Pareto Optimality for the Single-Stream Transmission in Multiuser Relay Networks

Ruixue Hu, *Student Member, IEEE*, and Tat-Ming Lok, *Senior Member, IEEE*

***Abstract*— In this paper, we study Pareto optimality for multiuser relay networks. We adopt single-stream transmission and amplify-and-forward relays. First, with fixed relay processing matrices and transmit and receive beamforming vectors, we study Pareto optimality with respect to the power of the transmitters. Based on the signal-to-noise-plus-interference ratio (SINR) balancing analysis, we give a necessary and sufficient condition for a set of SINRs to be Pareto optimal. Second, we consider Pareto optimality with respect to the relay processing matrices, where the power of the transmitters and the transmit and receive beamforming vectors is fixed. Taking advantage of multi-objective optimization analysis, we present a necessary and sufficient condition for a set of SINRs to be Pareto optimal. We also give a necessary condition to check whether Pareto optimality is fulfilled. Finally, with fixed relay processing matrices, we study Pareto optimality with respect to the transmit and receive beamforming vectors. Simulations show that our proposed algorithms outperform the compared schemes.**

***Index Terms*— Pareto optimality, relay networks, SINR balancing, beamforming.**

I. Introduction

THE dramatic growth of the requirement of data rates and the number of devices in communication networks has motivated the study of the fifth generation (5G) networks [1]. In 5G cellular networks, the communication between users can be accomplished via device to device (D2D) communication or with the help of relays when the direct D2D link can't provide satisfactory quality of service (QoS), instead of using the base stations as in traditional cellular networks [2]. Several users may be served simultaneously by a cluster of relays, resulting in a multiuser relay network. In the multiuser relay network studied in this paper, we focus on the Pareto optimality, which refers to a state of allocation of resources where it is impossible to increase the performance of one individual without decreasing that of another.

When the power of the transmitters is the only variable to be optimized, a popular approach to achieve Pareto optimality is the game-theoretic approach. The motivation of using game theory to solve power control problem is discussed in [3]. In [4] and [5], power control among selfish users is modeled as a noncooperative game where each user aims at optimizing its own utility. Besides, pricing techniques have been used to force selfish users to help improve system performance, such as revenue [6], [7] and social-welfare [8].

With fixed relay processing matrices, the system performance is affected by the number of antennas at the transmitter and the receiver [9]. If transmitters or receivers are equipped with multiple antennas, beamforming vectors can be optimized. Downlink beamforming for the multiuser multicell network is optimized by using the uplink-downlink duality in [10]. For a multiple-input and single-output (MISO) interference channel, when each receiver implements single-user detection, [11] shows that single stream transmission can achieve all points on the Pareto boundary of the rate region. Authors in [12] completely characterize the Pareto optimal transmit beamforming vectors under the assumption of single stream transmission. Besides, the rate profile approach [13] can be utilized to compute Pareto optimal beamforming vectors [14]. For Pareto optimality in multiple-input and multiple-output (MIMO) interference channels, authors in [15] characterize the Pareto optimal boundary by deducing a necessary condition, which reduces the search space. In [16] and [17], the transmit beamformers and receive beamformers are alternatively optimized towards Pareto optimality.

Fixing transmit and receive beamforming vectors, we consider the optimization of the relay processing matrices. There exist works of relay optimization for various objectives, such as minimizing the total transmit power [18], [19], minimizing the network error rate (NER) [20], satisfying the zero-forcing (ZF) or minimum-mean-square-error (MMSE) criteria [21], and so on. However, works of relay optimization for Pareto optimality are scarce. A relevant work is [22]. Therein, by utilizing the rate profile approach and semidefinite relaxation (SDR) method, the relay processing matrices are updated iteratively to achieve non-decreasing rates for all the users. When the iterative update converges, Pareto optimality would be approximately obtained. Different from [22], we now consider the relay optimization by taking advantage of SINR balancing analysis, based on which we come up with necessary and sufficient conditions for a set of SINRs to be Pareto optimal. Although the relay optimization algorithm in this paper also achieves Pareto optimality approximately,

Manuscript received October 1, 2016; revised February 21, 2017 and June 3, 2017; accepted July 3, 2017. Date of publication July 11, 2017; date of current version October 9, 2017. This work was supported by the General Research Fund from the Research Grants Council, Hong Kong Special Administrative Region, under Project CUHK 14203616. The associate editor coordinating the review of this paper and approving it for publication was G. C. Alexandropoulos. *(Corresponding author: Ruixue Hu.)*

The authors are with the Department of Information Engineering, The Chinese University of Hong Kong, Hong Kong (e-mail: hr012@ie.cuhk.edu.hk; tmlok@ie.cuhk.edu.hk).

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/TWC.2017.2724036

its performance is much better than that of [22], detailed in the simulation section. Besides, the relay optimization algorithm in this paper would result in a set of SINRs which is Pareto optimal with respect to the power of the transmitters.

In this paper, we study the Pareto optimality for a multiuser interference channel with relays. We first consider the Pareto optimality with respect to the power of the transmitters based on SINR balancing analysis [23]–[25] and come up with a necessary and sufficient condition for a set of SINRs to be Pareto optimal. Second, the optimization of relay processing matrices is taken into consideration. We optimize the corresponding SINR balancing problem by bisection search in which the feasibility problem is solved by the feasible point pursuit successive convex approximation (FPP-SCA) method [26]. Taking advantage of multi-objective optimization [27], we give a necessary and sufficient condition for a set of SINRs to be Pareto optimal with respect to the relay processing matrices. Combining the FPP-SCA method and the necessary and sufficient condition, we propose an algorithm towards Pareto optimality. The relay processing matrices are optimized through a cloud computing center connected to all relays. With the central unit, we assume a system power constraint (the same type of power constraint can be found in multiple-user multiple-relay networks [28]–[30] and single-user multiple-relay network [31]). Finally, we consider the Pareto optimality with respect to the power of the transmitters and beamforming vectors with fixed relay processing matrices. A necessary and sufficient condition is proposed and distributed update scheme of beamforming vectors is obtained.

Our main contributions are the necessary and sufficient conditions for the transmit power and the relay processing matrices to be on the Pareto optimal boundary of the SINR region, given respectively in Theorems 1 and 2. Another contribution is the relay optimization algorithm towards Pareto optimality. While there are many works [23]–[25] on SINR balancing in the literature, our goal is not just solving the SINR balancing problem. The typical SINR balancing problem considers a given set of SINRs, which may or may not be on the boundary of the SINR region. In this work, we aim at Pareto optimal SINRs. Moreover, instead of fixed SINRs, the SINRs are updated in our proposed optimization algorithms.

Notations: Vectors and matrices are written in boldface with matrices in capitals. All vectors are column vectors unless otherwise specified. For a matrix $\mathbf{A}$, we use $\mathbf{A}^T$, $\mathbf{A}^H$, $\mathrm{tr}(\mathbf{A})$, $\|\mathbf{A}\|$ and $\mathrm{vec}(\mathbf{A})$ to indicate the transpose, Hermitian transpose, trace, 2-norm and vectorization of $\mathbf{A}$ respectively. $\mathbf{I}_n$ denotes the identity matrix of size n (the subscript is omitted when unnecessary). For square matrices $\mathbf{B}_i, i = 1, 2, \ldots, n$, $\mathrm{diag}([\mathbf{B}_1, \mathbf{B}_2, \ldots, \mathbf{B}_n])$ means stacking the matrix blocks into the diagonal. $\mathbf{A} \otimes \mathbf{B}$ is the Kronecker product of $\mathbf{A}$ and $\mathbf{B}$. The inequality relation between two matrices with the same size, e.g. $\mathbf{A} \geq \mathbf{B}$, is component-wise. $\mathbf{A} \succ \mathbf{0}$ and $\mathbf{A} \succeq \mathbf{0}$ mean $\mathbf{A}$ is positive definite and positive semidefinite respectively.

II. SYSTEM MODEL

Consider the relay network with N transmitters, N receivers and K relays, where source nodes S_i $\forall i \in \mathcal{N} \triangleq \{1, 2, \ldots, N\}$ are transmitters, R_k $\forall k \in \mathcal{K} \triangleq \{1, \ldots, K\}$ are relays and destination nodes D_i $\forall i \in \mathcal{N}$ are receivers. Assume that D_i only desires messages from S_i and there is no direct link between S_i and D_i $\forall i$. The number of transmit antennas at S_i, transmit/receive antennas at R_k and receive antennas at D_j are N_{si}, N_{rk} and N_{dj} respectively. Without loss of generality, we assume $N_{si} = N_s, N_{rk} = N_r, N_{dj} = N_d, \ \forall i, k, j$. Here $\forall i, k, j$ is the short form for $\forall i \in \mathcal{N}, \forall k \in \mathcal{K}, \forall j \in \mathcal{N}$. In the following, we use $\forall i, \forall k, \forall j$ to stand for $\forall i \in \mathcal{N}, \forall k \in \mathcal{K}, \forall j \in \mathcal{N}$ respectively, unless specified otherwise.

We assume single stream transmission at each transmitter. Let x_i denote the message transmitted by S_i with $\mathbb{E}[\mathbf{x}\mathbf{x}^H] = \mathbf{I}$ and $\mathbf{x} = [x_1, x_2, \ldots, x_N]^T$. The transmit signal of S_i is $\mathbf{s}_i = \sqrt{p_i} x_i \mathbf{v}_i$ where p_i and $\mathbf{v}_i$ are the transmit power and the unit transmit beamforming vector of message x_i. The covariance matrix of $\mathbf{s}_i$ is $\boldsymbol{\Sigma}_i = p_i \mathbf{v}_i \mathbf{v}_i^H$.

Relays work in the mode of time-division duplex. They receive signals from the transmitters in the first time slot and forward signals to the receivers in the second time slot. Specifically, the received signal at R_k is

$$\mathbf{y}_{rk} = \sum_{i=1}^{N} \mathbf{F}_{ki} \mathbf{s}_i + \mathbf{n}_{rk}$$

where $\mathbf{F}_{ki} \in \mathbb{C}^{N_r \times N_s}$ is the channel matrix between S_i and R_k, and $\mathbf{n}_{rk} \sim \mathcal{CN}(0, \sigma^2 \mathbf{I}_{N_r})$ is the circularly symmetric complex Gaussian noise vector at relay R_k.

Let $\mathbf{F}_i = [\mathbf{F}_{1i}^T, \mathbf{F}_{2i}^T, \ldots, \mathbf{F}_{Ki}^T]^T$ be the channel matrix between S_i and the relay cluster, and $\mathbf{y}_R \triangleq [\mathbf{y}_{r1}^T, \mathbf{y}_{r2}^T, \ldots, \mathbf{y}_{rK}^T]^T$ the received signal of the relay cluster, which can be expressed as

$$\mathbf{y}_R = \sum_{i=1}^{N} \mathbf{F}_i \mathbf{s}_i + \mathbf{n}_R$$

where $\mathbf{n}_R = [\mathbf{n}_{r1}^T, \mathbf{n}_{r2}^T, \ldots, \mathbf{n}_{rK}^T]^T$ is the concatenation of noises and $\mathbb{E}[\mathbf{n}_R \mathbf{n}_R^H] = \sigma^2 \mathbf{I}_{N_R}$ with $N_R = K N_r$.

In the second time slot, relay R_k multiplies $\mathbf{y}_{rk}$ by $\mathbf{M}_k$, and then forwards $\mathbf{M}_k \mathbf{y}_{rk}$, where $\mathbf{M}_k \in \mathbb{C}^{N_r \times N_r}$ is the processing matrix of R_k. The overall transmitted signal of the relay cluster is $\mathbf{x}_R = \mathbf{M}\mathbf{y}_R$ with $\mathbf{M} = \mathrm{diag}([\mathbf{M}_1, \mathbf{M}_2, \ldots, \mathbf{M}_K])$. The covariance of the signal transmitted at the relay cluster is

$$\boldsymbol{\Sigma}_R = \mathbf{M}(\sum_{i=1}^{N} \mathbf{F}_i \boldsymbol{\Sigma}_i \mathbf{F}_i^H + \sigma^2 \mathbf{I}_{N_R})\mathbf{M}^H.$$

The total transmit power of the relay cluster is

$$\mathrm{tr}(\boldsymbol{\Sigma}_R) = \mathrm{tr}(\mathbf{M}(\sum_{i=1}^{N} \mathbf{F}_i \boldsymbol{\Sigma}_i \mathbf{F}_i^H + \sigma^2 \mathbf{I}_{N_R})\mathbf{M}^H). \quad (1)$$

We assume the power constraint is on the whole system, i.e., the sum of the transmit power and the relay power is upper limited by P_{sys},

$$\begin{aligned} &\sum_{i=1}^{N} \mathrm{tr}(\boldsymbol{\Sigma}_i) + \mathrm{tr}(\boldsymbol{\Sigma}_R) \\ &= \sum_{i=1}^{N} p_i + \sum_{i=1}^{N} \mathbf{v}_i^H \mathbf{F}_i^H \mathbf{M}^H \mathbf{M} \mathbf{F}_i \mathbf{v}_i p_i + \sigma^2 \mathrm{tr}(\mathbf{M}\mathbf{M}^H) \\ &\leq P_{sys}. \end{aligned} \quad (2)$$

Denote $\mathbf{G}_j = [\mathbf{G}_{j1}, \mathbf{G}_{j2}, \ldots, \mathbf{G}_{jK}]$ with $\mathbf{G}_{jk} \in \mathbb{C}^{N_d \times N_r}$ being the channel matrix between R_k and D_j, the received signal at D_j is

$$\mathbf{y}_j = \sum_{i=1}^{N} \mathbf{H}_{ji}\mathbf{s}_i + \mathbf{G}_j\mathbf{M}\mathbf{n}_R + \mathbf{n}_j \tag{3}$$

where $\mathbf{H}_{ji} = \mathbf{G}_j\mathbf{M}\mathbf{F}_i$ and $\mathbf{n}_j \sim \mathcal{CN}(0, \sigma^2\mathbf{I}_{N_d})$ is the circularly symmetric complex Gaussian noise vector at D_j. It is obvious that when $\mathbf{M}$ is fixed, the multiuser relay network can be viewed as a multiuser one-hop interference channel, where $\mathbf{H}_{ji}$ is the equivalent channel matrix from S_i to D_j and $\mathbf{z}_j = \mathbf{G}_j\mathbf{M}\mathbf{n}_R + \mathbf{n}_j$ is the equivalent noise received at D_j.

Let $\mathbf{u}_i \in \mathbb{C}^{N_d \times 1}$ be the unit decoding vector at D_i. The received SINR at D_i is

$$\text{SINR}_i = \frac{h_{ii}p_i}{\sum_{j \neq i} h_{ij}p_j + \sigma_i} \tag{4}$$

with $h_{ij} = |\mathbf{u}_i^H\mathbf{H}_{ij}\mathbf{v}_j|^2$ and $\sigma_i = \sigma^2\mathbf{u}_i^H(\mathbf{G}_i\mathbf{M}\mathbf{M}^H\mathbf{G}_i^H + \mathbf{I})\mathbf{u}_i$. We have $\sigma_i > 0\ \forall i$ since $\mathbf{G}_i\mathbf{M}\mathbf{M}^H\mathbf{G}_i^H + \mathbf{I} \succ \mathbf{0}$ and receive beamforming vector $\mathbf{u}_i$ is nonzero. We assume $h_{ii} \neq 0\ \forall i$ in the following. This assumption is based on the fact that $h_{ii} = 0$ will result in zero SINR and zero rate, which means user i doesn't work. So we can omit user i and consider the network with one fewer user. Actually, with Pareto optimality as the goal and SINR balancing analysis as the tool, optimization algorithms in our work will not return $\mathbf{u}_i$, $\mathbf{v}_i$ and $\mathbf{M}$ causing $h_{ii} = 0$, because $h_{ii} = 0$ will make the SINR balancing ratio to be zero and zero can't be the optimal value of the SINR balancing problem (e.g., (5)).

III. Pareto Optimality With Respect to the Power of the Transmitters

In this section, we give a necessary and sufficient condition for a SINR set to be Pareto optimal when the power of the transmitters is the only variable to be controlled. The theorem is based on SINR balancing analysis. So before stating the theorem, we first talk about the SINR balancing problem.

The SINR balancing problem is usually formulated to guarantee the fairness among users in the multiuser network, such as the multiuser multicell network [24], [25] and the multiuser downlink transmission in a single cell [23]. Denote $\gamma_i > 0$ as the individual target SINR at D_i, and consider the following SINR balancing problem:

$$\max_{\mathbf{U},\mathbf{V},\mathbf{M},\mathbf{p}} \min_{i \in \mathcal{N}} \frac{\text{SINR}_i}{\gamma_i} \quad \text{s.t.} \quad \boldsymbol{\omega}^T\mathbf{p} \leq P_{max} \tag{5}$$

where $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \ldots \mathbf{v}_N]$ and $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N]$ are the matrices containing all the transmit and receive beamforming vectors respectively, and $\mathbf{p} = [p_1, p_2, \ldots, p_N]^T$ is the power vector composed of the power at all transmitters. We have assumed the system power constraint (2), so we make the constraint of Problem (5) equivalent to (2) by letting $\boldsymbol{\omega} = [\omega_1, \omega_2, \ldots, \omega_N]^T$ with $\omega_i = 1 + \mathbf{v}_i^H\mathbf{F}_i^H\mathbf{M}^H\mathbf{M}\mathbf{F}_i\mathbf{v}_i$ and $P_{max} = P_{sys} - \sigma^2\text{tr}(\mathbf{M}\mathbf{M}^H)$. Notice that $\mathbf{M}$ should be chosen to satisfy $P_{sys} - \sigma^2\text{tr}(\mathbf{M}\mathbf{M}^H) > 0$.

In this section, we fix $\mathbf{U}$, $\mathbf{V}$ and $\mathbf{M}$, and deal with the optimization problem with respect to $\mathbf{p}$:

$$\max_{\mathbf{p}} \min_{i \in \mathcal{N}} \frac{\text{SINR}_i}{\gamma_i} \quad \text{s.t.} \quad \boldsymbol{\omega}^T\mathbf{p} \leq P_{max}. \tag{6}$$

Lemma 1: Denote α^* as the optimal value of Problem (6)*, then the solution has the following properties,*

$$\frac{SINR_i}{\gamma_i} = \alpha^*, \quad \forall i \tag{7}$$

$$\boldsymbol{\omega}^T\mathbf{p} = P_{max}. \tag{8}$$

The proof of this lemma is in the appendix.

Let $\mathbf{D} = \text{diag}([\gamma_1/h_{11}, \ldots, \gamma_N/h_{NN}])$, $\boldsymbol{\sigma} = [\sigma_1, \ldots, \sigma_N]^T$ and $\bar{\mathbf{H}}_I$ be an $N \times N$ matrix whose (ij)th element is h_{ij} if $i \neq j$ or 0 if $i = j$. Based on Lemma 1, the power vector $\mathbf{p}$ satisfies the following eigensystem

$$\boldsymbol{\Omega}\mathbf{p}_{ext} = \frac{1}{\alpha^*}\mathbf{p}_{ext} \tag{9}$$

where $\mathbf{p}_{ext} = [\mathbf{p}^T, 1]^T$ and

$$\boldsymbol{\Omega} = \begin{bmatrix} \mathbf{D}\bar{\mathbf{H}}_I & \mathbf{D}\boldsymbol{\sigma} \\ \frac{1}{P_{max}}\boldsymbol{\omega}^T\mathbf{D}\bar{\mathbf{H}}_I & \frac{1}{P_{max}}\boldsymbol{\omega}^T\mathbf{D}\boldsymbol{\sigma} \end{bmatrix}. \tag{10}$$

The detailed derivation of (9) is in the appendix. We notice that $\boldsymbol{\Omega}$ is a nonnegative matrix and $\frac{1}{\alpha^*}$ is a singular value of $\boldsymbol{\Omega}$. Based on the Perron-Frobenius theory (see, e.g. [32]), the spectral radius of $\boldsymbol{\Omega}$, denoted as $\rho(\boldsymbol{\Omega})$, is positive and the associated singular vector, denoted as $\mathbf{x}$, is also positive, i.e., $\mathbf{x} \geq \mathbf{0}$. On the other hand, according to Theorem 1 of [33], it is possible to choose $\mathbf{x} \geq \mathbf{0}$ and $\rho(\boldsymbol{\Omega})$ is the only singular value whose associated singular vector is positive. This indicates that there is only one solution of (9) considering that $\mathbf{p}_{ext}$ and α^* are positive. Hence, $\alpha^* = 1/\rho(\boldsymbol{\Omega})$.

Based on the SINR balancing analysis above, we have a necessary and sufficient condition for a set of SINRs to be Pareto optimal.

Theorem 1: For given $\mathbf{U}$, $\mathbf{V}$ *and* $\mathbf{M}$*, a set of SINRs* $\{\gamma_1, \ldots, \gamma_N\}$ *is Pareto optimal under the constraint* $\boldsymbol{\omega}^T\mathbf{p} \leq P_{max}$*, if and only if the optimal value of its associated SINR balancing problem*[1] *Problem* (6) *satisfies* $\alpha^* = 1$*, which is equivalent to*

$$\rho(\bar{\mathbf{H}}_I\mathbf{D} + \frac{1}{P_{max}}\boldsymbol{\sigma}\boldsymbol{\omega}^T\mathbf{D}) = 1. \tag{11}$$

The proof is in the appendix.

IV. Pareto Optimality With Respect to the Relay Processing Matrices

Based on the Pareto optimality analysis with respect to the power, we now take Pareto optimality with respect to the relay processing matrices into consideration. In this section, we optimize the relay processing matrices with fixed transmit power, transmit beamforming vectors and receive beamforming vectors.

[1]For a set of SINRs, its associated SINR balancing problem refers to the SINR balancing problem with this set of SINRs as the SINR requirement.

For given $\mathbf{p}$, $\mathbf{U}$ and $\mathbf{V}$, the SINR balancing Problem (5) becomes

$$\max_{\mathbf{M}} \min_{i\in\mathcal{N}} \frac{\text{SINR}_i}{\gamma_i} \quad \text{s.t.} \quad \boldsymbol{\omega}^T\mathbf{p} \le P_{max} \tag{12}$$

where the constraint is also affected by $\mathbf{M}$ since $\omega_i = 1 + \mathbf{v}_i^H\mathbf{F}_i^H\mathbf{M}^H\mathbf{M}\mathbf{F}_i\mathbf{v}_i$.

Lemma 2: Denote α^* as the optimal value associated with the optimal solution $\mathbf{M}^*$ of Problem (12) and $\boldsymbol{\omega}^*$ as the weight calculated with $\mathbf{M}^*$, i.e., $\omega_i^* = 1+\mathbf{v}_i^H\mathbf{F}_i^H(\mathbf{M}^*)^H\mathbf{M}^*\mathbf{F}_i\mathbf{v}_i$. Then the constraint holds with equality when (12) is solved, i.e.,

$$(\boldsymbol{\omega}^*)^T\mathbf{p} = P_{max}. \tag{13}$$

The proof of Lemma 2 is in the appendix. Based on Lemma 2 and the fact that $\frac{\text{SINR}_i}{\gamma_i} > 0\ \forall i$, we rewrite Problem (12) as

$$\min_{\mathbf{M}} \max_{i\in\mathcal{N}} \frac{\gamma_i}{\text{SINR}_i} \quad \text{s.t.} \quad \boldsymbol{\omega}^T\mathbf{p} = P_{max}. \tag{14}$$

In the following, we first reshape Problem (14) and then talk about the method to solve it.

A. Problem Transformation

Following the same steps in Section IV.A of [22], we reformulate the SINR expression (4) as

$$\text{SINR}_i = \frac{p_i\mathbf{m}^H\mathbf{h}_{ii}\mathbf{h}_{ii}^H\mathbf{m}}{\sum_{j\ne i} p_j\mathbf{m}^H\mathbf{h}_{ij}\mathbf{h}_{ij}^H\mathbf{m} + \mathbf{m}^H\mathbf{E}_i\mathbf{m} + \sigma^2}$$

where $\mathbf{m} = [\text{vec}(\mathbf{M}_1)^T, \ldots, \text{vec}(\mathbf{M}_K)^T]^T$, $\mathbf{h}_{ij} = [(\mathbf{v}_j^T\mathbf{F}_{1j}^T)\otimes(\mathbf{u}_i^H\mathbf{G}_{i1}), \ldots, (\mathbf{v}_j^T\mathbf{F}_{Kj}^T)\otimes(\mathbf{u}_i^H\mathbf{G}_{iK})]^H$ and $\mathbf{E}_i = \sigma^2\text{diag}([\mathbf{I}\otimes(\mathbf{G}_{i1}^H\mathbf{u}_i\mathbf{u}_i^H\mathbf{G}_{i1}), \ldots, \mathbf{I}\otimes(\mathbf{G}_{iK}^H\mathbf{u}_i\mathbf{u}_i^H\mathbf{G}_{iK})])$.

Similarly, we rewrite the constraint $\boldsymbol{\omega}^T\mathbf{p} = P_{max}$ as $\mathbf{m}^H\mathbf{Q}_{sum}\mathbf{m} = 1$ with $\mathbf{Q}_i = \text{diag}([(\mathbf{F}_{1i}\mathbf{v}_i\mathbf{v}_i^H\mathbf{F}_{1i}^H)^T\otimes\mathbf{I}, \ldots, (\mathbf{F}_{Ki}\mathbf{v}_i\mathbf{v}_i^H\mathbf{F}_{Ki}^H)^T\otimes\mathbf{I}])$ and $\mathbf{Q}_{sum} = \frac{1}{P_{max}-\sum_{i=1}^N p_i}(\sum_{i=1}^N p_i\mathbf{Q}_i + \sigma^2\mathbf{I})$. Notice that the power of transmitters should be chosen such that $P_{max} - \sum_{i=1}^N p_i > 0$. The objective of Problem (14) is equivalent to

$$\min_{\mathbf{m}} \max_{i\in\mathcal{N}} \frac{\gamma_i}{p_i}\frac{\sum_{j\ne i} p_j\mathbf{m}^H\mathbf{h}_{ij}\mathbf{h}_{ij}^H\mathbf{m} + \mathbf{m}^H\mathbf{E}_i\mathbf{m} + \sigma^2}{\mathbf{m}^H\mathbf{h}_{ii}\mathbf{h}_{ii}^H\mathbf{m}}. \tag{15}$$

Let $\mathbf{y} = \mathbf{Q}_{sum}^{\frac{1}{2}}\mathbf{m}$, the optimization problem becomes

$$\min_{\mathbf{y}} \max_{i\in\mathcal{N}} \frac{\gamma_i\mathbf{y}^H\mathbf{A}_i\mathbf{y}}{\mathbf{y}^H\mathbf{C}_i\mathbf{y}} \quad \text{s.t.} \quad \mathbf{y}^H\mathbf{y} = 1 \tag{16}$$

where $\mathbf{A}_i = \frac{1}{\sigma^2}\mathbf{Q}_{sum}^{-\frac{1}{2}}(\sum_{j\ne i} p_j\mathbf{h}_{ij}\mathbf{h}_{ij}^H + \mathbf{E}_i)\mathbf{Q}_{sum}^{-\frac{1}{2}} + \mathbf{I}$ and $\mathbf{C}_i = \frac{p_i}{\sigma^2}\mathbf{Q}_{sum}^{-\frac{1}{2}}\mathbf{h}_{ii}\mathbf{h}_{ii}^H\mathbf{Q}_{sum}^{-\frac{1}{2}}$.

Notice that for any $\mathbf{y} \ne \mathbf{0}$, we can always scale it to satisfy the norm-one constraint without affecting the value of objective function. So we replace the norm-one constraint with $\mathbf{y} \ne \mathbf{0}$ when solving Problem (16).

We introduce a new variable t. Problem (16) is equivalent to

$$\min_{\mathbf{y}\ne\mathbf{0},t} t \quad \text{s.t.} \quad \frac{\gamma_i\mathbf{y}^H\mathbf{A}_i\mathbf{y}}{\mathbf{y}^H\mathbf{C}_i\mathbf{y}} \le t \quad \forall i. \tag{17}$$

We utilize the bisection search method to find the optimal solution of Problem (17). Let $[a, b]$ be the search interval containing the optimal value t^*. If t_0 is a feasible solution, then the following problem is feasible:

$$\text{find}_{\mathbf{y}\ne\mathbf{0}}\ \mathbf{y} \quad \text{s.t.} \quad \frac{\gamma_i\mathbf{y}^H\mathbf{A}_i\mathbf{y}}{\mathbf{y}^H\mathbf{C}_i\mathbf{y}} \le t_0 \quad \forall i. \tag{18}$$

For any $\mathbf{y} \ne 0$, we have $\mathbf{y}^H\mathbf{A}_i\mathbf{y} > 0$ and $\mathbf{y}^H\mathbf{C}_i\mathbf{y} \ge 0$ since $\mathbf{A}_i \succ \mathbf{0}$ and $\mathbf{C}_i \succeq \mathbf{0}$. Zero denominator fails the inequality constraints of Problem (18), so $\mathbf{y}^H\mathbf{C}_i\mathbf{y} \ne 0$, namely $\mathbf{y}^H\mathbf{C}_i\mathbf{y} > 0\ \forall i$. Multiply both sides of the constraint by $\mathbf{y}^H\mathbf{C}_i\mathbf{y}$ and rearrange the terms, we have an equivalent form of Problem (18), i.e.,

$$\text{find}_{\mathbf{y}\ne\mathbf{0}}\ \mathbf{y} \quad \text{s.t.} \quad \mathbf{y}^H(\mathbf{A}_i - \frac{t_0}{\gamma_i}\mathbf{C}_i)\mathbf{y} \le 0 \quad \forall i \tag{19}$$

where the constraints $\mathbf{y}^H\mathbf{C}_i\mathbf{y} \ne 0\ \forall i$ are omitted. This is because given $\mathbf{y} \ne \mathbf{0}$, if $\mathbf{y}^H\mathbf{C}_i\mathbf{y} = 0$ for some i, then $\mathbf{y}^H\mathbf{A}_i\mathbf{y} \le 0$ which contradicts to $\mathbf{A}_i \succ \mathbf{0}$. Problem (19) is a quadratically constrained feasibility problem, which is NP-hard in general.

B. FPP-SCA Method to Solve Problem (19)

Given that finding a feasible solution to Problem (19) is NP-hard, we introduce FPP-SCA, which combines the convex approximation with utilization of slack penalty [26]. This method can find a feasible solution to Problem (19) with high probability without knowing an initial feasible point when it is feasible.

Problem (19) has N inequality constraints, thus we need N slack variables, $s_i\ \forall i \in \mathcal{N}$. Using the slack penalty as the objective function, we have the relaxed problem of Problem (19), i.e.,

$$\min_{\mathbf{y}\ne\mathbf{0},\{s_i\}} \sum_{i=1}^N s_i \quad \text{s.t.} \quad \mathbf{y}^H(\mathbf{A}_i - \frac{t_0}{\gamma_i}\mathbf{C}_i)\mathbf{y} - s_i \le 0 \quad \forall i, \quad s_i \ge 0 \quad \forall i. \tag{20}$$

The penalty $\sum_{i=1}^N s_i$ is to push the slack variables towards zero. We can also use other suitable penalty functions.

For a positive semidefinite matrix $\mathbf{C}$, we have $(\mathbf{x} - \mathbf{z})^H\mathbf{C}(\mathbf{x} - \mathbf{z}) \ge 0$ for any $\mathbf{x}$ and $\mathbf{z}$. Expanding the Left Hand Side (LHS) results in $\mathbf{x}^H\mathbf{C}\mathbf{x} - \mathbf{z}^H\mathbf{C}\mathbf{x} - \mathbf{x}^H\mathbf{C}\mathbf{z} + \mathbf{z}^H\mathbf{C}\mathbf{z} \ge 0$. Given any complex number a, the real part $\text{Re}\{a\} = (a + a^H)/2$. Hence we have the following linear restriction:

$$\mathbf{x}^H\mathbf{C}\mathbf{x} \ge 2\text{Re}\{\mathbf{z}^H\mathbf{C}\mathbf{x}\} - \mathbf{z}^H\mathbf{C}\mathbf{z}. \tag{21}$$

Algorithm 1 FPP-SCA Method to Find a Feasible Solution of Problem (19)

Set the tolerance criterion ϵ and the iteration number $n = 0$;
Generate a random point $\mathbf{z}^{(1)} \in \mathbb{C}^{KN_k^2 \times 1}$;
do
  $n = n + 1$, $\mathbf{z} = \mathbf{z}^{(n)}$;
  Solve the following approximation problem:

$$\min_{\mathbf{y} \neq \mathbf{0}, \{s_i\}} \sum_{i=1}^{N} s_i$$
$$\text{s.t. } \mathbf{y}^H \mathbf{A}_i \mathbf{y} - 2\frac{t_0}{\gamma_i}\mathrm{Re}\{\mathbf{z}^H \mathbf{C}_i \mathbf{y}\} \leq s_i - \frac{t_0}{\gamma_i}\mathbf{z}^H \mathbf{C}_i \mathbf{z} \quad \forall i$$
$$s_i \geq 0 \quad \forall i; \tag{22}$$

  Denote $\mathbf{y}^*$ as the optimal solution and set $\mathbf{z}^{(n+1)} = \mathbf{y}^*$;
while $|\mathbf{z}^{(n)} - \mathbf{z}^{(n+1)}| \geq \epsilon$
if $(\mathbf{y}^*)^H(\mathbf{A}_i - \frac{t_0}{\gamma_i}\mathbf{C}_i)\mathbf{y}^* \leq 0 \ \forall i$ **then**
  Problem (19) has a feasible solution $\mathbf{y}^*$;
else
  Conclude that Problem (19) is not feasible.
end if

Based on (21) and the fact that $\mathbf{C}_i \ \forall i$ are positive semidefinite, we have Algorithm 1 to solve Problem (19).

Setting $s_i = 0 \ \forall i$ in Problem (22), we have

$$\text{find}_{\mathbf{y} \neq \mathbf{0}} \ \mathbf{y}$$
$$\text{s.t. } \mathbf{y}^H \mathbf{A}_i \mathbf{y} - 2\frac{t_0}{\gamma_i}\mathrm{Re}\{\mathbf{z}^H \mathbf{C}_i \mathbf{y}\} \leq -\frac{t_0}{\gamma_i}\mathbf{z}^H \mathbf{C}_i \mathbf{z} \quad \forall i \tag{23}$$

which is the result of applying (21) to the terms containing $\mathbf{C}_i$ in Problem (19).

The meaning of linear approximation (21) lies in turning Problem (19) from NP-hard to convex. The application of slack variables is meaningful in two aspects. First, applying (21) to the terms containing $\mathbf{C}_i$ tightens the inequality constraint of Problem (19). Thus Problem (23) may be infeasible although Problem (19) is feasible. Adding the slack variables can avoid this risk since Problem (22) is always feasible. Second, solving Problem (23) by the SCA approach requires a feasible initial point which is NP-hard to find. Adding slack variables relieves this requirement and results in an easier start for the algorithm.

To run Algorithm 1, matrices $\mathbf{A}_i$ and $\mathbf{C}_i \ \forall i$ need to be calculated. Each relay needs to know all the channel matrices of the links connected with it, i.e., the channel matrices from all the transmitters to this relay and the channel matrices from this relay to all the receivers. Relays also need to know the transmit power and the beamforming vectors of all the transmitters as well as all the receive beamforming vectors. Then they send all the information to the cloud computing center connected to the relays. Algorithm 1 is convergent since it results in nonincreasing slack penalty and the penalty has zero as its lower bound. To be specific, we denote the optimal slack variables of the nth iteration as $s_i^{(n)} \ \forall i$. Since $\mathbf{z}^{(n+1)} = \mathbf{y}^*$ is the optimal solution, it satisfies

$$(\mathbf{z}^{(n+1)})^H \mathbf{A}_i \mathbf{z}^{(n+1)} - 2\frac{t_0}{\gamma_i}\mathrm{Re}\{\mathbf{z}^H \mathbf{C}_i \mathbf{z}^{(n+1)}\} + \frac{t_0}{\gamma_i}\mathbf{z}^H \mathbf{C}_i \mathbf{z} \leq s_i^{(n)}, \quad \forall i \tag{24}$$

with $\mathbf{z} = \mathbf{z}^{(n)}$. Using (21), we have

$$(\mathbf{z}^{(n+1)})^H \mathbf{A}_i \mathbf{z}^{(n+1)} - \frac{t_0}{\gamma_i}(\mathbf{z}^{(n+1)})^H \mathbf{C}_i \mathbf{z}^{(n+1)} \leq s_i^{(n)}, \quad \forall i. \tag{25}$$

At the $(n+1)$th iteration, we set $\mathbf{z} = \mathbf{z}^{(n+1)}$. With the inequality (25), it is easy to verify that $\mathbf{y} = \mathbf{z}^{(n+1)}$ and $s_i = s_i^{(n)} \ \forall i$ satisfy the constraints. This is to say, the optimal solution of the nth iteration is a feasible solution of the $(n+1)$th iteration. If we denote the optimal slack variables of the $(n+1)$th solution as $s_i^{(n+1)} \ \forall i$, then $\sum_i s_i^{(n+1)} \leq \sum_i s_i^{(n)}$.

C. Pareto Optimality Analysis Based on Multi-Objective Optimization

FPP-SCA can find a solution of Problem (19) with high probability when it is feasible. In this subsection, we try to improve the probability by taking advantage of multi-objective optimization. Denote t^* as the optimal value of Problem (17). According to Theorem 1, if a set of SINRs is Pareto optimal, then the optimal value of the associated SINR balancing problem (17) has to be $t^* = 1$. Aiming at Pareto optimality, we study Problem (19) with $t_0 = 1$ in this subsection. In other words, when mentioning Problem (19) in this subsection, we refer to Problem (19) with $t_0 = 1$, unless specified otherwise.

A set of SINRs $\{\gamma_1, \ldots, \gamma_N\}$ is achievable if and only if its associated feasibility problem (19) is feasible. Denote $f_i(\mathbf{y}) = \mathbf{y}^H(\mathbf{A}_i - \frac{1}{\gamma_i}\mathbf{C}_i)\mathbf{y}$. Problem (19) is equivalent to find a nonzero vector $\mathbf{y}$ such that $f_i(\mathbf{y}) \leq 0 \ \forall i$.

The definition of Pareto optimality is that there is no way to increase one SINR, say γ_i, without deceasing another SINR γ_j, $j \neq i$. Therefore, if $\boldsymbol{\gamma} = [\gamma_1, \ldots, \gamma_N]^T$ is Pareto optimal, any $\boldsymbol{\gamma}' \geq \boldsymbol{\gamma}$ with $\boldsymbol{\gamma}' \neq \boldsymbol{\gamma}$ is not achievable, i.e., the feasibility problem associated with $\boldsymbol{\gamma}'$ is infeasible. Based on this idea, we have the following theorem on Pareto optimality.

Theorem 2: For given $\mathbf{U}$, $\mathbf{V}$ *and* $\mathbf{p}$, *a set of SINRs* $\{\gamma_1, \ldots, \gamma_N\}$ *is Pareto optimal under the constraint* $\boldsymbol{\omega}^T \mathbf{p} \leq P_{max}$, *if and only if the solution of its associated feasibility problem* (19) *with* $t_0 = 1$ *satisfies that*

i) $\mathbf{y}$ *makes all the constraints active, i.e.,* $f_i(\mathbf{y}) = 0 \ \forall i$, *and*

ii) there does not exist a $\mathbf{y}'$ *such that* $\mathbf{f}(\mathbf{y}') \leq \mathbf{f}(\mathbf{y})$ *and* $\mathbf{f}(\mathbf{y}') \neq \mathbf{f}(\mathbf{y})$.

The proof is in the appendix.

The second condition in Theorem 2 is equivalent to that $\mathbf{y}$ achieves Pareto optimality of $\mathbf{f}(\mathbf{y})$. The global Pareto optimality is NP-hard to find since $f_i(\mathbf{y}) \ \forall i$ are nonconvex. Instead, we try to find a locally Pareto optimal solution $\mathbf{y}$, i.e., within a neighborhood of $\mathbf{y}$ there does not exist a point $\mathbf{y}' \neq \mathbf{y}$ with $\mathbf{f}(\mathbf{y}') \leq \mathbf{f}(\mathbf{y})$ and $\mathbf{f}(\mathbf{y}') \neq \mathbf{f}(\mathbf{y})$.

Denote the vector $\mathbf{r} \in \mathbb{R}^{2KN_k^2 \times 1}$ as $\mathbf{r} = (\mathrm{Re}\{\mathbf{y}\}, \mathrm{Im}\{\mathbf{y}\})$. Hence

$$y_j = r_j + \imath r_{j+KN_k^2}, \quad j = 1, \ldots, KN_k^2. \tag{26}$$

Here $\imath$ is the imaginary unit. The function vector $\mathbf{f}(\mathbf{y}) = \mathbf{f}(\text{Re}\{\mathbf{y}\} + \text{Im}\{\mathbf{y}\}\imath)$, and we use $\mathbf{f}(\mathbf{r})$ to stand for $\mathbf{f}(\text{Re}\{\mathbf{y}\} + \text{Im}\{\mathbf{y}\}\imath)$ in the following. Functions $f_i(\mathbf{r})\ \forall i$ are continuously differentiable. The Jacobian of $\mathbf{f}$ at $\mathbf{r}$ is denoted by $\mathbf{Jf}(\mathbf{r})$, an $N \times 2KN_k^2$ matrix with entries

$$(\mathbf{Jf}(\mathbf{r}))_{ij} = \frac{\partial f_i}{\partial r_j}(\mathbf{r}).$$

A necessary condition for $\mathbf{r}$ to be locally Pareto optimal is

$$\text{span}(\mathbf{Jf}(\mathbf{r})) \cap (-\mathbb{R}_{++})^N = \varnothing \tag{27}$$

where $\text{span}(\mathbf{Jf}(\mathbf{r}))$ denotes the column space spanned by matrix $\mathbf{Jf}(\mathbf{r})$ (see e.g. [27]).

If $\mathbf{r}$ does not satisfy (27), there exists a direction $\mathbf{d} \in \mathbb{C}$ satisfying $\mathbf{Jf}(\mathbf{r})\mathbf{d} \in (-\mathbb{R}_{++})^N$. A reasonable choice is to minimizing the largest change across the set of objectives, i.e., to reduce the objectives as much as possible [34].

$$\begin{aligned} &\min_{\beta,\mathbf{d}}\ \beta + \frac{1}{2}\|\mathbf{d}\|^2 \\ &\text{s.t. } \mathbf{Jf}(\mathbf{r})\mathbf{d} \le \beta\mathbf{1}. \end{aligned} \tag{28}$$

After finding a descent direction, we can apply the steepest descent method and update $\mathbf{r}$ as

$$\mathbf{r} = \mathbf{r}_c + \text{step} * \mathbf{d} \tag{29}$$

where $\mathbf{r}_c$ is the current solution and step is an appropriately chosen stepsize.

D. Algorithm Towards Pareto Optimality

In this subsection, we combine the FPP-SCA method and the multi-objective analysis, and come up with the algorithm for relay processing matrices towards the local Pareto optimality.

In Algorithm 2, Step 3-7 utilize the bisection search method and during each search the feasibility problem is solved by FPP-SCA method. Step 17-28 are based on multi-objective analysis. To run Algorithm 2, each relay needs to know all the channel matrices of the links connected with it, i.e., the channel matrices from all the transmitters to this relay and the channel matrices from this relay to all the receivers. Relays also need to know the transmit power and the beamforming vectors of all the transmitters as well as all the receive beamforming vectors. Then they send all the information to the computing center connected to the relays, which then runs Algorithm 2 and obtains the relay processing matrices towards Pareto optimality.

This algorithm results in a relay vector and a set of SINRs. The resultant SINR set $\{\gamma_1^*, \ldots, \gamma_N^*\}$ is approximately locally Pareto optimal with respect to the relay processing matrices, because (27) is a necessary condition for local Pareto optimality. If we can replace it with a necessary and sufficient condition, then Pareto optimality can be achieved.

Although the resultant SINR set achieves the local Pareto optimality with respect to the relay processing matrices approximately, it satisfies the necessary and sufficient condition stated in Theorem 1. This is to say, the update of relay processing matrices by Algorithm 2 can always keep the set of SINRs on the Pareto optimal boundary. The details are in the following theorem.

Algorithm 2 Relay Optimization Towards Pareto Optimality

1: Set the tolerance criterion ϵ;
2: Set the bisection search interval [a,b];
3: Initialize a set of SINRs $\{\gamma_1, .., \gamma_N\}$;
4: **while** $|a - b| \ge \epsilon$ **do**
5:     Set $t_0 = \frac{a+b}{2}$;
6:     Solve Problem (19) by Algorithm 1;
7:     **if** Problem (19) is feasible with solution $\mathbf{y}$ **then**
8:         set $b = t_0$;
9:     **else** Problem (19) is infeasible
10:         set $a = t_0$;
11:     **end if**
12: **end while**
13: **if** $t > 1$ **then**
14:     Conclude that the initial set of SINRs is unachievable;
15: **end if**
16: Set the current SINR of user i as $\frac{1}{t}\gamma_i$;
17: **for** $i = 1, \ldots, N$ **do**
18:     **if** $f_i(\mathbf{y}) < 0$ **then**
19:         Adapt γ_i such that $f_i(\mathbf{y}) = 0$;
20:     **end if**
21: **end for**
22: Set $\mathbf{r} = (\text{Re}\{\mathbf{y}\}, \text{Im}\{\mathbf{y}\})$ and solve Problem (28);
23: **if** Problem (28) is infeasible **then**
24:     Conclude the current set of SINRs is Pareto optimal;
25: **else** Problem (28) is feasible with solution $\mathbf{d}$
26:     Update $\mathbf{r}$ by (29) until Problem (28) is infeasible;
27:     Set $\mathbf{y}$ by (26) and go to Step 17;
28: **end if**
29: Normalize $\mathbf{y}$ and return $\mathbf{m} = \mathbf{Q}_{sum}^{-\frac{1}{2}}\mathbf{y}$ as the relay vector.

Theorem 3: With given transmit beamforming vectors and receive beamforming vectors, the relay processing matrices and the SINR set resulting from Algorithm 2 satisfy (11). *The resultant SINR set is on the Pareto optimal boundary of the SINR region with respect to the power.*

The proof is in the appendix.

In Step 22-28 of Algorithm 2, we utilize (27) to check whether the set achieves Pareto optimality. If not, we update $\mathbf{r}$ by (28) and (29). However, such an update method usually converges very slowly. So we would rather use (27) as a criterion to check whether the Pareto optimality is satisfied than as a computing method in practice. Even without Step 22-28, the resultant SINR set is also on the Pareto optimal boundary since $f_i = 0, \forall\ i$ hold. Besides, given a set of SINRs, if the purpose is to check whether it is feasible, we can stop at Step 15.

V. Pareto Optimality With Respect to the Receive and Transmit Beamforming Vectors

In this section, we fix the relay processing matrices and consider Pareto optimality with respect to the receive and transmit beamforming vectors. We study the following optimization problem:

$$\max_{\mathbf{U},\mathbf{V},\mathbf{p}} \min_{i\in\mathcal{N}} \frac{\text{SINR}_i}{\gamma_i} \quad \text{s.t.} \quad \boldsymbol{\omega}^T\mathbf{p} \le P_{max}. \tag{30}$$

Based on the analysis in Section III, the optimal value of the SINR balancing problem is $\alpha^* = 1/\rho(\boldsymbol{\Omega})$ for a certain $\mathbf{U}$ and $\mathbf{V}$. Hence we denote the optimal value of Problem (30) as $\alpha^*(\mathbf{U},\mathbf{V})$.

A set of SINR $\{\gamma_1,\ldots,\gamma_N\}$ is Pareto optimal if and only if $\max_{\mathbf{U},\mathbf{V}} \alpha^*(\mathbf{U},\mathbf{V}) = 1$. In order to achieve the Pareto optimality, we first need to find the method to solve $\max_{\mathbf{U},\mathbf{V}} \alpha^*(\mathbf{U},\mathbf{V})$ or equivalently $\min_{\mathbf{U},\mathbf{V}} \rho(\boldsymbol{\Omega})$. Here $\boldsymbol{\Omega}(\mathbf{U},\mathbf{V})$ is a function of $\mathbf{U}$ and $\mathbf{V}$ and calculated as (10).

A. Optimization of Receive Beamforming Vectors

Recall that $\rho(\boldsymbol{\Omega}) = \rho\big(\mathbf{D}(\bar{\mathbf{H}}_I + \frac{1}{P_{max}}\boldsymbol{\sigma}\boldsymbol{\omega}^T)\big)$ from Section III. Denote $\mathbf{B} = \mathbf{D}(\bar{\mathbf{H}}_I + \frac{1}{P_{max}}\boldsymbol{\sigma}\boldsymbol{\omega}^T)$, so $\rho(\boldsymbol{\Omega}) = \rho(\mathbf{B})$ and the (i,j)th element of $\mathbf{B}$ is $\frac{\gamma_i}{h_{ii}}(h_{ij} + \frac{\sigma_i\omega_j}{P_{max}})$ if $i \neq j$ and $\frac{\gamma_i}{h_{ii}}\frac{\sigma_i\omega_i}{P_{max}}$ if $i = j$. The Collatz-Wielandt formula [32] asserts that for any $N \times N$ nonnegative matrix $\mathbf{A}$,

$$\rho(\mathbf{A}) = \min_{\mathbf{x}>0} \max_{i\in\mathcal{N}} \frac{(\mathbf{A}\mathbf{x})_i}{x_i}. \tag{31}$$

Fixing the transmit vectors $\mathbf{V}$ and only considering the optimization of $\mathbf{U}$, we have

$$\begin{aligned}\min_{\mathbf{U}} \rho(\boldsymbol{\Omega}) &= \min_{\mathbf{x}>0,\mathbf{U}} \max_{i\in\mathcal{N}} \frac{(\mathbf{B}\mathbf{x})_i}{x_i} \\ &= \min_{\mathbf{x}>0,\mathbf{U}} \max_{i\in\mathcal{N}} \frac{\gamma_i}{h_{ii}x_i}\Big(\sum_{j\neq i} h_{ij}x_j + \frac{\sigma_i}{P_{max}}\sum_{j\in\mathcal{N}}\omega_j x_j\Big) \\ &\overset{(a)}{=} \min_{\mathbf{x}>0,\mathbf{U}} \max_{i\in\mathcal{N}} \frac{\gamma_i}{h_{ii}x_i}\Big(\sum_{j\neq i} h_{ij}x_j + \sigma_i\Big) \\ &= \min_{\mathbf{x}>0,\mathbf{U}} \max_{i\in\mathcal{N}} \frac{\gamma_i\mathbf{u}_i^H\mathbf{J}_i\mathbf{u}_i}{x_i\mathbf{u}_i^H\mathbf{H}_{ii}\mathbf{v}_i\mathbf{v}_i^H\mathbf{H}_{ii}^H\mathbf{u}_i}\end{aligned} \tag{32}$$

where $\mathbf{J}_i = \sum_{j\neq i} x_j\mathbf{H}_{ij}\mathbf{v}_j\mathbf{v}_j^H\mathbf{H}_{ij}^H + \sigma^2(\mathbf{G}_i\mathbf{M}\mathbf{M}^H\mathbf{G}_i^H + \mathbf{I})$ and (a) is because we can always set $\sum_{j\in\mathcal{N}}\omega_j x_j = P_{max}$ since scaling $\mathbf{x}$ does not affect the result according to (31).

Optimizing $\mathbf{x}$ and $\mathbf{U}$ simultaneously is very difficult, so we optimize them iteratively. For a certain $\mathbf{x} > 0$ and a certain i,

$$\begin{aligned}&\min_{\mathbf{U}} \frac{\gamma_i\mathbf{u}_i^H\mathbf{J}_i\mathbf{u}_i}{x_i\mathbf{u}_i^H\mathbf{H}_{ii}\mathbf{v}_i\mathbf{v}_i^H\mathbf{H}_{ii}^H\mathbf{u}_i} \\ &= 1/\Big(\max_{\mathbf{u}_i} \frac{x_i\mathbf{u}_i^H\mathbf{H}_{ii}\mathbf{v}_i\mathbf{v}_i^H\mathbf{H}_{ii}^H\mathbf{u}_i}{\gamma_i\mathbf{u}_i^H\mathbf{J}_i\mathbf{u}_i}\Big) \\ &\overset{(b)}{=} \frac{\gamma_i}{x_i\mathbf{v}_i^H\mathbf{H}_{ii}^H\mathbf{J}_i^{-1}\mathbf{H}_{ii}\mathbf{v}_i}\end{aligned} \tag{33}$$

where (b) is because the optimal solution is

$$\mathbf{u}_i = \frac{\mathbf{J}_i^{-1}\mathbf{H}_{ii}\mathbf{v}_i}{\|\mathbf{J}_i^{-1}\mathbf{H}_{ii}\mathbf{v}_i\|}. \tag{34}$$

The solution $\mathbf{u}_i$ is a direct result from the fact that for any vector $\mathbf{a}$ and matrix $\mathbf{A} \succ \mathbf{0}$,

$$\frac{\mathbf{x}^H\mathbf{a}\mathbf{a}^H\mathbf{x}}{\mathbf{x}^H\mathbf{A}\mathbf{x}} \le \mathbf{a}^H\mathbf{A}^{-1}\mathbf{a} \tag{35}$$

where the equality holds if and only if $\mathbf{x} = c\mathbf{A}^{-1}\mathbf{a}$ with c being a nonzero real scalar.

For a certain $\mathbf{U}$, Problem (32) is to find a $\mathbf{x}$ such that

$$\rho(\boldsymbol{\Omega}) = \min_{\mathbf{x}>0} \max_{i\in\mathcal{N}} \frac{(\mathbf{B}\mathbf{x})_i}{x_i}. \tag{36}$$

This is fulfilled by the singular vector associated with $\rho(\mathbf{B}) = \rho(\boldsymbol{\Omega})$. In detail, if $\boldsymbol{\Omega}\mathbf{p}_{ext} = \rho(\boldsymbol{\Omega})\mathbf{p}_{ext}$ with the last element $\mathbf{p}_{ext}$ scaled to 1, we stack the first N elements in $\mathbf{p}$. Then it is easy to check that $\mathbf{B}\mathbf{p} = \rho(\mathbf{B})\mathbf{p}$. Therefore $\mathbf{p}$ is the optimal solution of Problem (32) when $\mathbf{U}$ is fixed.

To update the receive beamforming vector, each receiver needs to know the transmit beamforming vectors of all the transmitters, the relay processing matrices, all the channel matrices from the transmitters to the relays and all the channel matrices from the relays to this receiver.

B. Optimization of Transmit Beamforming Vectors

To optimize the transmit beamforming vectors, we first reformulate $\rho(\boldsymbol{\Omega})$. Recall that $\rho(\mathbf{A}) = \rho(\mathbf{A}^T)$ and $\rho(\mathbf{A}\mathbf{B}) = \rho(\mathbf{B}\mathbf{A})$ for any matrices $\mathbf{A}$ and $\mathbf{B}$, we have

$$\rho(\boldsymbol{\Omega}) = \rho\Big(\mathbf{D}(\bar{\mathbf{H}}_I^T + \frac{1}{P_{max}}\boldsymbol{\omega}\boldsymbol{\sigma}^T)\Big).$$

Denote $\tilde{\mathbf{B}} = \mathbf{D}(\bar{\mathbf{H}}_I^T + \frac{1}{P_{max}}\boldsymbol{\omega}\boldsymbol{\sigma}^T)$. The ijth element of $\tilde{\mathbf{B}}$ is $\frac{\gamma_i}{h_{ii}}(h_{ji} + \frac{\sigma_j\omega_i}{P_{max}})$ if $i \neq j$ and $\frac{\gamma_i}{h_{ii}}\frac{\sigma_i\omega_i}{P_{max}}$ if $i = j$. Fixing $\mathbf{U}$, we optimize $\mathbf{V}$ and $\tilde{\mathbf{x}}$ iteratively. For a certain $\tilde{\mathbf{x}}$ and a certain i, utilizing (31), we have

$$\begin{aligned}\min_{\mathbf{V}} \rho(\boldsymbol{\Omega}) &= \min_{\mathbf{V}} \frac{(\tilde{\mathbf{B}}\tilde{\mathbf{x}})_i}{\tilde{x}_i} \\ &= \min_{\mathbf{V}} \frac{\gamma_i}{h_{ii}\tilde{x}_i}\Big(\sum_{j\neq i} h_{ji}\tilde{x}_j + \frac{\omega_i}{P_{max}}\sum_{j\in\mathcal{N}}\sigma_j\tilde{x}_j\Big) \\ &\overset{(c)}{=} \min_{\mathbf{V}} \frac{\gamma_i}{h_{ii}\tilde{x}_i}\Big(\sum_{j\neq i} h_{ji}\tilde{x}_j + \omega_i\Big) \\ &\overset{(d)}{=} \frac{\gamma_i}{\tilde{x}_i\mathbf{u}_i^H\mathbf{H}_{ii}\tilde{\mathbf{J}}_i^{-1}\mathbf{H}_{ii}^H\mathbf{u}_i}\end{aligned} \tag{37}$$

where $\tilde{\mathbf{J}}_i = \sum_{j\neq i}\tilde{x}_j\mathbf{H}_{ji}^H\mathbf{u}_j\mathbf{u}_j^H\mathbf{H}_{ji} + \mathbf{F}_i^H\mathbf{M}^H\mathbf{M}\mathbf{F}_i + \mathbf{I}$. (c) is because we can set $\sum_{j\in\mathcal{N}}\sigma_j\tilde{x}_j = P_{max}$ without affecting the optimal value. (d) results from (35) and the optimal solution is

$$\mathbf{v}_i = \frac{\tilde{\mathbf{J}}_i^{-1}\mathbf{H}_{ii}^H\mathbf{u}_i}{\|\tilde{\mathbf{J}}_i^{-1}\mathbf{H}_{ii}^H\mathbf{u}_i\|}. \tag{38}$$

For fixed $\mathbf{V}$, as analyzed in last subsection, we choose $\tilde{\mathbf{x}}$ to be the singular vector associated with $\rho(\tilde{\mathbf{B}})$.

To update the transmit beamforming vector, each transmitter needs to know the receive beamforming vectors of all the receivers, the relay processing matrices, all the channel matrices from the relays to the receivers and all the channel matrices from this transmitter to the relays.

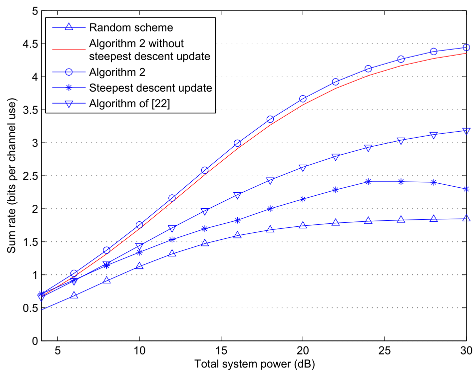

Fig. 1. Relay optimization.

After updating the transmit and receive beamforming vectors, we calculate the current $\rho(\mathbf{\Omega})$ and set the new SINR as $\gamma_i/\rho(\mathbf{\Omega}),\ \forall i$. Then the necessary and sufficient condition is satisfied, i.e., with the new SINR $\max_{\mathbf{U},\mathbf{V}} \alpha^*(\mathbf{U},\mathbf{V}) = 1$.

VI. Simulation

In this section, we present the numerical result to analyze the performance of our proposed algorithm. All the channel coefficients are independent and identically distributed (i.i.d.) zero mean unit variance circularly symmetric complex Gaussian. The noise power for all nodes is normalized to unit. All the simulation are conducted for the three-user network with two relays, i.e., $N = 3, K = 2$.

A. Optimization of Relay Processing Matrices

First, we focus on the performance of Algorithm 2, the algorithm to optimize relay processing matrices towards Pareto optimality. We assume all nodes are equipped with two antennas, i.e., $N_s = N_d = N_r = 2$. We plot the simulation results in Fig. 1.

In Fig. 1, the horizontal axis is the total system power in dB. The vertical axis is the sum rate of all the users. With an initial set of SINRs, random beamformers and relay matrices, we optimize the power of users and calculate the sum rate. The obtained plot is the plot of random scheme. If we apply step 17-28 in Algorithm 2 to the initial set of SINRs, we get the plot of steepest descent update, which is the result of steepest descent update based on a random initialization. The plot of Algorithm 2 without steepest descent update results from bisection search without the steepest descent update, i.e., all the steps before 21. The plot of Algorithm 2 is the final result of the whole Algorithm 2.

We can see the plot of Algorithm 2 without steepest descent update is very close to that of Algorithm 2. So in practice we can omit step 22-28. These two plots outperform all the other plots in the figure. Besides, the plot of the steepest descent update is not monotonically increasing. This is because the steepest descent method starts with random initialization and may end at a local optimal point.

B. Joint Optimization

Joint optimization refers to updating the beamformers by (34) and (38) and the relay matrices by Algorithm 2 iteratively until converging. The simulation of this part is conducted for the case with two antennas at all the nodes, i.e., $N_s = N_d = N_r = 2$. We first give examples to show the convergence, where the total system power is $P_{sys} = 20$(dB).

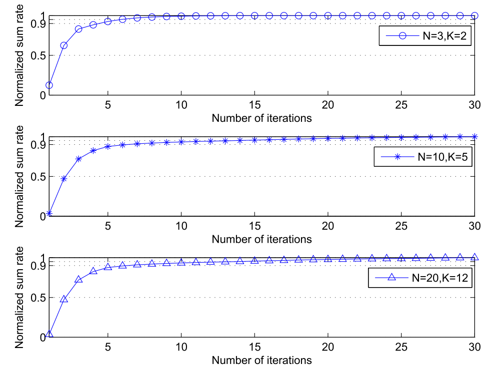

Fig. 2. Convergence of joint optimization.

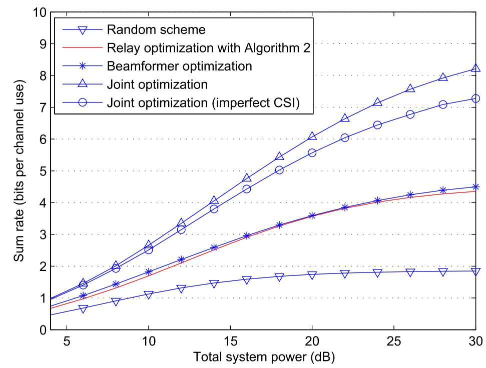

Fig. 3. Joint optimization.

In Fig. 2, the vertical axis is rate normalized to the rate at convergence. One iteration means one update of the beamformers and the relay matrices. The three plots show the convergence performances for 3 users with 2 relays, 10 users with 5 relays and 20 users with 12 relays respectively. We see that the further gain is less than 5% beyond 15 iterations for all the three cases.

In Fig. 3, the horizontal axis is the total system power P_{sys} in dB. The vertical axis is the sum rate of all the users over one time slot. The plot of random scheme is obtained by choosing the relay processing matrices and beamforming vectors randomly and setting the power as (9). The red plot shows the result of Algorithm 2. The plot of beamformer optimization refers to updating beamformers by (34) and (38) iteratively until converging. Combining Algorithm 2 and the beamforming update, we have the plot of joint optimization. Joint optimization outperforms the others. To balance the computation complexity and performance, we can run the joint optimization within a limited number of iterations considering that the improvement of the sum rate is small after 15 iterations in Fig. 2.

We also plot the performance of joint optimization with imperfect CSI. The difference between the accurate channel coefficient and estimated channel coefficient is modelled as a complex Gaussian variable with zero mean and 1/16 variance. Notice that both joint optimization with perfect CSI and imperfect CSI outperform all the other compared schemes.

VII. Conclusions

In this paper, we study how to check and obtain a Pareto optimal set of SINRs in a multiuser relay network. We give various necessary and sufficient conditions for a set of SINRs to be Pareto optimal. Specifically, we deal with the Pareto optimality problem with respect to the power, the relay processing matrices and the beamforming vectors respectively, by taking advantage of SINR balancing analysis. One key optimization work of this paper is the Pareto optimality with respect to the relay processing matrices, where FPP-SCA and multi-objective analysis are used to calculate a set of SINRs which is approximately Pareto optimal. Simulations show that our proposed algorithms outperform the compared schemes. We have developed these results based on the assumption of single stream transmission. We will consider generalizing the results to multi-stream transmission in the future.

Appendix A
Proof of Lemma 1

With α^* as the optimal value of Problem (6), we have

$$\frac{\mathrm{SINR}_i}{\gamma_i} \geq \alpha^* \ \forall i \tag{39}$$

$$\boldsymbol{\omega}^T \mathbf{p} \leq P_{max} \tag{40}$$

and at least one equality holds in (39). Otherwise, if $\frac{\mathrm{SINR}_i}{\gamma_i} > \alpha^* \ \forall i$, then we can find a bigger α^* while (39) and (40) are satisfied, which contradicts with the optimality of α^*.

First, we prove $\frac{\mathrm{SINR}_i}{\gamma_i} = \alpha^* \ \forall i$ by contradiction. Suppose $\frac{\mathrm{SINR}_i}{\gamma_i} > \alpha^*$ for some i, we can always decrease p_i such that $\frac{\mathrm{SINR}_i}{\gamma_i} > \alpha^*$ is still satisfied. Decreasing p_i results in increasing of $\mathrm{SINR}_j, \forall j \neq i$. Consequently, $\frac{\mathrm{SINR}_j}{\gamma_j} > \alpha^* \ \forall j$. Moreover, since $p_j, \forall j \neq i$ don't change and p_i decreases, the power constraint $\boldsymbol{\omega}^T \mathbf{p} \leq P_{max}$ is still satisfied. As analyzed above, a bigger optimal value can be found and α^* is not optimal. Hence there doesn't exist any i with $\frac{\mathrm{SINR}_i}{\gamma_i} > \alpha^*$.

Second, we prove $\boldsymbol{\omega}^T \mathbf{p} = P_{max}$ by contradiction. Suppose $\boldsymbol{\omega}^T \mathbf{p} < P_{max}$. Let $P_{sum} = \boldsymbol{\omega}^T \mathbf{p}$ and $a = \frac{P_{max}}{P_{sum}}$, then $P_{sum} < P_{max}$ and $a > 1$. Set $a\mathbf{p}$ as the new power vector, we have $\boldsymbol{\omega}^T (a\mathbf{p}) = P_{max}$. The new SINR of user i is

$$\begin{aligned} \mathrm{SINR}'_i &= \frac{h_{ii} p_i a}{\sum_{j\neq i} h_{ij} p_j a + \sigma_i} \\ &= \frac{h_{ii} p_i}{\sum_{j\neq i} h_{ij} p_j + \sigma_i \frac{1}{a}} \\ &> \frac{h_{ii} p_i}{\sum_{j\neq i} h_{ij} p_j + \sigma_i} \\ &= \mathrm{SINR}_i, \ \forall i. \end{aligned} \tag{41}$$

Consequently, $\frac{\mathrm{SINR}'_i}{\gamma_i} > \alpha^* \ \forall i$, which means α^* is not optimal. Thus the assumption $\boldsymbol{\omega}^T \mathbf{p} < P_{max}$ can't hold. We have $\boldsymbol{\omega}^T \mathbf{p} = P_{max}$.

Appendix B
Derivation of (9)

Substituting the SINR expression (4) into (7), we can rewrite as $\frac{1}{\alpha^*}\mathbf{p} = \mathbf{D}\bar{\mathbf{H}}_I \mathbf{p} + \mathbf{D}\boldsymbol{\sigma}$. Multiplying both sides by $\boldsymbol{\omega}^T$ and dividing both sides by P_{max}, we have $\frac{1}{\alpha^*} = \frac{1}{P_{max}}\boldsymbol{\omega}^T \mathbf{D}\bar{\mathbf{H}}_I \mathbf{p} + \frac{1}{P_{max}}\boldsymbol{\omega}^T \mathbf{D}\boldsymbol{\sigma}$. Combining the above two equations, we obtain (9).

Appendix C
Proof of Theorem 1

Before proving the above theorem, we give two lemmas that are needed in the proof.

Lemma 3: The optimal value of Problem (6) α^* *can be expressed as* $\alpha^* = 1/\rho(\bar{\mathbf{H}}_I \mathbf{D} + \frac{1}{P_{max}}\boldsymbol{\sigma}\boldsymbol{\omega}^T \mathbf{D})$.

Proof: Based on the mathematical rule that for any $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $\mathbf{B} \in \mathbb{R}^{n\times m}$, $\rho(\mathbf{AB}) = \rho(\mathbf{BA})$, we can rewrite the spectral radius of $\boldsymbol{\Omega}$ as

$$\begin{aligned} \rho(\boldsymbol{\Omega}) &= \rho(\begin{bmatrix} \mathbf{D} \\ \frac{1}{P_{max}}\boldsymbol{\omega}^T \mathbf{D} \end{bmatrix} \begin{bmatrix} \bar{\mathbf{H}}_I & \boldsymbol{\sigma} \end{bmatrix}) \\ &= \rho(\begin{bmatrix} \bar{\mathbf{H}}_I & \boldsymbol{\sigma} \end{bmatrix} \begin{bmatrix} \mathbf{D} \\ \frac{1}{P_{max}}\boldsymbol{\omega}^T \mathbf{D} \end{bmatrix}). \end{aligned} \tag{42}$$

Hence $\alpha^* = 1/\rho(\boldsymbol{\Omega}) = 1/\rho(\bar{\mathbf{H}}_I \mathbf{D} + \frac{1}{P_{max}}\boldsymbol{\sigma}\boldsymbol{\omega}^T \mathbf{D})$. ■

Lemma 4: Denote $f_\rho(\boldsymbol{\gamma}) = \rho\big((\bar{\mathbf{H}}_I + \frac{1}{P_{max}}\boldsymbol{\sigma}\boldsymbol{\omega}^T)\mathbf{D}_h \boldsymbol{\Gamma}\big)$ *as a function about SINR vector* $\boldsymbol{\gamma} \triangleq [\gamma_1, \ldots, \gamma_N]^T$ *with* $\mathbf{D}_h = diag([1/h_{11}, \ldots, 1/h_{NN}])$ *and* $\boldsymbol{\Gamma} = diag(\boldsymbol{\gamma})$. *Then* $f_\rho(\boldsymbol{\gamma})$ *is strictly increasing in* $\boldsymbol{\gamma}$*, i.e., for any two SINR vectors* $\boldsymbol{\gamma}$ *and* $\boldsymbol{\gamma}'$ *with* $\boldsymbol{\gamma}' \geq \boldsymbol{\gamma}$ *and* $\boldsymbol{\gamma}' \neq \boldsymbol{\gamma}$, $f_\rho(\boldsymbol{\gamma}') > f_\rho(\boldsymbol{\gamma})$.

Proof: We first introduce the definition of the irreducible matrix (see, e.g. [35]).

Definition 1: An $n \times n$ *nonnegative matrix* $\mathbf{A}$ *is* irreducible *if for every pair* i, j *of its index set, there exists a positive integer* $m \triangleq m(i, j)$ *such that* $A_{ij}^m > 0$.

In the definition, $m(i, j)$ depends on the index pair i, j and A_{ij}^m refers to the (i, j)th element of $\mathbf{A}^m$, where $\mathbf{A}^m$ is the multiplication of m $\mathbf{A}'s$. Every positive matrix is irreducible according to [32]. Next we prove Lemma 4. It is obvious that $\mathbf{T} \triangleq (\bar{\mathbf{H}}_I + \frac{1}{P_{max}}\boldsymbol{\sigma}\boldsymbol{\omega}^T)\mathbf{D}_h$ is a positive matrix. Hence it is irreducible. According to the corollary (3.29) in Chapter 1 of [32], if $\mathbf{0} \leq \mathbf{A} \leq \mathbf{B}$, where $\mathbf{A}$ is irreducible and $\mathbf{A} \neq \mathbf{B}$, then $\rho(\mathbf{A}) < \rho(\mathbf{B})$. Suppose $\boldsymbol{\gamma}$ is an arbitrary SINR vector and $\boldsymbol{\gamma}'$ is another arbitrary SINR vector. If $\boldsymbol{\gamma}' \geq \boldsymbol{\gamma}$ and $\boldsymbol{\gamma}' \neq \boldsymbol{\gamma}$, then $\mathbf{0} \leq \mathbf{T}\boldsymbol{\Gamma} \leq \mathbf{T}\boldsymbol{\Gamma}'$ and $\mathbf{T}\boldsymbol{\Gamma} \neq \mathbf{T}\boldsymbol{\Gamma}'$. Consequently $\rho(\mathbf{T}\boldsymbol{\Gamma}) < \rho(\mathbf{T}\boldsymbol{\Gamma}')$, i.e., $f_\rho(\boldsymbol{\gamma}') > f_\rho(\boldsymbol{\gamma})$. ■

Now we start to prove Theorem 1. For given $\mathbf{U}$, $\mathbf{V}$ and $\mathbf{M}$, it is self-explanatory that a set of SINRs $\{\gamma_1, \ldots, \gamma_N\}$ is achievable if and only if the optimal value of its associated SINR balancing problem (6) satisfies $\alpha^* \geq 1$. With this result, we prove $\alpha^* = 1$ by two parts.

i) *Only if.* Given that a set of SINRs $\{\gamma_1, \ldots, \gamma_N\}$ is Pareto optimal, we prove the *only if* part by contradiction. Suppose $\alpha^* > 1$, then we can always set a new SINR requirement $\tilde{\gamma}_i$ as $\tilde{\gamma}_i = \alpha^* \gamma_i \ \forall i$. With the new set of SINRs, the optimal value of (6) is $\tilde{\alpha}^* = 1$.

This indicates the new SINR requirement is achievable. Furthermore, $\{\tilde{\gamma}_1, \dots, \tilde{\gamma}_N\}$ outperforms the original SINR requirement $\{\gamma_1, \dots, \gamma_N\}$, which contradicts the definition of Pareto optimality. Hence if $\{\gamma_1, \dots, \gamma_N\}$ is Pareto optimal, the optimal value of the corresponding SINR balancing problem (6) has to be $\alpha^* = 1$, i.e., $\rho(\bar{\mathbf{H}}_I\mathbf{D} + \frac{1}{P_{max}}\boldsymbol{\sigma}\boldsymbol{\omega}^T\mathbf{D}) = 1$ according to Lemma 3.

ii) *If.* Given that a set of SINRs $\{\gamma_1, \dots, \gamma_N\}$ satisfies $\rho(\bar{\mathbf{H}}_I\mathbf{D} + \frac{1}{P_{max}}\boldsymbol{\sigma}\boldsymbol{\omega}^T\mathbf{D}) = 1$, i.e., $f_\rho(\boldsymbol{\gamma}) = 1$. Suppose there exists some $\boldsymbol{\gamma}'$ such that $\boldsymbol{\gamma}' \geq \boldsymbol{\gamma}$ and $\boldsymbol{\gamma}' \neq \boldsymbol{\gamma}$. Then according to Lemma 4, $f_\rho(\boldsymbol{\gamma}') > f_\rho(\boldsymbol{\gamma})$, i.e., $\rho\big((\bar{\mathbf{H}}_I + \frac{1}{P_{max}}\boldsymbol{\sigma}\boldsymbol{\omega}^T)\mathbf{D}_h\boldsymbol{\Gamma}'\big) > 1$ where $\boldsymbol{\Gamma}' = \mathrm{diag}(\boldsymbol{\gamma}')$. This results in $\alpha^* < 1$, which means that the new SINR vector $\boldsymbol{\gamma}'$ is not achievable. Hence there doesn't exist a set of SINRs outperforming $\{\gamma_1, \dots, \gamma_N\}$. We can conclude that $\{\gamma_1, \dots, \gamma_N\}$ satisfying $\rho(\bar{\mathbf{H}}_I\mathbf{D} + \frac{1}{P_{max}}\boldsymbol{\sigma}\boldsymbol{\omega}^T\mathbf{D}) = 1$ is Pareto optimal.

Appendix D
Proof of Lemma 2

We prove (13) by contradiction. Suppose $(\boldsymbol{\omega}^*)^T\mathbf{p} = \sum_i \mathbf{v}_i^H\mathbf{F}_i^H(\mathbf{M}^*)^H\mathbf{M}^*\mathbf{F}_i\mathbf{v}_i p_i < P_{max}$. Set $\beta\mathbf{M}^*$ as the new $\mathbf{M}$, where $\beta = \sqrt{\frac{P_{max}}{(\boldsymbol{\omega}^*)^T\mathbf{p}}} > 1$. Recall the SINR expression (4) and write it as a function of $\mathbf{M}$ since it is the only variable considered here. We can check that all the SINRs are increased, i.e.,

$$\begin{aligned}&\mathrm{SINR}_i(\beta\mathbf{M}^*)\\ &= \frac{|\mathbf{u}_i^H\mathbf{G}_i\mathbf{M}^*\mathbf{F}_i\mathbf{v}_i|^2 p_i}{\sum\limits_{j\neq i}|\mathbf{u}_i^H\mathbf{G}_i\mathbf{M}^*\mathbf{F}_j\mathbf{v}_j|^2 p_j + \sigma^2\mathbf{u}_i^H\mathbf{G}_i\mathbf{M}^*(\mathbf{M}^*)^H\mathbf{G}_i^H\mathbf{u}_i + \frac{\sigma^2}{\beta^2}}\\ &> \mathrm{SINR}_i(\mathbf{M}^*). \end{aligned} \tag{43}$$

Since $\mathrm{SINR}_i(\beta\mathbf{M}^*) > \mathrm{SINR}_i(\mathbf{M}^*)\ \forall i$ and $\beta\mathbf{M}^*$ satisfies the power constraint, $\beta\mathbf{M}^*$ results in a bigger optimal value. Thus α^* is not optimal. Consequently, the assumption $(\boldsymbol{\omega}^*)^T\mathbf{p} = \sum_i(1 + \mathbf{v}_i^H\mathbf{F}_i^H(\mathbf{M}^*)^H\mathbf{M}^*\mathbf{F}_i\mathbf{v}_i)p_i < P_{max}$ can't hold. We have $\boldsymbol{\omega}^T\mathbf{p} = P_{max}$ when (12) is solved.

Appendix E
Proof of Theorem 2

We prove the above theorem by looking at the cases when it is possible to increase an SINR without decreasing another.

i) *There is at least one inactive constraint.* Without loss of generality, we assume the ith constraint is inactive, i.e., $f_i(\mathbf{y}) < 0$. Keeping $\mathbf{y}$ and γ_j, $\forall j \neq i$, unchanged, and replacing γ_i with γ_i' where $\gamma_i < \gamma_i' \leq \frac{\mathbf{y}^H\mathbf{C}_i\mathbf{y}}{\mathbf{y}^H\mathbf{A}_i\mathbf{y}}$, the ith constraint is still satisfied. Hence we can increase the ith SINR without decreasing any other SINR if the ith constraint is inactive.

ii) *All the constraints are active.* If we can find a new $\mathbf{y}'$ such that $\mathbf{f}(\mathbf{y}') \leq \mathbf{f}(\mathbf{y}) = \mathbf{0}$ and $\mathbf{f}(\mathbf{y}') \neq \mathbf{f}(\mathbf{y})$, then a certain constraint is inactive and we can increase the corresponding SINR, as in case i).

Based on the above analysis, the case when there is no possibility to increase some SINR without decreasing the others is that $f_i(\mathbf{y}) = 0\ \forall i$ and there does not exist a $\mathbf{y}'$ such that $\mathbf{f}(\mathbf{y}') \leq \mathbf{f}(\mathbf{y})$ and $\mathbf{f}(\mathbf{y}') \neq \mathbf{f}(\mathbf{y})$.

Appendix F
Proof of Theorem 3

With the resultant SINR set, $f_i = 0, \forall\ i$. Step 16 of Algorithm 2 results in $t = 1$, so $f_i = 0$ is equivalent to $\gamma_i\mathbf{y}^H\mathbf{A}_i\mathbf{y} = \mathbf{y}^H\mathbf{C}_i\mathbf{y}$. By substituting $\mathbf{y} = \mathbf{Q}_{sum}^{\frac{1}{2}}\mathbf{m}$ and the specific expression of $\mathbf{A}_i$, $\mathbf{C}_i$ and $\mathbf{Q}_{sum}$, and using $\mathbf{m}^H\mathbf{Q}_{sum}\mathbf{m} = 1$, we have

$$\begin{aligned}\gamma_i\sum_{j\neq i}p_jh_{ij} + \gamma_i\sigma_i &= p_ih_{ii}, \forall\ i\\ \boldsymbol{\omega}^T\mathbf{p} &= P_{max}\end{aligned} \tag{44}$$

whose compact form is $\boldsymbol{\Omega}\mathbf{p}_{ext} = \mathbf{p}_{ext}$ with $\boldsymbol{\Omega}$ expressed in (10). For the eigensystem (9), we have mentioned $\rho(\boldsymbol{\Omega})$ is the only singular value whose associated singular vector is also positive. So $\rho(\boldsymbol{\Omega}) = 1$, which is equivalent to (11).

References

[1] J. G. Andrews *et al.*, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.

[2] M. N. Tehrani, M. Uysal, and H. Yanikomeroglu, "Device-to-device communication in 5G cellular networks: Challenges, solutions, and future directions," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 86–92, May 2014.

[3] A. B. MacKenzie and S. B. Wicker, "Game theory in communications: Motivation, explanation, and application to power control," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, vol. 2. Nov. 2001, pp. 821–826.

[4] D. Goodman and N. Mandayam, "Power control for wireless data," *IEEE Pers. Commun.*, vol. 7, no. 2, pp. 48–54, Apr. 2000.

[5] H. Ji and C.-Y. Huang, "Non-cooperative uplink power control in cellular radio systems," *Wireless Netw.*, vol. 4, no. 3, pp. 233–240, Mar. 1998.

[6] A. Al Daoud, T. Alpcan, S. Agarwal, and M. Alanyali, "A Stackelberg game for pricing uplink power in wide-band cognitive radio networks," in *Proc. 47th IEEE Conf. Decision Control (CDC)*, Dec. 2008, pp. 1422–1427.

[7] S. Ren and M. van der Schaar, "Pricing and distributed power control in wireless relay networks," *IEEE Trans. Signal Process.*, vol. 59, no. 6, pp. 2913–2926, Jun. 2011.

[8] C. U. Saraydar, N. B. Mandayam, and D. J. Goodman, "Efficient power control via pricing in wireless data networks," *IEEE Trans. Commun.*, vol. 50, no. 2, pp. 291–303, Feb. 2002.

[9] H. Ding, C. He, and L. Jiang, "Performance analysis of fixed gain MIMO relay systems in the presence of co-channel interference," *IEEE Commun. Lett.*, vol. 16, no. 7, pp. 1133–1136, Jul. 2012.

[10] H. Dahrouj and W. Yu, "Coordinated beamforming for the multicell multi-antenna wireless system," *IEEE Trans. Wireless Commun.*, vol. 9, no. 5, pp. 1748–1759, May 2010.

[11] X. Shang, B. Chen, and H. V. Poor, "Multiuser MISO interference channels with single-user detection: Optimality of beamforming and the achievable rate region," *IEEE Trans. Inf. Theory*, vol. 57, no. 7, pp. 4255–4273, Jul. 2011.

[12] E. A. Jorswieck, E. G. Larsson, and D. Danev, "Complete characterization of the Pareto boundary for the MISO interference channel," *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 5292–5296, Oct. 2008.

[13] M. Mohseni, R. Zhang, and J. M. Cioffi, "Optimized transmission for fading multiple-access and broadcast channels with multiple antennas," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 8, pp. 1627–1639, Aug. 2006.

[14] J. Qiu, R. Zhang, Z.-Q. Luo, and S. Cui, "Optimal distributed beamforming for MISO interference channels," *IEEE Trans. Signal Process.*, vol. 59, no. 11, pp. 5638–5643, Nov. 2011.

[15] J. Park and Y. Sung, "On the Pareto-optimal beam structure and design for multi-user MIMO interference channels," *IEEE Trans. Signal Process.*, vol. 61, no. 23, pp. 5932–5946, Dec. 2013.

[16] R. Mochaourab, P. Cao, and E. Jorswieck, "Alternating rate profile optimization in single stream MIMO interference channels," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2013, pp. 4834–4838.

[17] P. Cao, E. A. Jorswieck, and S. Shi, "Pareto boundary of the rate region for single-stream MIMO interference channels: Linear transceiver design," *IEEE Trans. Signal Process.*, vol. 61, no. 20, pp. 4907–4922, Oct. 2013.
[18] V. Havary-Nassab, S. Shahbazpanahi, and A. Grami, "Optimal distributed beamforming for two-way relay networks," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1238–1250, Mar. 2010.
[19] M. R. A. Khandaker and Y. Rong, "Interference MIMO relay channel: Joint power control and transceiver-relay beamforming," *IEEE Trans. Signal Process.*, vol. 60, no. 12, pp. 6509–6518, Dec. 2012.
[20] E. Koyuncu and H. Jafarkhani, "Distributed beamforming in wireless multiuser relay-interference networks with quantized feedback," *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4538–4576, Jul. 2012.
[21] J. Joung and A. H. Sayed, "Multiuser two-way amplify-and-forward relay processing and power control methods for beamforming systems," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1833–1846, Mar. 2010.
[22] R. Hu and T.-M. Lok, "Toward Pareto optimality in multiuser relay networks," *IEEE Trans. Veh. Technol.*, vol. 66, no. 1, pp. 246–255, Jan. 2017.
[23] M. Schubert and H. Boche, "Solution of the multiuser downlink beamforming problem with individual SINR constraints," *IEEE Trans. Veh. Technol.*, vol. 53, no. 1, pp. 18–28, Jan. 2004.
[24] A. Tölli, H. Pennanen, and P. Komulainen, "SINR balancing with coordinated multi-cell transmission," in *Proc. WCNC*, 2009, pp. 1–6.
[25] G. Dartmann, W. Afzal, X. Gong, and G. Ascheid, "Low complexity cooperative downlink beamforming in multiuser multicell networks," in *Proc. 12th IEEE Int. Conf. Commun. Technol. (ICCT)*, Nov. 2010, pp. 717–721.
[26] O. Mehanna, K. Huang, B. Gopalakrishnan, A. Konar, and N. D. Sidiropoulos, "Feasible point pursuit and successive approximation of non-convex QCQPs," *IEEE Signal Process. Lett.*, vol. 22, no. 7, pp. 804–808, Jul. 2015.
[27] C. Tammer and A. Göpfert, "Theory of vector optimization," in *Multiple Criteria Optimization: State of the Art Annotated Bibliographic Surveys*. New York, NY, USA: Kluwer, 2003, pp. 1–70.
[28] K. T. Truong, P. Sartori, and R. W. Heath, Jr., "Cooperative algorithms for MIMO amplify-and-forward relay networks," *IEEE Trans. Signal Process.*, vol. 61, no. 5, pp. 1272–1287, Mar. 2013.
[29] Y. Shi, J. Zhang, and K. B. Letaief, "Coordinated relay beamforming for amplify-and-forward two-hop interference networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2012, pp. 2408–2413.
[30] A. Liu, V. K. N. Lau, and Y. Liu, "Duality and optimization for generalized multi-hop MIMO amplify-and-forward relay networks with linear constraints," *IEEE Trans. Signal Process.*, vol. 61, no. 9, pp. 2356–2365, May 2013.
[31] W. Cheng, M. Ghogho, Q. Huang, D. Ma, and J. Wei, "Maximizing the sum-rate of amplify-and-forward two-way relaying networks," *IEEE Signal Process. Lett.*, vol. 18, no. 11, pp. 635–638, Nov. 2011.
[32] A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*. Philadelphia, PA, USA: SIAM, 1994.
[33] W. Yang and G. Xu, "Optimal downlink power assignment for smart antenna systems," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, vol. 6. May 1998, pp. 3337–3340.
[34] J. Fliege and B. F. Svaiter, "Steepest descent methods for multicriteria optimization," *Math. Methods Oper. Res.*, vol. 51, no. 3, pp. 479–494, 2000.
[35] E. Seneta, *Non-Negative Matrices and Markov Chains*. New York, NY, USA: Springer, 2006.

Ruixue Hu (S'15) received the B.Eng. degree in communication engineering from the Harbin Institute of Technology, Harbin, Heilongjiang, China, in 2012. She is currently pursuing the Ph.D. degree with the Chinese University of Hong Kong, Hong Kong.

She is currently involved in the system modeling, analysis, and algorithm design of multiple-input multiple-output transmission in multiuser relay networks. Her research interests are mainly in wireless communications and networking.

Tat-Ming Lok (SM'03) received the B.Sc. degree in electronic engineering from the Chinese University of Hong Kong, and the M.S.E.E. degree and the Ph.D. degree in electrical engineering from Purdue University. He was a Post-Doctoral Research Associate with Purdue University.

He then joined the Chinese University of Hong Kong, where he is currently an Associate Professor. His research interests include communication theory, communication networks, signal processing for communications, and wireless systems. He was a Co-Chair of the Wireless Access Track of the IEEE VTC04 Fall. He has served on the TPCs of different international conferences, including IEEE ICC, VTC, Globecom, WCNC, and ISIT. He also served as an Associate Editor of the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY from 2002 to 2008. He has been serving as an Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS since 2015.