network Optimization Subcase I-2 Using the Proposed Algorithm in Table I. The specific setup for this example is as follows. The number of antennas n_1, n_2 , and n_r are set to be 6, 4, and 8, respectively. Power limits for the source nodes are $P_1^{\max} = 2$ W, $P_2^{\max} = 2.5$ W. The relay's power limit is set to $P_r^{\max} = 3$ W. Since the optimality of the solution derived using the algorithm has been proved analytically by the insights from Theorem 2, we focus on demonstrating the iterative process and the convergence of the algorithm. Fig. 1(a) shows instantaneous $R^{\text{tw}}(\mathbf{B}, \mathbf{D})$, $R^{\text{ma}}(\mathbf{D})$, and $R(\mathbf{B}, \mathbf{D})$ versus the number of iterations. From the figure, it can be seen that the above three rates converge very fast. Fig. 1(b) shows the instantaneous $\bar{R}_{ir}(\mathbf{D}_i)$, $\hat{R}_{ri}(\lambda_i), \forall i$ and the power consumption of the source nodes 1 and 2, denoted as P_1 and P_2 , respectively. Two observations can be drawn from Fig. 1(b). First, $\hat{R}_{r2}(\lambda_2) < \bar{R}_{1r}(\mathbf{D}_1)$ and $\hat{R}_{r1}(\lambda_1) = \bar{R}_{2r}(\mathbf{D}_2)$ at optimality since the sum-rate is bounded

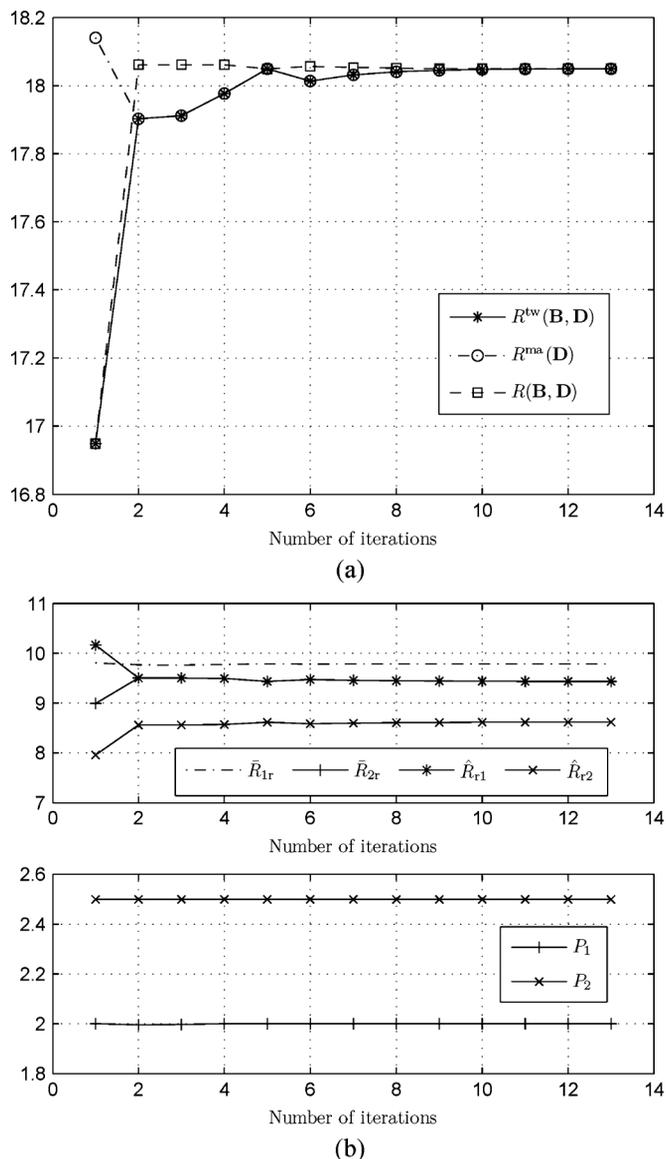


Fig. 1. Illustration of the algorithm in Table I for Subcase I-2. (a) Convergence of $R^{tw}(\mathbf{B}, \mathbf{D})$, $R^{ma}(\mathbf{D})$, and $R(\mathbf{B}, \mathbf{D})$; (b) $\bar{R}_{ir}(\mathbf{D}_i)$, $\hat{R}_{ri}(\lambda_i)$, $\forall i$, and the power consumption of the source nodes over the number of iterations.

by $R^{ma}(\mathbf{D}) < \bar{R}_{1r}(\mathbf{D}_1) + \bar{R}_{2r}(\mathbf{D}_2)$. Second, both source nodes use all available power at optimality. The latter observation verifies the conclusion that for Case I the optimal power allocation in Subcase I-2 is inefficient for using relatively more power and achieving relatively less sum-rate compared to that in Subcase I-1.

Example 2: Comparison with Relay Optimization in Part I of This Two-Part Paper. The specific setup for this example is as follows. The number of antennas at the relay, i.e., n_r , is set to be 5. The power limit of the relay, i.e., P_r^{\max} , is set to be 3 W. The total number of antennas at both source nodes is fixed so that $n_1 + n_2 = 5$. The total available power at both source nodes is also fixed so that $P_1^{\max} + P_2^{\max} = 2$ W. Given the above total number of antennas and total available power at the source nodes, both the relay optimization and the network optimization problems are solved for different combinations of n_1 ,

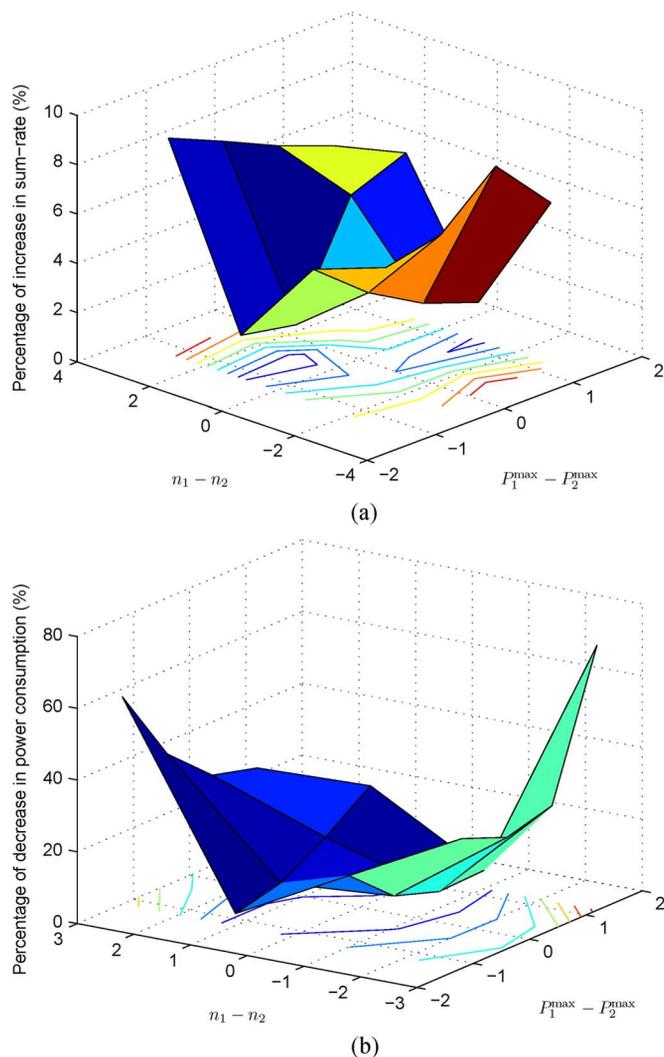


Fig. 2. Improvements as compared to the relay optimization scenario. (a) Percentage of the increase in sum-rate at optimality; (b) percentage of the decrease in power consumption at optimality.

n_2 , P_1^{\max} , and P_2^{\max} each with 100 channel realizations. The percentage of the increase in the average sum-rate and the percentage of the decrease in the average power consumption at optimality of the network optimization problem compared to those at optimality of the relay optimization problem are plotted in Figs. 2(a) and 2(b), respectively. These percentages are shown versus the difference between the number of antennas and the difference between the power limits at the source nodes. From these two figures, it can be seen that although the optimal solution of the network optimization problem on average consumes much less power than that of the relay optimization problem, it still achieves larger sum-rate. Moreover, it can also be seen that the improvements, in either sum-rate or power consumption of the optimal solution of the network optimization problem as compared to that of the relay optimization problem, become more obvious when there is more asymmetry in the system. This is because the source nodes and the relay can jointly optimize their power allocations and therefore cope to some extent with the negative effect of the asymmetry in the system in the network optimization scenario. In contrast, the relay optimization

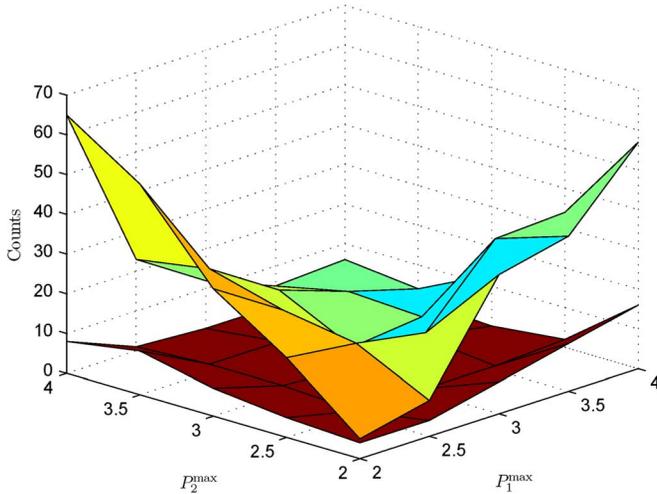


Fig. 3. Number of channel realizations that Subcases I-2 and II-4 appear depending on the asymmetry in P_1^{\max} and P_2^{\max} .

scenario does not have such capability to combat the negative effect of asymmetry.

Example 3: The Effect of Asymmetry in the Scenario of Network Optimization. First, we solve the network optimization problem for different P_1^{\max} and P_2^{\max} given that P_r^{\max} is fixed at 4 W. The number of antennas at the relay is set to 8 and the number of antennas at each source node is set to 4. For each combination of P_1^{\max} and P_2^{\max} , we use 200 channel realizations and solve the resulting 200 network optimization problems. The number of channel realizations that Subcases I-2 and II-4 appear are plotted in Fig. 3. In this figure, the points in the upper surface correspond to the counts of Subcase I-2 while the points in the lower surface correspond to the counts of Subcase II-4. From Fig. 3, it can be seen that in general the count of either Subcase I-2 or Subcase II-4 is the smallest when $P_1^{\max} = P_2^{\max}$. Moreover, for any given P_1^{\max} or P_2^{\max} , the largest count of either Subcase I-2 or Subcase II-4 mostly appears when the difference between P_1^{\max} and P_2^{\max} is the largest.⁹ The above two observations are accurate for most of the times in Fig. 3, which shows that the asymmetry of P_i^{\max} leads to the rise of the occurrence of Subcases I-2 and II-4.

Next we demonstrate the effect of asymmetry in the number of antennas at the source nodes. The number of antennas at the relay is still 8 and P_r^{\max} is still 4 W. However, the number of antennas at the source nodes 1 and 2 are first set to 4 and 6, respectively, and then set to both 6. The network optimization problem is solved for different P_1^{\max} and P_2^{\max} and the sum of the counts of Subcases I-2 and II-4 in 200 channel realizations is plotted in Fig. 4 for each combination of P_1^{\max} and P_2^{\max} . From Fig. 4(a), it can be seen that the sum of the counts of Subcases I-2 and II-4 substantially increases when $n_1 = 4$ and $n_2 = 6$ as compared to the sum of the counts in Fig. 3 on most of the points. However, as shown in Fig. 4(b), when $n_1 = n_2 = 6$, the sum of the counts of Subcases I-2 and II-4 drops to the same level as the sum of the counts in Fig. 3. Therefore, it can be seen that asymmetry in the number of antennas at the source nodes leads to larger chance of Subcases I-2 and II-4.

⁹Note, however, that subcases are also determined by the number of antennas at the relay and the source nodes, the power limit P_r^{\max} , the channel realizations, and other factors instead of only by P_i^{\max} , $\forall i$.

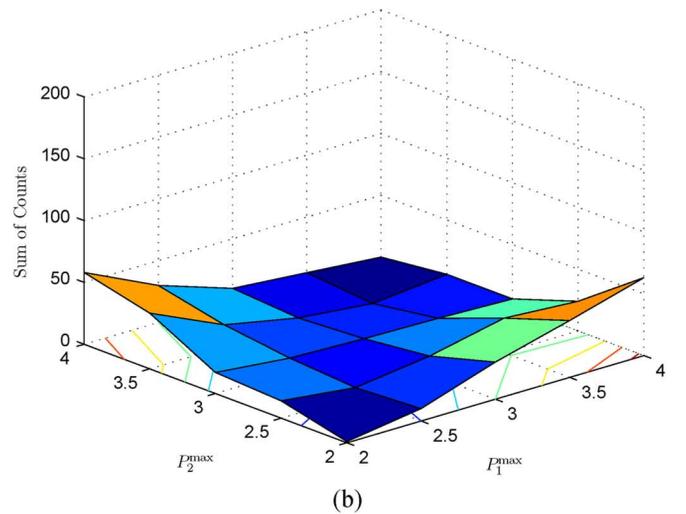
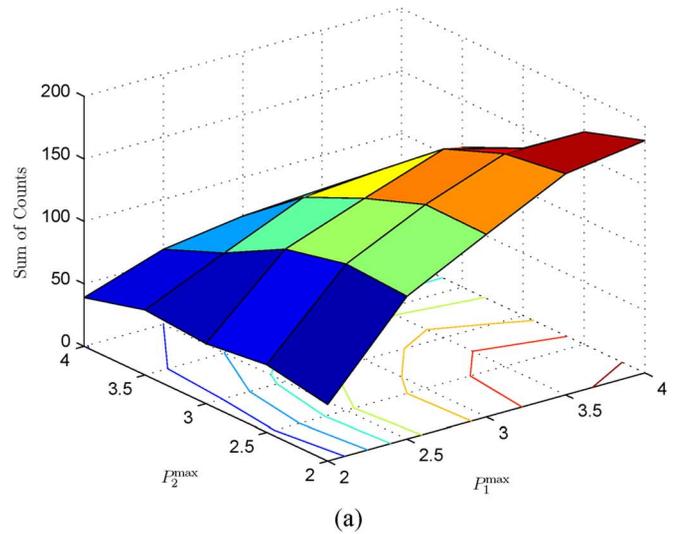


Fig. 4. Illustration of the effect of asymmetry in the number of antennas at the source nodes. (a) Sum of the counts of Subcases I-2 and II-4 versus P_1^{\max} and P_2^{\max} , $n_1 = 4$, $n_2 = 6$; (b) sum of the counts of Subcases I-2 and II-4 versus P_1^{\max} and P_2^{\max} , $n_1 = n_2 = 6$.

Lastly, we show the effect of asymmetry in channels. Instead of generating the real and imaginary parts of each element of \mathbf{H}_{ir} , $\forall i$ from Gaussian distributions with zero mean and unit variance, we use here Gaussian distribution with zero mean and variance v_i to generate the real and imaginary parts of each element of \mathbf{H}_{ir} , $\forall i$. For each combination of v_1 and v_2 , we use 200 channel realizations and solve the resulting 200 network optimization problems. The number of antennas at the relay is set to 6 and the number of antennas at each source node is set to 4. The power limits are $P_r^{\max} = 5$ W and $P_i^{\max} = 3$ W, $\forall i$. The sum of the counts of Subcases I-2 and II-4 is plotted in Fig. 5 versus v_1 and v_2 . In Fig. 5(a), channel reciprocity is not assumed and the real and imaginary parts of each element of \mathbf{H}_{ri} , $\forall i$ are generated from Gaussian distribution with zero mean and unit variance. In Fig. 5(b), channel reciprocity is assumed i.e., $\mathbf{H}_{ri} = \mathbf{H}_{ir}^T$, $\forall i$ where $(\cdot)^T$ represents the transpose. It can be seen from both Figs. 5(a) and 5(b) that the sum of the counts of Subcases I-2 and II-4 tends to increase when the difference between v_1 and v_2 becomes larger. Therefore, Fig. 5 clearly shows that the asymmetry in the channels also leads to larger chance of the inefficient Subcases I-2 and II-4.

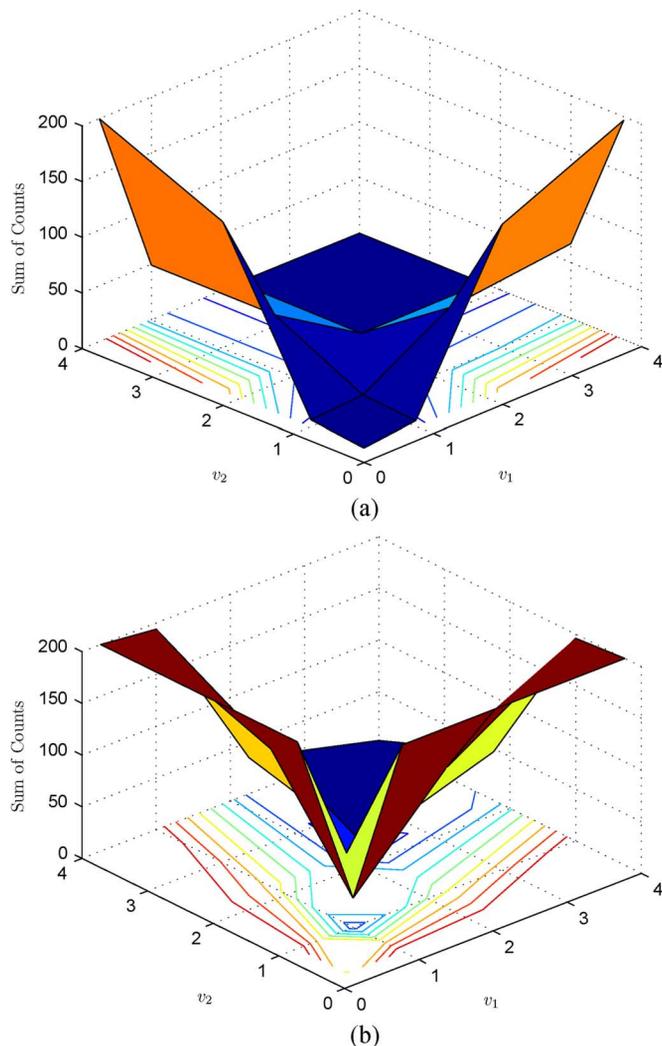


Fig. 5. Illustration of the effect of asymmetry in channel statistics. (a) Sum of the counts of Subcases I-2 and II-4 versus v_1 and v_2 (without assuming channel reciprocity); (b) sum of the counts of Subcases I-2 and II-4 versus v_1 and v_2 (assuming channel reciprocity).

V. CONCLUSION

In Part II of this two-part paper, we have solved the problem of sum-rate maximization using minimum total transmission power for MIMO DF TWR in the scenario of network optimization. For finding the optimal solution, we have studied the original problem in two cases each of which has several subcases. It has been shown that for all except two subcases, the originally nonconvex problem can be simplified into corresponding convex optimization problem(s). For the remaining two subcases, we have found the properties that the optimal solution must satisfy and have proposed the algorithm to find the optimal solution based on these properties. We have shown that the optimal power allocation in these two subcases are inefficient in the sense that it always consumes all the available power of the relay (and sometimes all the available power of the source nodes as well) yet cannot achieve the maximum sum-rate of either the MA or BC phase. We have also shown that the asymmetry in the power limits, number of antennas, and channels leads to a higher probability of the above-mentioned two inefficient subcases. Together with Part I of this work, we have provided a

complete and detailed study of the problem of sum-rate maximization using minimum power consumption for MIMO DF TWR.

APPENDIX

A. Proof of Lemma 1

Proof for claim 1: Given $\tilde{\mathbf{D}}$ as defined in the lemma, it follows that $R^{ma}(\tilde{\mathbf{D}}) = \bar{R}_{jr}(\tilde{\mathbf{D}}_j)$. From the definitions (9a)–(9c), it can be seen that $R^{ma}(\tilde{\mathbf{D}}) > \bar{R}_{jr}(\tilde{\mathbf{D}}_j) = \bar{R}_{jr}(\mathbf{D}_j)$ if $1/\mu_{ma}(\tilde{\mathbf{D}}) > 1/\mu_j$. Therefore, it is necessary that $1/\mu_{ma}(\tilde{\mathbf{D}}) \leq 1/\mu_j$.

Proof for claim 2: First, note that $R^{ma}(\hat{\mathbf{D}})$ is a continuous and strictly increasing function of t in $[0, 1]$. Second, based on the definition (9c), it follows that $R^{ma}(\hat{\mathbf{D}})$ is a strictly increasing function of $1/\mu_{ma}(\hat{\mathbf{D}})$ when $1/\mu_{ma}(\hat{\mathbf{D}}) > \min\{1/\alpha_i(k), \forall i, \forall k\}$, or equivalently, $R^{ma}(\hat{\mathbf{D}}) > 0$. Since $\text{Tr}\{\mathbf{D}_1\} > 0$ and $\text{Tr}\{\mathbf{D}_2\} > 0$, we have $R^{ma}(\hat{\mathbf{D}}) > 0$ for any $t \in [0, 1]$. Thus, given the fact that $1/\mu_{ma}(\hat{\mathbf{D}}) \leq 1/\mu_j$ when $t = 0$ and that $1/\mu_{ma}(\hat{\mathbf{D}}) = 1/\mu_{ma}(\mathbf{D}) > 1/\mu_j$ when $t = 1$, it can be seen that there exists $\hat{t} \in [0, 1]$ such that $1/\mu_{ma}(\hat{\mathbf{D}}) = 1/\mu_j$ when $t = \hat{t}$. Using the result of Lemma 1 in Part I of this two-part paper [2], i.e., $1/\mu_{ma} < \max\{1/\mu_1, 1/\mu_2\}$, it can be seen that $1/\mu_i(\hat{\mathbf{D}}_i) > 1/\mu_{ma}(\hat{\mathbf{D}}) = 1/\mu_j$ when $t = \hat{t}$. ■

B. Proof of Lemma 2

First we prove $\lambda_j = \mu_i > \mu_{ma}$ if $\lambda_i < \lambda_j$. Using Lemma 2 in Part I [2], it can be seen that we have $\lambda_1 = \lambda_2$ if $\min\{1/\mu_1, 1/\mu_2\} \geq 1/\mu_{ma}$ at optimality. Therefore, we have $\min\{1/\mu_1, 1/\mu_2\} < 1/\mu_{ma}$ given that $\lambda_1 \neq \lambda_2$. Using the same lemma and the constraint (13a), it can be further concluded that $1/\mu_i < 1/\mu_{ma}$ at optimality given that $\lambda_i < \lambda_j$. Otherwise, the constraint (13b) cannot be satisfied. Therefore, $1/\mu_j > 1/\mu_{ma}$ according to Lemma 1 in Part I [2]. Due to the constraint (13a), we must have $1/\lambda_j \leq 1/\mu_i$ at optimality. Moreover, from Lemma 2 in Part I and the assumption that $\lambda_i < \lambda_j$, it can be seen that $1/\lambda_j < 1/\mu_i$ is not optimal. Therefore, $1/\lambda_j = 1/\mu_i$ if $\lambda_i < \lambda_j$. Following the same approach, we can prove $\lambda_j = \mu_i > \mu_{ma}$ if $\mu_i > \mu_{ma}$.

C. Proof of Theorem 1

Recall the definitions of μ_1 , μ_2 , and μ_{ma} in (9a)–(9c). Considering the constraints (16b)–(16d) in the problem (16), it can be seen that at optimality we must have $\mu_{ma}^* \leq \lambda^0$, $\mu_1^* \leq \lambda^0$, and $\mu_2^* \leq \lambda^0$. Otherwise, the above mentioned constraints cannot be satisfied. We will prove Theorem 1 by contradiction.

Assume that $\mu_{ma}^* \neq \lambda^0$ at optimality, then $\mu_{ma}^* < \lambda^0$ according to the above paragraph. Using the result of Lemma 1 in Part I of this two-part paper [2], i.e., $1/\mu_{ma} < \max\{1/\mu_1, 1/\mu_2\}$, and given that $\mu_1^* \leq \lambda^0$ and $\mu_2^* \leq \lambda^0$, there are only two possible situations as follows: a) $\max\{1/\mu_1^*, 1/\mu_2^*\} > 1/\mu_{ma}^* > \min\{1/\mu_1^*, 1/\mu_2^*\} \geq 1/\lambda^0$ and b) $\max\{1/\mu_1^*, 1/\mu_2^*\} \geq \min\{1/\mu_1^*, 1/\mu_2^*\} \geq 1/\mu_{ma}^* > 1/\lambda^0$. Assume without loss of generality that $\max\{1/\mu_1^*, 1/\mu_2^*\} = 1/\mu_1^*$ and $\min\{1/\mu_1^*, 1/\mu_2^*\} = 1/\mu_2^*$. If it is Situation a), then we have $1/\mu_1^* > 1/\mu_{ma}^* > 1/\mu_2^* \geq 1/\lambda^0$. Use Lemma 1 of this paper with $\hat{\mathbf{D}}_i = t\mathbf{D}_1^*$ and $\hat{\mathbf{D}}_j = \mathbf{D}_2^*$. As proved in Lemma 1, there exists $t \in [0, 1)$ such that $1/\mu_{ma}([t\mathbf{D}_1^*, \mathbf{D}_2^*]) = 1/\mu_2^*$. Since $1/\mu_2^* \geq 1/\lambda^0$, we have $1/\mu_1([t\mathbf{D}_1^*]) > 1/\mu_{ma}([t\mathbf{D}_1^*, \mathbf{D}_2^*]) = 1/\mu_2^* \geq 1/\lambda^0$, which

indicates that $\hat{\mathbf{D}} = [t\mathbf{D}_1^*, \mathbf{D}_2^*]$ also satisfies (16b)–(16e) while $\text{Tr}\{t\mathbf{D}_1^*\} + \text{Tr}\{\mathbf{D}_2^*\} < \text{Tr}\{\mathbf{D}_1^*\} + \text{Tr}\{\mathbf{D}_2^*\}$. It contradicts the fact that $\mathbf{D}^* = [\mathbf{D}_1^*, \mathbf{D}_2^*]$ is the optimal solution to the problem (16). Therefore, Situation a) is impossible. If it is Situation b), there exist two following possible sub-situations: Sub-situation b-1) there exists $i \in \{1, 2\}$ such that $1/\mu_{\text{ma}}(\hat{\mathbf{D}}) = 1/\lambda^0$ and $1/\mu_i(\hat{\mathbf{D}}_i) \geq 1/\mu_{\text{ma}}(\hat{\mathbf{D}})$ where $\hat{\mathbf{D}} = [\hat{\mathbf{D}}_1, \hat{\mathbf{D}}_2]$ with $\hat{\mathbf{D}}_i = t_i\mathbf{D}_i^*$ and $\hat{\mathbf{D}}_j = \mathbf{D}_j^*$ for some $t_i \in [0, 1)$ and Sub-situation b-2) there does not exist $t_i \in [0, 1)$ such that $1/\mu_{\text{ma}}(\hat{\mathbf{D}}) = 1/\lambda^0$ and $1/\mu_i(\hat{\mathbf{D}}_i) \geq 1/\mu_{\text{ma}}(\hat{\mathbf{D}})$ where $\hat{\mathbf{D}} = [\hat{\mathbf{D}}_1, \hat{\mathbf{D}}_2]$ with $\hat{\mathbf{D}}_i = t_i\mathbf{D}_i^*$ and $\hat{\mathbf{D}}_j = \mathbf{D}_j^*$ for either $i = 1$ or $i = 2$. In Sub-situation b-1), it can be seen that $\hat{\mathbf{D}}$ satisfies (16b)–(16e) while $\text{Tr}\{t_i\mathbf{D}_i^*\} + \text{Tr}\{\mathbf{D}_j^*\} < \text{Tr}\{\mathbf{D}_1^*\} + \text{Tr}\{\mathbf{D}_2^*\}$. It contradicts the fact that $\mathbf{D}^* = [\mathbf{D}_1^*, \mathbf{D}_2^*]$ is the optimal solution to the problem (16). Therefore, Sub-situation b-1) is impossible. If it is Sub-situation b-2), it indicates that with $t_i \in [0, 1)$, for either $i = 1$ or $i = 2$, such that $1/\mu_{\text{ma}}(\hat{\mathbf{D}}) = 1/\lambda^0$, we have $1/\mu_i(\hat{\mathbf{D}}_i) = 1/\mu_i(t_i\mathbf{D}_i^*) < 1/\mu_{\text{ma}}(\hat{\mathbf{D}}) = 1/\lambda^0$. As a result, there exists $t'_i \in (t_i, 1)$ such that $1/\mu_i(t'_i\mathbf{D}_i^*) = 1/\lambda^0$ and $1/\mu_{\text{ma}}(\mathbf{D}') > 1/\lambda^0$ where $\mathbf{D}' = [\mathbf{D}'_1, \mathbf{D}'_2]$ with $\mathbf{D}'_i = t'_i\mathbf{D}_i^*$ and $\mathbf{D}'_j = \mathbf{D}_j^*$. Note that $1/\mu_{\text{ma}}(\mathbf{D}') > 1/\lambda^0$ because if $1/\mu_i(\mathbf{D}'_i) = 1/\lambda^0$ and $1/\mu_{\text{ma}}(\mathbf{D}') = 1/\lambda^0$ then we have Sub-situation b-1) instead of Sub-situation b-2). Recalling that $1/\mu_j(\mathbf{D}_j^*) > 1/\mu_{\text{ma}}(\mathbf{D}^*) > 1/\mu_{\text{ma}}(\mathbf{D}')$, we have $1/\mu_j(\mathbf{D}_j^*) > 1/\mu_{\text{ma}}(\mathbf{D}') > 1/\mu_i(\mathbf{D}'_i) = 1/\lambda^0$. It indicates that by changing \mathbf{D}_i^* at optimality to $\mathbf{D}'_i = t'_i\mathbf{D}_i^*$, and thus, using less power than $\text{Tr}\{\mathbf{D}_1^*\} + \text{Tr}\{\mathbf{D}_2^*\}$ while satisfying (16b)–(16e), Sub-situation b-2) changes to Situation a). As it is proved that Situation a) is impossible at optimality, so it is Sub-situation b-2).

Therefore, it is proved that the assumption $\mu_{\text{ma}}^* \neq \lambda^0$ must lead to either of two situations both of which are impossible at optimality. Thus, it is impossible that $\mu_{\text{ma}}^* \neq \lambda^0$. As a result, we must have $\mu_{\text{ma}}^* = \lambda^0$. This completes the proof. ■

D. Proof of Theorem 2

Proof of Property 1: First we show that $1/\mu_{\text{ma}}^* < 1/\lambda^0$. Since the maximum $\bar{R}_{\text{r}}(\mathbf{D}_i)$, as the objective function of the problem (17), cannot achieve $\hat{R}_{\text{r}}(\lambda^0)$ in Subcase I-2, it can be seen that $1/\mu_l < 1/\lambda^0$ whenever $1/\mu_{\text{ma}} \geq 1/\lambda^0$ and $1/\mu_{\text{r}} \geq 1/\lambda^0$. As a result, any \mathbf{D} that leads to $1/\mu_{\text{ma}} \geq 1/\lambda^0$ is not optimal. The reason is that in such a case the optimal relay power allocation requires $1/\lambda_{\text{r}} = 1/\mu_l < 1/\lambda^0$ according to Lemma 2 and such relay power allocation leads to a BC phase sum-rate $\sum_i \hat{R}_{\text{r}i}(\lambda_i)$ which is less than $\hat{R}_{\text{r}1}(\lambda^0) + \hat{R}_{\text{r}2}(\lambda^0)$ according to Lemma 2 in Part I [2]. Since $1/\mu_{\text{ma}} \geq 1/\lambda^0$ implies that $R^{\text{ma}}(\mathbf{D}) \geq \hat{R}_{\text{r}1}(\lambda^0) + \hat{R}_{\text{r}2}(\lambda^0)$, it can be seen that the constraint (12b) is not satisfied and therefore such strategies cannot be optimal. Next we show that $\min_i \{1/\mu_i^*\} < 1/\mu_{\text{ma}}^*$. Assuming that $1/\mu_{\text{ma}}^* \leq \min_i \{1/\mu_i^*\}$, it leads to $1/\mu_{\text{ma}}^* < 1/\lambda^0$ given that the problem (16) is infeasible. Moreover, it also leads to the result that $\lambda_i^* = \mu_{\text{ma}}^*$, $\forall i$. However, it is not difficult to see that $R^{\text{ma}}(\mathbf{D})$, $\sum_i \hat{R}_{\text{r}i}(\lambda_i)$, and eventually $R^{\text{tw}}(\mathbf{B}, \mathbf{D})$ can be increased in this case through appropriately increasing $1/\mu_{\text{ma}}$, which is feasible since $1/\mu_{\text{ma}}^0 > 1/\lambda^0 > 1/\mu_{\text{ma}}^*$, and also increasing at least one of $1/\lambda_{\text{r}}^0$ and $1/\lambda_l$, which is also feasible

since $1/\lambda_{\text{r}}^* = 1/\mu_{\text{ma}}^* < 1/\lambda^0$, given that $1/\mu_{\text{ma}}^* \leq \min_i \{1/\mu_i^*\}$ and $1/\mu_{\text{ma}}^* < 1/\lambda^0$. It contradicts the assumption that \mathbf{D}^* and \mathbf{B}^* are the optimal solution. Therefore, $1/\mu_{\text{ma}}^* > \min_i \{1/\mu_i^*\}$.

Proof of Property 2: Given the fact that $1/\mu_{\text{ma}}^* < 1/\lambda^0$, the problem boils down to finding the maximum $1/\mu_{\text{ma}}$ such that the corresponding rate $R^{\text{ma}}(\mathbf{D})$ can also be achieved by the BC phase sum-rate $\sum_i \hat{R}_{\text{r}i}(\lambda_i)$ subject to the first constraint in (10) and the constraint that $\min_i \{1/\lambda_i\} = \min_i \{1/\mu_i\}$ as stated in Lemma 2. Since the maximum $\sum_i \hat{R}_{\text{r}i}(\lambda_i)$ cannot achieve $R^{\text{ma}}(\mathbf{D})$ subject to the above-mentioned constraints as long as $1/\mu_{\text{ma}} \geq 1/\lambda^0$, the problem is equivalent to finding the $\lambda_i, \forall i$ to maximize $\sum_i \hat{R}_{\text{r}i}(\lambda_i)$ such that the resulting $\sum_i \hat{R}_{\text{r}i}(\lambda_i)$ is achievable by $R^{\text{ma}}(\mathbf{D})$ subject to the constraint that $\min_i \{1/\lambda_i\}$ is achievable by $\min_i \{1/\mu_i\}$ (in addition to the power constraints). Consider the problem of maximizing $R^{\text{ma}}(\mathbf{D})$ subject to the constraint that $\min_i \{1/\mu_i\} \geq C$ where C is a constant. Note that the maximum of this problem is a non-increasing function of C as long as the problem is feasible and $C < 1/\mu_{\text{ma}}$. Recall from Property 1 that $\min_i \{1/\mu_i^*\} < 1/\mu_{\text{ma}}^* < 1/\lambda^0$. Assume that the relay does not use full power at optimality, then the maximum achievable $\sum_i \hat{R}_{\text{r}i}(\lambda_i)$ and the maximum achievable $R^{\text{ma}}(\mathbf{D})$ can be both increased subject to all the above constraints by appropriately decreasing $\min_i \{1/\mu_i\}$ (and thereby increasing the maximum achievable $R^{\text{ma}}(\mathbf{D})$) while letting the relay decrease $\min_i \{1/\lambda_i\}$ accordingly and at the same time using all the remaining power to increase $\max_i \{1/\lambda_i\}$ (and thereby increasing the maximum achievable $\sum_i \hat{R}_{\text{r}i}(\lambda_i)$). It contradicts the assumption, which infers that the relay must use full power at optimality.

Proof of Property 3: Define the index $i^- \triangleq \arg \min_i \{1/\mu_i\}$. Recall from the proof of Property 1 that $1/\mu_{\text{ma}}^* < 1/\lambda^0$. As a result, $R^{\text{ma}}(\mathbf{D}^*)$ is not the maximum $R^{\text{ma}}(\mathbf{D})$ that can be achieved, which implies that there exists \mathbf{D}^{s} such that $R^{\text{ma}}(\mathbf{D}^{\text{s}}) > R^{\text{ma}}(\mathbf{D}^*)$ and $\bar{R}_{i-\text{r}}(\mathbf{D}_{i^-}^{\text{s}}) > \bar{R}_{i-\text{r}}(\mathbf{D}_{i^-}^*) - \delta$ where δ is a positive number. Define $\mathbf{Z} \triangleq \mathbf{H}_{1\text{r}}\mathbf{D}_1\mathbf{H}_{1\text{r}}^{\text{H}} + \mathbf{H}_{2\text{r}}\mathbf{D}_2\mathbf{H}_{2\text{r}}^{\text{H}}$. It can be seen that $R^{\text{ma}}(\mathbf{D})$ is a concave function of \mathbf{Z} . If \mathbf{D}^* is not the optimal solution to the problem of maximizing $\min_i \{1/\mu_i\}$ subject to the constraints in (18), there exists \mathbf{D}^{q} such that $R^{\text{ma}}(\mathbf{D}^{\text{q}}) \geq R^{\text{ma}}(\mathbf{D}^*)$ and $\bar{R}_{i-\text{r}}(\mathbf{D}_{i^-}^{\text{q}}) > \bar{R}_{i-\text{r}}(\mathbf{D}_{i^-}^*)$. Then, for any $0 < \alpha < 1$, there is a \mathbf{D}^{c} such that $\mathbf{D}_i^{\text{c}} = \alpha\mathbf{D}_i^{\text{q}} + (1 - \alpha)\mathbf{D}_i^{\text{s}}, \forall i$. Moreover, for any α such that

$$\frac{\bar{R}_{i-\text{r}}(\mathbf{D}_{i^-}^{\text{s}}) - \bar{R}_{i-\text{r}}(\mathbf{D}_{i^-}^{\text{s}_-})}{\bar{R}_{i-\text{r}}(\mathbf{D}_{i^-}^{\text{q}}) - \bar{R}_{i-\text{r}}(\mathbf{D}_{i^-}^{\text{s}_-})} < \alpha < 1 \quad (28)$$

it can be shown that $\bar{R}_{i-\text{r}}(\mathbf{D}_{i^-}^{\text{c}}) > \bar{R}_{i-\text{r}}(\mathbf{D}_{i^-}^*)$ using the fact that $\bar{R}_{i-\text{r}}(\mathbf{D}_i)$ is concave with respect to $\mathbf{D}_i, \forall i$. Denoting $\mathbf{Z}^{\text{q}} = \mathbf{H}_{1\text{r}}\mathbf{D}_1^{\text{q}}\mathbf{H}_{1\text{r}}^{\text{H}} + \mathbf{H}_{2\text{r}}\mathbf{D}_2^{\text{q}}\mathbf{H}_{2\text{r}}^{\text{H}}$ and $\mathbf{Z}^{\text{s}} = \mathbf{H}_{1\text{r}}\mathbf{D}_1^{\text{s}}\mathbf{H}_{1\text{r}}^{\text{H}} + \mathbf{H}_{2\text{r}}\mathbf{D}_2^{\text{s}}\mathbf{H}_{2\text{r}}^{\text{H}}$, it can be shown that $\mathbf{D}_i^{\text{c}}, \forall i$ lead to $\mathbf{Z}^{\text{c}} = \alpha\mathbf{Z}^{\text{q}} + (1 - \alpha)\mathbf{Z}^{\text{s}}$ and therefore $R^{\text{ma}}(\mathbf{D}^{\text{c}}) \geq \alpha R^{\text{ma}}(\mathbf{D}^{\text{q}}) + (1 - \alpha)R^{\text{ma}}(\mathbf{D}^{\text{s}}) > R^{\text{ma}}(\mathbf{D}^*)$. Hence, if \mathbf{D}^* does not maximize $\bar{R}_{i-\text{r}}(\mathbf{D}_{i^-})$ subject to the constraints in (18), then $\bar{R}_{i-\text{r}}(\mathbf{D}_{i^-})$ and $R^{\text{ma}}(\mathbf{D})$ can be simultaneously increased. The fact that $\bar{R}_{i-\text{r}}(\mathbf{D}_{i^-})$ can be

increased means that $\min_i \{1/\mu_i\}$ can be increased, which implies that the BC phase sum-rate $\sum_i \hat{R}_{r_i}(\lambda_i)$ can be increased according to Lemma 2 in Part I [2] subject to the constraint that $\min_i \{1/\lambda_i\} = \min_i \{1/\mu_i\}$ as implied by Lemma 2. Given this result, the fact that $R^{\text{ma}}(\mathbf{D})$ can be simultaneously increased suggests that $R^{\text{tw}}(\mathbf{B}, \mathbf{D})$ can be increased. This contradicts the fact that \mathbf{D}^* is the optimal solution that maximizes $R^{\text{tw}}(\mathbf{B}, \mathbf{D})$ with \mathbf{D}^* subject to the related constraints. Therefore, \mathbf{D}^* must maximize $\min_i \{1/\mu_i\}$ subject to (18).

Proof of Property 4: It can be seen that the maximum achievable $1/\mu_l$ subject to the constraints

$$R^{\text{ma}}(\mathbf{D}) \geq R^{\text{obj}}, \quad \text{Tr}(\mathbf{D}_i) \leq P_i^{\text{max}}, \quad \forall i \quad (29)$$

is a non-increasing function of R^{obj} . If $1/\mu_l^* \leq 1/\mu_l^*$, according to Property 1 of this theorem and the fact that $1/\mu_{\text{ma}} < \max_i \{1/\mu_i\}$, it indicates that $1/\mu_l^* > 1/\mu_{\text{ma}}^*$. Since $1/\mu_{\text{ma}}^0 > 1/\mu_l^0$ and the maximum achievable $1/\mu_l$ is a non-increasing function of R^{obj} , there exists $\tilde{\mathbf{D}}$ such that $1/\mu_l^* \geq 1/\tilde{\mu}_l$ and $1/\tilde{\mu}_l = 1/\tilde{\mu}_{\text{ma}} \geq 1/\mu_{\text{ma}}^*$. Using $1/\mu_{\text{ma}} < \max_i \{1/\mu_i\}$ from Lemma 1 in Part I [2], it infers that $1/\tilde{\mu}_l > 1/\tilde{\mu}_l = 1/\tilde{\mu}_{\text{ma}}$ at this point. Since the maximum $\hat{R}_{r_i}(\mathbf{D}_i)$ cannot achieve $\hat{R}_{r_i}(\lambda^0)$ in the problem (17), it can be seen that $1/\tilde{\mu}_l = 1/\tilde{\mu}_{\text{ma}} < 1/\lambda^0$. In such a case, the optimal strategy of the relay is to use $1/\lambda_i = 1/\tilde{\mu}_{\text{ma}} < 1/\lambda^0, \forall i$, which does not consume the full power of the relay. Therefore, according to Property 2 of this theorem, when $1/\tilde{\mu}_l = 1/\tilde{\mu}_{\text{ma}}$, the $R^{\text{tw}}(\mathbf{B}, \mathbf{D})$ that can be achieved, specifically $R^{\text{ma}}(\tilde{\mathbf{D}})$, is not the maximum that $R^{\text{tw}}(\mathbf{B}, \mathbf{D})$ can achieve. Moreover, since $1/\tilde{\mu}_{\text{ma}} \geq 1/\mu_{\text{ma}}^*$, it can be seen that $R^{\text{ma}}(\mathbf{D}^*) \leq R^{\text{ma}}(\tilde{\mathbf{D}})$. As a result, $R^{\text{tw}}(\mathbf{B}^*, \mathbf{D}^*) = R^{\text{ma}}(\mathbf{D}^*) \leq R^{\text{ma}}(\tilde{\mathbf{D}})$. Using the above-proved fact that $R^{\text{ma}}(\tilde{\mathbf{D}})$ is not the maximum that $R^{\text{tw}}(\mathbf{B}, \mathbf{D})$ can achieve, this result obtained under the assumption $1/\mu_l^* \leq 1/\mu_l^*$ contradicts the assumption that \mathbf{B}^* and \mathbf{D}^* are optimal. Therefore, the assumption that $1/\mu_l^* \leq 1/\mu_l^*$ must be invalid. ■

E. Proof of Theorem 3

The proof follows the same route as the proof of Theorem 2.

Proof of Property 1: As there exists no λ_l which satisfies the constraints in (27), it can be seen that $\sum_i \hat{R}_{r_i}(\lambda_i)$ cannot achieve

$R^{\text{ma}}(\mathbf{D}^0)$ subject to the constraint $\lambda_l = \mu_l^0$, which is necessary as stated in Lemma 2. Therefore, it is necessary that $1/\mu_{\text{ma}} < 1/\mu_{\text{ma}}^0$. Given that $1/\mu_{\text{ma}} < 1/\mu_{\text{ma}}^0$, it can be shown that the resulting $R^{\text{tw}}(\mathbf{B}, \mathbf{D})$ is not maximized if $1/\mu_{\text{ma}} \leq \min_i \{1/\mu_i^*\}$.

Therefore, it is necessary that $1/\mu_{\text{ma}} > \min_i \{1/\mu_i^*\}$.

Proof of Properties 2–3 from Appendix D can be applied here after we substitute all λ^0 therein to μ_{ma}^0 . Proof of Property 4 of Theorem 2 can be directly applied here. ■

REFERENCES

- [1] B. Rankov and A. Wittneben, "Spectral efficient protocols for half-duplex fading relay channels," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 2, pp. 379–389, Feb. 2007.
- [2] J. Gao, S. A. Vorobyov, H. Jiang, J. Zhang, and M. Haardt, "Sum-rate maximization with minimum power consumption for MIMO DF two-way relaying—Part I: Relay optimization," *IEEE Trans. Signal Process.*, vol. 61, no. 14, pp. 3563–3577, 2013.

- [3] T. J. Oechtering, R. F. Wyrembelski, and H. Boche, "Multiantenna bidirectional broadcast channels—Optimal transmit strategies," *IEEE Trans. Signal Process.*, vol. 57, no. 5, pp. 1948–1958, May 2009.
- [4] 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.
- [5] C. Y. Leow, Z. Ding, and K. K. Leung, "Joint beamforming and power management for nonregenerative MIMO two-way relaying channels," *IEEE Trans. Veh. Technol.*, vol. 60, no. 9, pp. 4374–4383, Nov. 2011.
- [6] 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.
- [7] J. Zou, H. Luo, M. Tao, and R. Wang, "Joint source and relay optimization for non-regenerative MIMO two-way relay systems with imperfect CSI," *IEEE Trans. Wireless Commun.*, vol. 11, no. 9, pp. 3305–3315, Sep. 2012.
- [8] S. J. Kim, N. Devroye, P. Mitran, and V. Tarokh, "Achievable rate regions and performance comparison of half duplex bi-directional relaying protocols," *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6405–6418, Oct. 2011.
- [9] J. Gao, J. Zhang, S. A. Vorobyov, H. Jiang, and M. Haardt, "Power allocation/beamforming for DF MIMO two-way relaying: Relay and network optimization," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Anaheim, CA, USA, Dec. 2012, pp. 5657–5662.
- [10] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 5, pp. 684–702, Jun. 2003.
- [11] M. Chen and A. Yener, "Power allocation for F/TDMA multiuser two-way relay networks," *IEEE Trans. Wireless Commun.*, vol. 9, no. 2, pp. 546–551, Feb. 2010.
- [12] C.-H. Liu and F. Xue, "Network coding for two-way relaying: Rate region, sum rate and opportunistic scheduling," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Beijing, China, May 2008, pp. 1044–1049.
- [13] J. Liu, M. Tao, Y. Xu, and X. Wang, "Superimposed XOR: A new physical layer network coding scheme for two-way relay channels," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Honolulu, HI, USA, Dec. 2009.
- [14] W. Yu, R. Wong, S. Boyd, and J. M. Cioffi, "Iterative water-filling for Gaussian vector multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 1, pp. 145–152, Jan. 2004.



Jie Gao (S'09) received the B.Eng. degree in electronics and information engineering from Huazhong University of Science and Technology, Wuhan, China, in 2007 and the M.Sc. degree in electrical engineering from the University of Alberta, Edmonton, Alberta, Canada, in 2009. He is currently working towards his Ph.D. degree at the University of Alberta. His research interests include applications of optimization and game theoretic methods in signal processing for multiuser communications and cooperative networks.



Sergiy A. Vorobyov (M'02–SM'05) received the M.Sc. and Ph.D. degrees in systems and control from Kharkiv National University of Radio Electronics, Ukraine, in 1994 and 1997, respectively.

He is a Professor (currently on leave) with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB, Canada and with the Department of Signal Processing and Acoustics, Aalto University, Finland. He has been with the University of Alberta as an Assistant Professor (2006–2010), Associate Professor (2010–2012), and Full Professor since 2012. Since his graduation, he also held various research and faculty positions at Kharkiv National University of Radio Electronics, Ukraine; the Institute of Physical and Chemical Research (RIKEN), Japan; McMaster University, Canada; Duisburg-Essen University and Darmstadt University of Technology, Germany; and the Joint Research Institute between Heriot-Watt University and Edinburgh University, U.K. He has also held short-term visiting positions at Technion, Haifa, Israel and

Ilmenau University of Technology, Ilmenau, Germany. His research interests include statistical and array signal processing, applications of linear algebra, optimization, and game theory methods in signal processing and communications, estimation, detection, and sampling theories, and cognitive systems.

Dr. Vorobyov is a recipient of the 2004 IEEE Signal Processing Society Best Paper Award, the 2007 Alberta Ingenuity New Faculty Award, the 2011 Carl Zeiss Award (Germany), the 2012 NSERC Discovery Accelerator Award, and other awards. He served as an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 2006 to 2010 and for the IEEE SIGNAL PROCESSING LETTERS from 2007 to 2009. He was a member of the Sensor Array and Multi-Channel Signal Processing Committee of the IEEE Signal Processing Society from 2007 to 2012. He is a member of the Signal Processing for Communications and Networking Committee since 2010. He has served as the Track Chair for Asilomar 2011, Pacific Grove, CA, the Technical Co-Chair for IEEE CAMSAP 2011, Puerto Rico, and the Tutorial Chair of ISWCS 2013, Ilmenau, Germany.



Hai Jiang (M'07) received the B.Sc. and M.Sc. degrees in electronics engineering from Peking University, Beijing, China, in 1995 and 1998, respectively, and the Ph.D. degree (with an Outstanding Achievement in Graduate Studies Award) in electrical engineering from the University of Waterloo, Waterloo, Ontario, Canada, in 2006.

Since July 2007, he has been a faculty member with the University of Alberta, Edmonton, Alberta, Canada, where he is currently an Associate Professor at the Department of Electrical & Computer

Engineering. His research interests include radio resource management, cognitive radio networking, and cross-layer design for wireless multimedia communications.

Dr. Jiang is an Editor for the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY and the IEEE WIRELESS COMMUNICATIONS LETTERS. He served as a Co-Chair for the Wireless and Mobile Networking Symposium at the IEEE International Conference on Communications (ICC) in 2010. He received an Alberta Ingenuity New Faculty Award in 2008 and a Best Paper Award from the IEEE Global Telecommunications Conference (GLOBECOM) in 2008.



Jianshu Zhang (S'09) received the B.E. degree in computer science from Sichuan University, Chengdu, China, the M.B.A. degree in global management and the M.Sc. degree in electrical engineering from Bremen University of Applied Sciences, Bremen, Germany, the M.Sc. degree in computer science and communications engineering from University of Duisburg-Essen, Duisburg, Germany, in 2003, 2004, 2006, and 2009, respectively. Since January 2010 he has been a research assistant in the Communications Research Laboratory at

Ilmenau University of Technology. His research interests include multidimensional signal processing, optimization theory as well as multi-user MIMO precoding and relaying.



Martin Haardt (S'90–M'98–SM'99) has been a Full Professor in the Department of Electrical Engineering and Information Technology and Head of the Communications Research Laboratory at Ilmenau University of Technology, Germany, since 2001. Since 2012, he has also served as an Honorary Visiting Professor in the Department of Electronics at the University of York, UK. After studying electrical engineering at the Ruhr-University Bochum, Germany, and at Purdue University, USA, he received his Diplom-Ingenieur (M.S.) degree from the

Ruhr-University Bochum in 1991 and his Doktor-Ingenieur (Ph.D.) degree from Munich University of Technology in 1996. In 1997 he joined Siemens Mobile Networks in Munich, Germany, where he was responsible for strategic research for third generation mobile radio systems. From 1998 to 2001 he was the Director for International Projects and University Cooperations in the mobile infrastructure business of Siemens in Munich, where his work focused on mobile communications beyond the third generation. During his time at Siemens, he also taught in the international Master of Science in Communications Engineering program at Munich University of Technology. Martin Haardt has received the 2009 Best Paper Award from the IEEE Signal Processing Society, the Vodafone (formerly Mannesmann Mobilfunk) Innovations-Award for outstanding research in mobile communications, the ITG best paper award from the Association of Electrical Engineering, Electronics, and Information Technology (VDE), and the Rohde & Schwarz Outstanding Dissertation Award. In the fall of 2006 and the fall of 2007 he was a visiting professor at the University of Nice in Sophia-Antipolis, France, and at the University of York, UK, respectively. His research interests include wireless communications, array signal processing, high-resolution parameter estimation, as well as numerical linear and multi-linear algebra. Prof. Haardt has served as an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING (2002–2006 and since 2011), the IEEE SIGNAL PROCESSING LETTERS (2006–2010), the *Research Letters in Signal Processing* (2007–2009), the *Hindawi Journal of Electrical and Computer Engineering* (since 2009), the *EURASIP Signal Processing Journal* (since 2011), and as a guest editor for the *EURASIP Journal on Wireless Communications and Networking*. He has also served as an elected member of the Sensor Array and Multichannel (SAM) technical committee of the IEEE Signal Processing Society (since 2011), as the technical co-chair of the IEEE International Symposia on Personal Indoor and Mobile Radio Communications (PIMRC) 2005 in Berlin, Germany, as the technical program chair of the IEEE International Symposium on Wireless Communication Systems (ISWCS) 2010 in York, UK, as the general chair of ISWCS 2013 in Ilmenau, Germany, and as the general co-chair of the 5-th IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP) 2013 in Saint Martin, French Caribbean.