Joint Optimization of Source Power Allocation and Distributed Relay Beamforming in Multiuser Peer-to-Peer Relay Networks

Yong Cheng, *Student Member, IEEE*, and Marius Pesavento, *Member, IEEE*

***Abstract*—In this paper, we consider the joint optimization of the source power allocation and relay beamforming weights in distributed multiuser peer-to-peer (MUP2P) relay networks applying the amplify-and-forward (AF) protocol. We adopt a quality-of-service (QoS) based approach, in which the total power transmitted from all sources and relays is minimized while guaranteeing the prescribed QoS requirement of each source-destination pair. The QoS is modeled as a function of the receive signal-to-interference-plus-noise ratio (SINR) at the destinations. Unlike the existing contributions, the transmitted powers of the sources and the beamforming weights of the relays are optimized jointly in this paper. Introducing an appropriate transformation of variables, the QoS based source power allocation and distributed relay beamforming (PADB) problem can be equivalently transformed into a difference of convex (DC) program, which can be efficiently solved with local optimality using the constrained concave convex procedure (CCCP). Based on this procedure, we also propose an iterative feasibility search algorithm (IFSA) to find an initial feasible point of the DC program. The analytic study of the proposed solution confirms that it converges to a local optimum of the PADB problem. Numerical results show that our solution outperforms (in terms of the total transmitted power) the alternating optimization procedure and the exact penalty based DC algorithm. In addition, the proposed IFSA outperforms the alternating optimization algorithm in finding feasible points of the DC program (i.e., the equivalence of the PADB problem).**

***Index Terms*—Constrained concave–conves procedure (CCCP), difference of convex (DC) program, distributed relay beamforming, iterative feasibility search algorithm, joint optimization, multiuser peer-to-peer relay networks, source power allocation.**

I. Introduction

Multiuser peer-to-peer (MUP2P) relay networks, where K source-destination pairs communicate in a pairwise manner with the help of L relays, have recently received increasing attention in the wireless community [1], [3]–[17]. In MUP2P relay networks, when all network nodes operate in half-duplex mode, the concurrent information transfer between the K source-destination pairs over the L relays take place in two phases, carried out, e.g., in two consecutive time-slots. In the first time-slot, all sources simultaneously transmit to the relay nodes on the same frequency band, and in the second time-slot, all relays simultaneously transmit to the destinations, also on the same frequency band. In such MUP2P relay networks, the concurrent transmissions of the K data streams between the K source-destination pairs (i.e., K peers) are facilitated by the use of distributed relay beamforming (also referred to as network beamforming [14]) techniques[1] applied at the L relays [1], [3]–[14].

Early works on MUP2P relay networks have focused on the optimization of the distributed relay beamforming weights with fixed source transmit power allocations [4]–[10], but not the joint optimization of both, source transmit powers and relay beamforming weights. Apparently, a joint approach, though more challenging due to the nonconvex nature of the problem, can provide a significant performance improvement over the existing approaches [4]–[10], due to the increased number of degrees of freedom in the joint design. The authors of [15]–[17] have considered the optimization of the source and relay transmit powers. In these works [15]–[17], the relay scaling factors are confined to be real and nonnegative, which induces a performance loss compared with distributed relay beamforming with complex relay weighting coefficients [1], [4]–[14].

More recently, the authors of [1], [11], and [13] have studied the joint optimization of the source power allocation and relay beamforming weights in distributed MUP2P relay networks. However, rather than a truly joint optimization, in these works [1], [11], [13], the alternating optimization approach has been taken, in which the source transmit powers are optimized while the relay beamforming weights are fixed, and vice versa. This procedure is carried out iteratively until the convergence of the objective function value [1], [11], [13]. Although superior performance of the alternating optimization approach over the existing relay beamforming designs [5], [8] which do not involve source power allocation has been confirmed through simulations in [1], [11], and [13], its computational complexity is very high and an analytic characterization of the alternating optimization framework [1], [11], [13] seems intractable.

Manuscript received July 15, 2011; revised October 30, 2011 and February 08, 2012; accepted February 13, 2012. Date of publication February 28, 2012; date of current version May 11, 2012. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Dominic K. C. Ho. This work was supported by the European Research Council (ERC) Advanced Investigator Grants Program under Grant 227477-ROSE, and the LOEWE Priority Program Cocoon.

The authors are with the Communication Systems Group, Technische Universität Darmstadt, D-64283 Darmstadt, Germany (e-mail: yong.cheng@nt.tu-darmstadt.de; marius.pesavento@nt.tu-darmstadt.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.2012.2189388

[1]In this work, the term ”distributed relay beamforming” refers to the scenario where the L relays collaboratively form beams towards the destinations by means of coherent data processing. However, the payload data are not exchanged between the relays. The relay beamforming weights are, e.g., computed at a central unit and then fed, e.g., to the relays [1], [3]–[14].

In this paper, unlike the alternating optimization framework [1], [11], [13], we take a truly joint optimization approach to address the source power allocation and distributed relay beamforming (PADB) problem in MUP2P relay networks under amplify-and-forward (AF) relaying, which is based on the difference of convex (DC) programming[2] technique [2], [18], [19]. Specifically, we model the QoS of each source-destination pair by means of its receive signal-to-interference-plus-noise ratio (SINR) at the destination node, and the design objective is to minimize the overall power transmitted from the source and relay nodes, while guaranteeing that the receive SINR of each destination node is above a predefined threshold. In summary, the main contributions of this paper are as follows:

- We propose a QoS based approach for the PADB problem in AF MUP2P relay networks, in which the source transmitted powers and the relay beamforming weights are optimized jointly (and simultaneously), rather than in an alternating procedure [1], [11], [13].
- We transform the PADB problem into a DC program [2], [18], [19], by applying an appropriate transformation of variables.
- We propose a low-complexity iterative algorithm to solve the DC program, which is based on the constrained concave-convex procedure (CCCP) [20]–[22]. We further introduce a novel initialization procedure that is based on a *feasible* point of the PADB problem obtained from a novel iterative feasibility search procedure, rather than an arbitrary (infeasible) point as in the conventional CCCP [20]–[22]. This initialization procedure can also be applied to improve the CCCP in other DC programming problems [19].
- We prove analytically that the proposed solution converges to a local optimum of the DC program, and thus also a local optimum of the PADB problem.

Numerical results show that the proposed low-complexity solution outperforms (in terms of the total transmitted power) the alternating optimization method proposed in [1] and the exact penalty based DC algorithm (EP-DCA), which is widely used for solving DC constrained DC programs [2]. In addition, the proposed iterative feasibility search algorithm (IFSA) performs better than the alternating optimization algorithm [1] in finding feasible points of the DC program, while the performance of the proposed IFSA and the EP-DCA of [2] is comparable.

The rest of this paper is organized as follows. The MUP2P relay network model and the two-hop data transmission schemes are introduced in Section II. In Section III, the QoS based PADB problem is formulated and the problem is then transformed into a DC program using an appropriate transformation of variables. In Section IV, we propose a low-complexity solution of the DC program and the IFSA, together with the analytical studies of the proposed algorithms. Numerical results and discussions are presented in Section V. Finally, Section VI concludes the paper with a summary of the main results.

[2]DC programs are optimization problems whose objective and/or constraint functions are functions that can be written as DC functions [2], [18], [19].

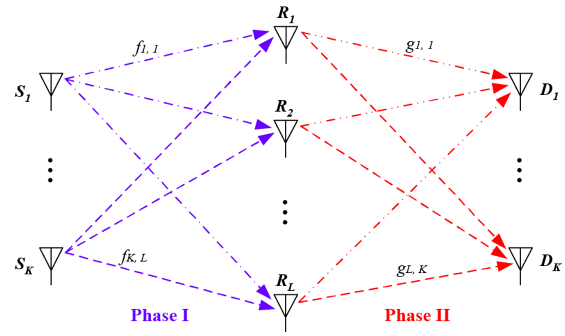


Fig. 1. A multiuser peer-to-peer (MUP2P) relay network, with K single-antenna sources $\{S_k\}_{k=1}^K$, L single-antenna relays $\{R_l\}_{l=1}^L$, and K single-antenna destinations $\{D_k\}_{k=1}^K$, sharing the same time and frequency radio resources. The source S_k intends to convey information to its corresponding destination D_k with the help of the L relay nodes.

Notations: Throughout this paper, vectors and matrices are denoted with boldface lowercase and uppercase letters, respectively. $\mathbb{R}_+$, $\mathbb{R}$, and $\mathbb{C}$ denote the set of *positive* real numbers, the set of real numbers, and the set of complex numbers, respectively. The operator $\mathrm{diag}\{\mathbf{q}\}$ constructs a diagonal matrix from the elements of the vector $\mathbf{q}$. $\mathrm{E}(\cdot)$ denotes the statistical expectation. $\mathbf{q}^T, (\mathbf{Q}^T)$ and $\mathbf{q}^H, (\mathbf{Q}^H)$ denote transpose and conjugate transpose of vector $\mathbf{q}$, (matrix $\mathbf{Q}$), respectively. $\mathrm{Re}\{\cdot\}$ denotes the real part of a variable. The vector $\mathbf{1}$ denotes a column vector whose elements are all ones. The symbol $\odot$ denotes the element-wise (Hadamard) product of two matrices. Finally, the symbol $\otimes$ denotes the Cartesian product of two sets.

II. System Model

In this section, we introduce the MUP2P relay network model, as well as the two-hop data transmission schemes under AF relaying.

A. Multiuser Peer-to-Peer Relay Network Model

Consider a MUP2P relay network with K sources $\{S_k\}_{k=1}^K$, K destinations $\{D_k\}_{k=1}^K$, and L relays $\{R_l\}_{l=1}^L$, as shown in Fig. 1, with all nodes equipped with a single antenna each.

The source S_k conveys independent information to its intended destination D_k, with the help of the L relays. When all the nodes operate in half-duplex mode, the two-hop data transmission takes place in two phases, e.g., in two consecutive time-slots. Specifically, in Phase I, all sources simultaneously transmit to the relays over a common frequency band, and in Phase II, all relays simultaneously transmit to the destinations, also sharing the same frequency band. In this paper, we consider the case where there is no direct link between sources and destinations and there is no message passing between any of the relays. Moreover, it is assumed that the source transmit powers and the relay beamforming weights are computed at a central node and then fed back to the sources and relays, respectively.

B. Two-Hop Data Transmission Under AF Relaying

Suppose that the AF relaying protocol is adopted at the relays. Then the relays simply amplify and forward their received signals to the destinations without decoding the information bits.

Let $r_l \in \mathbb{C}$ denote the received signal at the l^{th} relay in Phase I, under a frequency-flat channel model, the vector $\mathbf{r} \in \mathbb{C}^{L\times 1}$ of received signals at the L relays is given by

$$\mathbf{r} \triangleq [r_1,\ r_2,\ \ldots,\ r_L]^T = \sum_{j=1}^{K} \mathbf{f}_j \sqrt{p_j} x_j + \boldsymbol{\eta} \tag{1}$$

where

- the vector $\mathbf{f}_j \triangleq [f_{j,1},\ f_{j,2},\ \ldots,\ f_{j,L}]^T \in \mathbb{C}^{L\times 1}$, with $f_{j,l} \in \mathbb{C}$ denoting the frequency-flat channel coefficient from the j^{th} source to the l^{th} relay, $\forall j \in \mathcal{K},\ \forall l \in \mathcal{L}$, with $\mathcal{K} \triangleq \{1, 2, \ldots, K\}$ and $\mathcal{L} \triangleq \{1, 2, \ldots, L\}$;
- the vector $\mathbf{p} \triangleq [p_1,\ p_2,\ \ldots,\ p_K]^T \in \mathbb{R}_+^{K\times 1}$, with $p_j \in \mathbb{R}_+$ denoting the transmitted power of the j^{th} source, $\forall j \in \mathcal{K}$;
- the variable $x_j \in \mathbb{C}$ denotes the transmitted symbol of the j^{th} source, which is assumed to have unit power, i.e., $\mathrm{E}\{x_j x_j^H\} = 1,\ j \in \mathcal{K}$;
- the vector $\boldsymbol{\eta} \triangleq [\eta_1,\ \eta_2,\ \ldots,\ \eta_L]^T \in \mathbb{C}^{L\times 1}$, with $\eta_l \in \mathbb{C}$ denotes the additive white Gaussian noise (AWGN) at the l^{th} relay, with mean zero and variance[3] $\sigma_\eta^2,\ \forall l \in \mathcal{L}$.

In Phase II, the relays amplify and forward their received signals to the destinations. The received signal y_k at the k^{th} destination is given by

$$y_k = \mathbf{g}_k^T \mathbf{W}^H \mathbf{r} + \nu_k,\ \forall k \in \mathcal{K} \tag{2}$$

where

- the vector $\mathbf{g}_k \triangleq [g_{1,k},\ g_{2,k},\ \ldots,\ g_{L,k}]^T \in \mathbb{C}^{L\times 1}$, with $g_{l,k} \in \mathbb{C}$ denoting the frequency-flat channel coefficient from the l^{th} relay to the k^{th} destination, $\forall l \in \mathcal{L},\ \forall k \in \mathcal{K}$;
- the matrix $\mathbf{W} \triangleq \mathrm{diag}\{\mathbf{w}\}$, with the beamformer $\mathbf{w} \triangleq [w_1,\ w_2,\ \ldots,\ w_L]^T \in \mathbb{C}^{L\times 1}$ and w_l denoting the beamforming weight used by the l^{th} relay, $\forall l \in \mathcal{L}$;
- the variable $\nu_k \in \mathbb{C}$ denotes the AWGN at the k^{th} destination, with mean zero and variance $\sigma_\nu^2,\ \forall k \in \mathcal{K}$.

Substitute (1) into (2), we obtain the end-to-end input-output relationship of the source-destination pair $S_k \rightarrow D_k$ as

$$y_k = \mathbf{w}^H \mathbf{G}_k \mathbf{f}_k \sqrt{p_k} x_k + \underbrace{\sum_{j=1, j\neq k}^{K} \mathbf{w}^H \mathbf{G}_k \mathbf{f}_j \sqrt{p_j} x_j}_{\text{CCI}} + \underbrace{\mathbf{w}^H \mathbf{G}_k \boldsymbol{\eta} + \nu_k}_{\text{Noise}},\ \forall k \in \mathcal{K} \tag{3}$$

where the *diagonal* matrix $\mathbf{G}_k \in \mathbb{C}^{L\times L}$ is defined as: $\mathbf{G}_k \triangleq \mathrm{diag}\{\mathbf{g}_k\},\ \forall k \in \mathcal{K}$.

Similar to [1], [4]–[11], [13], and [15]–[17], we further assume that all the channel coefficients, data symbols, and the noises at the relays and destinations are statistically independent, and the co-channel interference (CCI) at the destinations is treated as noise, i.e., the destinations are assumed to perform single user detection. The receive SINR at the k^{th} destination, denoted as $\gamma_k(\mathbf{p}, \mathbf{w})$, can then be written as [1], [5], [8], [10], [11], and [13]:

$$\gamma_k(\mathbf{p}, \mathbf{w}) = \frac{p_k \mathbf{w}^H \mathbf{Q}_{k,k} \mathbf{w}}{\sum_{j=1, j\neq k}^{K} p_j \mathbf{w}^H \mathbf{Q}_{k,j} \mathbf{w} + \mathbf{w}^H \mathbf{D}_k \mathbf{w} + \sigma_\nu^2} \tag{4}$$

where the *Hermitian positive-semidefinite* matrices $\mathbf{Q}_{k,j} \in \mathbb{C}^{L\times L}$ and $\mathbf{D}_k \in \mathbb{C}^{L\times L}$ are defined as $\mathbf{Q}_{k,j} \triangleq \mathrm{E}\{\mathbf{G}_k \mathbf{f}_j \mathbf{f}_j^H \mathbf{G}_k^H\} = \mathrm{E}\{\mathbf{f}_j \mathbf{f}_j^H\} \odot \mathrm{E}\{\mathbf{g}_k \mathbf{g}_k^H\}$, and $\mathbf{D}_k \triangleq \sigma_\eta^2 \mathrm{E}\{\mathbf{G}_k \mathbf{G}_k^H\}$ (which is diagonal), respectively, $\forall k, j \in \mathcal{K}$.

To compute the optimal source power allocation $\mathbf{p}$ and relay beamforming weight vector $\mathbf{w}$, the central processing node requires the knowledge of the second order statistics of all the channel coefficients. To be specific, the central processing node generally requires the knowledge of the co-variance matrices $\mathrm{E}\{\mathbf{f}_k \mathbf{f}_k^H\}$ and $\mathrm{E}\{\mathbf{g}_k \mathbf{g}_k^H\},\ \forall k \in \mathcal{K}$. However, in practice channel models as introduced in [4]–[6] can be applied, in which the central processing node only requires the knowledge of the channel means, i.e., $\mathrm{E}\{f_{j,l}\}$ and $\mathrm{E}\{g_{m,k}\},\ \forall j, k \in \mathcal{K},\ \forall l, m \in \mathcal{L}$, as detailed in Section V. After computing the optimal source power allocation $\mathbf{p}$ and relay beamforming weight vector $\mathbf{w}$, the central processing node feeds the transmit power p_k and beamforming weight w_l forward to the k^{th} source and l^{th} relay, respectively [5], [7]–[10].

III. QoS Based PADB Problem Formulation

In this section, we first formulate the QoS based PADB problem for MUP2P relay networks and then transform the problem into a DC program through the use of an appropriate transformation of variables.

A. PADB Problem Formulation

In the QoS based PADB problem, the objective is to jointly minimize the total transmitted power of all sources and relays, while maintaining a minimum QoS level for each source–destination pair [1], [5], [7]–[11], [13], [15]–[17]. Similar to [1], [5], [7]–[11], [13], and [15]–[17], for the K source-destination pairs, we define the following QoS constraints:

$$\gamma_k(\mathbf{p}, \mathbf{w}) \geq \gamma_k^{(\min)},\ \forall k \in \mathcal{K} \tag{5}$$

where $\gamma_k^{(\min)} \in \mathbb{R}$ is the predefined receive SINR threshold at the destination node $D_k,\ \forall k \in \mathcal{K}$. Then, the QoS based PADB problem in the MUP2P relay network can be formulated as the following optimization problem:

$$\min_{\mathbf{p},\mathbf{w}}\ \psi(\mathbf{p}, \mathbf{w}) \triangleq \mathbf{1}^T \mathbf{p} + \mathbf{w}^H \left(\sum_{k=1}^{K} p_k \mathbf{F}_k + \sigma_\eta^2 \mathbf{I} \right) \mathbf{w} \tag{6a}$$

$$\text{s.t.}\ \ \gamma_k(\mathbf{p}, \mathbf{w}) \geq \gamma_k^{(\min)},\ \forall k \in \mathcal{K} \tag{6b}$$

$$\mathbf{p} \in \mathbb{R}_+^{K\times 1}, \mathbf{w} \in \mathbb{C}^{L\times 1} \tag{6c}$$

where the term $\mathbf{1}^T \mathbf{p}$ and the term $\mathbf{w}^H \left(\sum_{k=1}^{K} p_k \mathbf{F}_k + \sigma_\eta^2 \mathbf{I} \right) \mathbf{w}$ denote the total power transmitted from the K sources

[3]For notational simplicity, we assume that the noises $\{\eta_l,\ \forall l \in \mathcal{L}\}$ at the L relays are independent and identically distributed (i.i.d.), and the noises $\{\nu_k,\ \forall k \in \mathcal{K}\}$ at the K destinations are also i.i.d. The results of this paper hold, however, also for nonidentically distributed noises at the relays (destinations).

and the L relays, respectively, with the diagonal Hermitian positive semidefinite matrix $\mathbf{F}_k$ defined as $\mathbf{F}_k \triangleq \mathrm{E}\left\{\mathrm{diag}\left\{\mathbf{f}_k\right\}\mathrm{diag}\left\{\mathbf{f}_k^H\right\}\right\}$. In this paper, it is assumed that the matrices $\{\mathbf{Q}_{k,j}, \mathbf{D}_k, \mathbf{F}_k, \forall k, j \in \mathcal{K}\}$, i.e., the second-order statistics of the channel coefficients, are known perfectly at the central node, which is a common assumption widely used in the existing literature [4]–[6], [14].

The PADB problem in (6) is computationally difficult[4] to solve because the objective function (6a) and the constraint set (6b) are nonconvex, and the problem cannot be easily converted into a convex problem [1], [11], [13]. As a result, finding a global optimum of problem (6) is computationally expensive or even intractable. In this case, designing low-complexity algorithms to compute local minima of problem (6) is more meaningful in practice. In the following, we shall transform the PADB problem (6) into an equivalent DC program using an appropriate transformation of variables. Then a low-complexity algorithm is proposed to solve the DC program, which yields a local minimum of the DC program, and thus also a local minimum of the PADB problem (6).

It is worth mentioning that our approach naturally extends to the cases where additional total and/or individual source transmit power constraints, as well as the total and/or individual relay transmit power constraints are included in the problem formulation (6). This is because such transmit power constraints are generally convex [10], [14] and are therefore easy to incorporate in the framework considered in this paper. For simplicity of presentation, we omit these additional constraints here.

B. DC Program Re-formulation of the PADB Problem

We assume without loss of generality that the minimum requirement of the receive SINR at the k^{th} destination is strictly larger than zero, i.e., $\gamma_k^{(\min)} > 0$. Otherwise, when $\gamma_k^{(\min)} = 0$, the k^{th} source-destination pair is actually not scheduled to be served and the variable p_k can be ignored. Therefore, we can focus on the case that all elements of the vector $\mathbf{p}$ are positive. Hence, we can introduce the following variable transformations:

$$q_k = \frac{1}{p_k}, \ \forall k \in \mathcal{K} \tag{7}$$

and we further define the vector

$$\mathbf{q} \triangleq \left[\frac{1}{p_1}, \ \frac{1}{p_2}, \ \ldots, \ \frac{1}{p_K}\right]^T \in \mathbb{R}_+^{K\times 1}. \tag{8}$$

With the notation introduced in (7), we can transform the objective function in (6a) into a strictly convex function and the SINR constraints in (6b) into inequality constraints of DC form [2], [18], [19]. This leads to a DC program reformulation of the PADB problem (6). To be specific, with the vector $\mathbf{q}$ defined in (8), the objective function (6a) can be transformed into

$$\psi(\mathbf{q}, \mathbf{w}) \triangleq \sum_{k=1}^{K} \frac{1}{q_k} + \left(\sum_{k=1}^{K} \frac{\mathbf{w}^H \mathbf{F}_k \mathbf{w}}{q_k} + \sigma_\eta^2 \mathbf{w}^H \mathbf{w}\right) \tag{9}$$

[4]Even with fixed source power allocation $\mathbf{p}$, problem (6) is still a nonconvex quadratically constrained quadratic program (QCQP) in the beamformer $\mathbf{w}$, which is very difficult to solve, see, e.g., [5], [8]–[10], and [23] for details.

which represents the sum of the convex functions $\left\{\frac{1}{q_k}, \ \forall k \in \mathcal{K}\right\}$ and $\mathbf{w}^H\mathbf{w}$, and the quadratic-over-linear functions of type $\left\{\frac{\mathbf{w}^H \mathbf{F}_k \mathbf{w}}{q_k}, \ \forall k \in \mathcal{K}\right\}$, where the latter functions have the following nice properties.

Lemma 1 (Convex Functions): Given the Hermitian positive semidefinite matrix $\mathbf{A} \in \mathbb{C}^{L\times L}$, the quadratic-over-linear function of type $\frac{\mathbf{w}^H \mathbf{A} \mathbf{w}}{q_k}$ is jointly convex[5] in the variables $(q_k, \mathbf{w}) \in \mathbb{R}_+ \otimes \mathbb{C}^{L\times 1}$.

Proof: Please refer to [24, Sec. 3.2.6] for the proof. ■

From the definitions of the matrices $\mathbf{D}_k$, $\mathbf{F}_k$ and $\mathbf{Q}_{k,j}$ we know that the matrices $\mathbf{D}_k, \mathbf{F}_k$, and $\mathbf{Q}_{k,j}$ are Hermitian positive semidefinite, $\forall k, j \in \mathcal{K}$. Therefore, as a direct application of Lemma 1, we have the following corollary.

Corollary 1 (strictly Convex Objective Function): The objective function given in (9) is strictly jointly convex in the variables $(\mathbf{q}, \mathbf{w}) \in \mathbb{R}_+^{K\times 1} \otimes \mathbb{C}^{L\times 1}$.

Proof: Please refer to Appendix A for the proof. ■

Similar to the transformation of the objective function in (9), substituting the vector $\mathbf{q}$ in (8) into (6b), the minimum SINR constraints in (6b) can be rewritten as

$$\beta_k(\mathbf{q}, \mathbf{w}) - \alpha_k(\mathbf{q}, \mathbf{w}) \triangleq \varphi_k(\mathbf{q}, \mathbf{w}) \leq 0, \ \forall k \in \mathcal{K} \tag{11}$$

where the functions $\alpha_k(\mathbf{q}, \mathbf{w})$ and $\beta_k(\mathbf{q}, \mathbf{w})$ are defined as

$$\alpha_k(\mathbf{q}, \mathbf{w}) \triangleq \frac{\mathbf{w}^H \mathbf{Q}_{k,k} \mathbf{w}}{q_k}, \tag{12}$$

$$\beta_k(\mathbf{q}, \mathbf{w}) \triangleq \gamma_k^{(\min)}\left(\sum_{j=1, j\neq k}^{K} \frac{\mathbf{w}^H \mathbf{Q}_{k,j} \mathbf{w}}{q_j} + \mathbf{w}^H \mathbf{D}_k \mathbf{w} + \sigma_\nu^2\right). \tag{13}$$

We remark that the functions $\alpha_k(\mathbf{q}, \mathbf{w})$ and $\beta_k(\mathbf{q}, \mathbf{w})$ in (12) and (13), respectively, are both *convex* functions jointly in the variables $(\mathbf{q}, \mathbf{w}) \in \mathbb{R}_+^{K\times 1} \otimes \mathbb{C}^{L\times 1}$, as can be observed from Lemma 1 and the fact that the summation of convex functions is also a convex function [24, Sec. 3.2]. Hence, the function $\varphi_k(\mathbf{q}, \mathbf{w})$ defined in (11) represents the difference of two convex functions, i.e., a DC decomposition [2], [18], [19].

As a summary, with the DC decomposition given in (11), the PADB problem (6) can be equivalently transformed into the following DC program:

$$\min_{\mathbf{q}, \mathbf{w}} \quad \psi(\mathbf{q}, \mathbf{w}) \tag{14a}$$

$$\text{s.t.} \quad \beta_k(\mathbf{q}, \mathbf{w}) - \alpha_k(\mathbf{q}, \mathbf{w}) \leq 0, \ \forall k \in \mathcal{K} \tag{14b}$$

$$\mathbf{q} \in \mathbb{R}_+^{K\times 1}, \mathbf{w} \in \mathbb{C}^{L\times 1} \tag{14c}$$

where the constraints given in (14b) represent inequality constraints of DC type [2], [18], [19]. For ease of elaboration, we

[5]A function $U(\mathbf{q}, \mathbf{w}) : \mathbb{R}_+^{K\times 1} \otimes \mathbb{C}^{L\times 1} \mapsto \mathbb{R}$ is jointly convex in the variables $(\mathbf{q}, \mathbf{w}) \in \mathbb{R}_+^{K\times 1} \otimes \mathbb{C}^{L\times 1}$, if

$$U\left(\tau\mathbf{q}^{(1)} + (1-\tau)\mathbf{q}^{(2)}, \tau\mathbf{w}^{(1)} + (1-\tau)\mathbf{w}^{(2)}\right) \leq \tau U\left(\mathbf{q}^{(1)}, \mathbf{w}^{(1)}\right) + (1-\tau) U\left(\mathbf{q}^{(2)}, \mathbf{w}^{(2)}\right), \tag{10}$$

for all $\mathbf{q}^{(1)}, \mathbf{q}^{(2)} \in \mathbb{R}_+^{K\times 1}$, $\mathbf{w}^{(1)}, \mathbf{w}^{(2)} \in \mathbb{C}^{L\times 1}$, $0 \leq \tau \leq 1$. Furthermore, the function is strictly jointly convex if strict inequality holds in (10) for $\left[\left(\mathbf{q}^{(1)}\right)^T, \ \left(\mathbf{w}^{(1)}\right)^T\right]^T \neq \left[\left(\mathbf{q}^{(2)}\right)^T, \ \left(\mathbf{w}^{(2)}\right)^T\right]^T$ and $0 < \tau < 1$ [24, Sec. 3.1], [25].

further introduce the compact notation $\mathbf{z} \triangleq \left[\mathbf{q}^T,\ \mathbf{w}^T\right]^T \in \mathbb{R}_+^{K\times 1} \otimes \mathbb{C}^{L\times 1}$ that will be used in the sequel.

Remark 1 (Equivalent Transformation): The PADB problem in (6) and the DC program in (14) are equivalent in the sense that all solutions of problem (14) can be transferred directly into the corresponding solutions of problem (6), and vice versa.

IV. The Proposed Iterative Algorithms and Convergence Analysis

In this section, we propose a low-complexity algorithmic solution for the DC program (14), which is based on the CCCP [21], [22]. In contrast to the conventional CCCP as proposed in [21], here we introduce a novel initialization with a feasible point of the DC program that is obtained from the proposed IFSA (see Section IV-B for details). Through analytic studies, we show that the proposed solution always yields a local minimum of the DC program (14), and thus a local minimum of the PADB problem (6).

A. The Proposed Low-Complexity Solution

The CCCP, first proposed in [21], describes an sequential convex programming method that is widely adopted for solving DC programs [21], [22]. The main idea of the CCCP based algorithm is to iteratively approximate the originally nonconvex feasible set in (14b) around the current point by a convex subset and then solve the resulting convex approximation in each iteration using, e.g., the standard primal-dual interior point methods [24, Sec. 11.7]. As the nonconvex part in problem (14) stems from the fact that the function $\alpha_k(\mathbf{z})$ is convex but not concave, we approximate this function in the n^{th} iteration by its first-order Taylor expansion $\hat{\alpha}_k\left(\mathbf{z}^{(n)}, \mathbf{z}\right)$ around the current point $\mathbf{z}^{(n)}$. According to [26], [27], the first-order Taylor expansion $\hat{\alpha}_k\left(\mathbf{z}^{(n)}, \mathbf{z}\right)$ of the real-valued function $\alpha_k(\mathbf{z})$ of the complex-valued vector $\mathbf{z}$ is given by (see, e.g., [26, Theorems 3 and 4]),

$$\hat{\alpha}_k\left(\mathbf{z}^{(n)}, \mathbf{z}\right) = \alpha_k\left(\mathbf{z}^{(n)}\right) + 2\mathrm{Re}\left\{\nabla\alpha_k\left(\mathbf{z}^{(n)}\right)^H\left(\mathbf{z}-\mathbf{z}^{(n)}\right)\right\},\ \forall k\in\mathcal{K} \quad (15)$$

which is an affine function in $\mathbf{z}$. Here, $\nabla\alpha_k\left(\mathbf{z}^{(n)}\right)$ denotes the *conjugate derivative* of the function $\alpha_k(\mathbf{z})$ with respect to (w.r.t.) the complex vector $\mathbf{z}$, evaluated at the point $\mathbf{z}^{(n)} \triangleq \left[\left(\mathbf{q}^{(n)}\right)^T,\ \left(\mathbf{w}^{(n)}\right)^T\right]^T$. With the variable $q_k^{(n)}$ being real and positive, the conjugate derivative $\nabla\alpha_k\left(\mathbf{z}^{(n)}\right)$ is given by [26], [27], [28, Appendix B]:

$$\nabla\alpha_k\left(\mathbf{z}^{(n)}\right) = \left[\frac{\left(\mathbf{w}^{(n)}\right)^H\mathbf{Q}_{k,k}\mathbf{w}^{(n)}}{-2\left(q_k^{(n)}\right)^2}, \left(\frac{\mathbf{Q}_{k,k}\mathbf{w}^{(n)}}{q_k^{(n)}}\right)^T\right]^T. \quad (16)$$

Inserting (16) into (15), the affine approximation $\hat{\alpha}_k\left(\mathbf{z}^{(n)}, \mathbf{z}\right)$ of the function $\alpha_k(\mathbf{z})$ can be written as, with $\mathbf{z} \triangleq \left[(\mathbf{q})^T,\ (\mathbf{w})^T\right]^T$,

$$\begin{aligned}\hat{\alpha}_k\left(\mathbf{z}^{(n)}, \mathbf{z}\right) &= \frac{\left(\mathbf{w}^{(n)}\right)^H\mathbf{Q}_{k,k}\mathbf{w}^{(n)}}{q_k^{(n)}} - \frac{\left(\mathbf{w}^{(n)}\right)^H\mathbf{Q}_{k,k}\mathbf{w}^{(n)}}{\left(q_k^{(n)}\right)^2}\left(q_k - q_k^{(n)}\right) + 2\mathrm{Re}\left\{\frac{\left(\mathbf{w}-\mathbf{w}^{(n)}\right)^H\mathbf{Q}_{k,k}\mathbf{w}^{(n)}}{q_k^{(n)}}\right\} \\ &= \mathrm{Re}\left\{\frac{\mathbf{w}^H\mathbf{Q}_{k,k}\mathbf{w}^{(n)}}{q_k^{(n)}}\right\} - \frac{\left(\mathbf{w}^{(n)}\right)^H\mathbf{Q}_{k,k}\mathbf{w}^{(n)}}{\left(q_k^{(n)}\right)^2}\left(q_k - q_k^{(n)}\right) - \frac{\left(\mathbf{w}^{(n)}\right)^H\mathbf{Q}_{k,k}\mathbf{w}^{(n)}}{q_k^{(n)}},\ \forall k\in\mathcal{K}. \end{aligned} \quad (17)$$

Then, in the n^{th} iteration of the proposed CCCP based iterative algorithm, the following *convex* optimization problem:

$$\min_{\mathbf{z}} \quad \psi(\mathbf{z}) \quad (18a)$$

$$\text{s.t.} \quad \beta_k(\mathbf{z}) - \hat{\alpha}_k\left(\mathbf{z}^{(n)}, \mathbf{z}\right) \le 0,\ \forall k\in\mathcal{K}, \quad (18b)$$

$$\mathbf{z}\in\mathbb{R}_+^{K\times 1}\otimes\mathbb{C}^{L\times 1}, \quad (18c)$$

is solved and the solution is denoted by $\mathbf{z}^{(n+1)}$. This procedure is carried out iteratively until convergence or until the maximum number of allowable iterations is reached.

We remark that, since the function $\alpha_k(\mathbf{z})$ is convex in the variable $\mathbf{z}\in\mathbb{R}_+^{K\times 1}\otimes\mathbb{C}^{L\times 1}$, it is minorized by its first-order Taylor expansion $\hat{\alpha}_k\left(\mathbf{z}^{(n)}, \mathbf{z}\right)$ [24, Sec. 3.1], i.e.,

$$\alpha_k(\mathbf{z}) \ge \hat{\alpha}_k\left(\mathbf{z}^{(n)}, \mathbf{z}\right),\ \forall k\in\mathcal{K} \quad (19)$$

which suggests that

$$\beta_k(\mathbf{z}) - \alpha_k(\mathbf{z}) \le \beta_k(\mathbf{z}) - \hat{\alpha}_k\left(\mathbf{z}^{(n)}, \mathbf{z}\right) \triangleq \hat{\varphi}_k(\mathbf{q}, \mathbf{w}). \quad (20)$$

From (20) above, we know that the convex constraints in (18b) can be considered as a strengthening of the original nonconvex constraints in (14b). In other words, the feasible set defined in (18b) is a *subset* of the true feasible set defined in (14b). As a result, provided that the initial point $\mathbf{z}^{(0)}$ is feasible for the DC program (14), then all the iterates, $\left\{\mathbf{z}^{(n)}\right\}$ generated by iteratively solving the convex optimization problem (18) with the affine approximation in (17), always belong to the true feasible set defined in (14b).

We summarize the proposed low-complexity solution as Algorithm 1 given in the table below, where we assume that an initial feasible point $\mathbf{z}^{(0)}$ of the DC program (14) is available (see Section IV-B for details on how to obtain an initial feasible point).

Algorithm 1: The Proposed Low-Complexity Solution

Initialization: Define the tolerance of accuracy ϵ_1 and the maximum number of iterations $N_1^{(\max)}$; Initialize the algorithm with a *feasible* point $\mathbf{z}^{(0)}$; Set the iteration number $n = 0$.

Repeat:

1: Compute the affine approximation $\hat{\alpha}_k\left(\mathbf{z}^{(n)}, \mathbf{z}\right)$ according to (17).

2: Solve problem (18), and assign the solution to $\mathbf{z}^{(n+1)}$.

3: Update the iteration number: $n \longleftarrow n + 1$.

Until: The sequence $\left\{\psi\left(\mathbf{z}^{(n)}\right)\right\}$ converges, i.e., $\left|\psi\left(\mathbf{z}^{(n+1)}\right) - \psi\left(\mathbf{z}^{(n)}\right)\right| < \epsilon_1$; or the maximum number of iterations is reached, i.e., $n \geq N_1^{(\max)}$.

Remark 2 (Low-Complexity Solution): Algorithm 1 provides a low-complexity solution, in the sense that in each step a simple convex optimization problem is solved. The proposed Algorithm 1 converges to a local optimum after a few iterations as can be observed from the numerical experiments (see Section V-C).

B. The Proposed Iterative Feasibility Search Algorithm

As a novel modification of the CCCP [21], the proposed Algorithm 1 presented in the previous subsection is initialized with a *feasible* point $\mathbf{z}^{(0)}$ of the DC program (14), rather than an arbitrary point as in the conventional CCCP [21]. The main advantage of the proposed new initialization method stems from the fact that, once the proposed algorithm starts with a point in the feasible set of the DC program (14), all the iterates $\left\{\mathbf{z}^{(n)}\right\}$ generated by the algorithm (i.e., the algorithm trajectory) remain within the original feasible set of the DC program (14). In addition, if the CCCP is initialized with a random (infeasible) point, the CCCP may fail at the first iteration due to the infeasibility of problem (18). However, the task of computing a feasible point of a nonconvex optimization problem, e.g., the PADB problem (14), is NP-hard in general [29]. This observation motivates the development of suboptimal, however, low-complexity feasibility search procedures.

Inspired by the phase I method [24, Sec. 11.4] and the IFSA of [10], we propose an IFSA to find an initial feasible point of the DC program (14). We remark that the proposed initialization method and the IFSA can straightforwardly be applied to solve other problems that can be formulated as DC programs [19].

The proposed IFSA is based on similar iterative affine approximations of the originally nonconvex constraints as used in Algorithm 1, but with the following two modifications: i) the proposed IFSA starts with an arbitrary (e.g., a random) point $\mathbf{z}^{(0)}$, and ii) in the n^{th} iteration, instead of minimizing the total transmitted power as in problem (18), we minimize the slack parameter $s \geq 0$, which can be regarded as an abstract measure of the constraint violations. The feasibility problem can then be expressed as the following convex program:

$$\min_{\mathbf{z},s} \quad s \tag{21a}$$

$$\text{s.t.} \quad \beta_k(\mathbf{z}) - \hat{\alpha}_k\left(\mathbf{z}^{(n)}, \mathbf{z}\right) \leq s, \ \forall k \in \mathcal{K} \tag{21b}$$

$$s \in \mathbb{R}_+, \ \mathbf{z} \in \mathbb{R}_+^{K\times 1} \otimes \mathbb{C}^{L\times 1} \tag{21c}$$

with $\hat{\alpha}_k\left(\mathbf{z}^{(n)}, \mathbf{z}\right)$ defined according to (17). If the current objective value $s^{(n+1)}$ is zero, then the algorithm stops; otherwise, the algorithm continues until convergence or until the maximum number of allowable iterations is reached[6]. If no feasible point could be found with the proposed method, some admission control mechanisms can be adopted to reduce the number of source-destination pairs, which, however, is out of the scope of this paper. The proposed CCCP based IFSA is summarized as Algorithm 2 given in the table below.

Algorithm 2: Iterative Feasibility Search Algorith (IFSA)

Initialization: Define the tolerance of accuracy ϵ_2 and the maximum number of iterations $N_2^{(\max)}$; *Initialize the algorithm with an arbitrary random point* $\mathbf{z}^{(0)} = \mathbf{z}^{(rnd)}$; Set the iteration number $n = 0$.

Repeat:

1: Compute the affine approximation $\hat{\alpha}_k\left(\mathbf{z}^{(n)}, \mathbf{z}\right)$ according to (17).

2: Solve problem (21), and assign the solution to $\mathbf{z}^{(n+1)}$ and $s^{(n+1)}$, respectively.

3: Check whether the current objective value $s^{(n+1)}$ is zero or not. *If it is indeed zero, then the algorithm stops*.

4: Update the iteration number: $n \longleftarrow n + 1$

Until: The sequence $\left\{s^{(n)}\right\}$ converges, i.e., $\left|s^{(n+1)} - s^{(n)}\right| < \epsilon_2$; or the maximum number of iterations is reached, i.e., $n \geq N_2^{(\max)}$.

We remark that a solution of problem (21) with $s = 0$ obtained, e.g., from Algorithm 2, is always feasible for the DC program (14). Conversely, however, if the proposed Algorithm 2 fails to provide a feasible point of problem (14), then this does not imply that this problem is infeasible as Algorithm 2 operates only on a subset of the original feasible set of the DC program in (14).

Remark 3 (Two-Stage Algorithm): The proposed IFSA in Algorithm 2 together with the low-complexity solution in Algorithm 1 (which is in fact a novel modification of the conventional CCCP [20]–[22]) forms a *two-stage* algorithm for solving the DC program in (14). In the first stage, the IFSA is applied to find a feasible point of the DC program (14), staring with an arbitrary (e.g., random) infeasible point. If the proposed IFSA fails to obtain a feasible point, the algorithm declares failure and stops. In the second stage, the proposed low-complexity solution in Algorithm 1 is applied, starting with the feasible point found in the first stage.

C. Convergence Analysis of the Proposed Algorithms

In general, the CCCP based iterative algorithms converge to stationary points of DC programs [20]–[22], which are *not* necessarily local optima of DC programs. However, we prove in

[6]As one final step, after convergence or the maximum number of allowable iterations is reached, the point $\mathbf{z}^{(n+1)}$ will be substituted back into (14b) to check whether it is feasible or not.

this section that, for the DC program (14), the proposed Algorithm 1 indeed converges to one of its *local minima*. The convergence proof is carried out in three steps, namely, i) proof of convergence of Algorithm 1, ii) proof of convergence of Algorithm 1 to a stationary point, and iii) proof that stationary points are local minima. We start the convergence analysis with several important observations.

1) Observations (Partial-Monotonicity): The following important properties of the objective function and the constraint functions regarding the DC program (14) and the convex approximation (18) can be formulated:

O1) We observe from (9) that the objective function $\psi(\mathbf{q},\mathbf{w})$ is strictly decreasing[7] in variable $q_k,\ \forall k\in\mathcal{K}$.

O2) From (11)–(13), we observe that the constraint function $\varphi_k(\mathbf{q},\mathbf{w})$ is *strictly increasing* in the variable q_k, and *strictly decreasing* in the variable $q_j,\ \forall j\neq k,\ \forall j,k\in\mathcal{K}$.

O3) From (13), (17), and (20), we observe that, the approximated constraint function $\hat{\varphi}_k(\mathbf{q},\mathbf{w})$ is *strictly increasing* in the variable q_k, and *strictly decreasing* in the variable $q_j,\ \forall j\neq k,\ \forall j,k\in\mathcal{K}$.

We now analyze the convergence of the proposed Algorithm 1. The convergence behavior of Algorithm 2 can be inferred accordingly. From (15), we know that the point $\mathbf{z}^{(n)}$ is a *feasible* point of the convex optimization problem in (18), provided that the initial point $\mathbf{z}^{(0)}$ is feasible for the DC program (14). As a consequence, the sequence $\left\{\psi\left(\mathbf{z}^{(n)}\right)\right\}$ *monotonically decreases* as the iteration number n increases. Since the sequence $\left\{\psi\left(\mathbf{z}^{(n)}\right)\right\}$ is lower-bounded by zero, the convergence of the sequence $\left\{\psi\left(\mathbf{z}^{(n)}\right)\right\}$, and thus the convergence of Algorithm 1 is guaranteed[8] for any initial feasible point $\mathbf{z}^{(0)}$.

Moreover, since the objective function $\psi(\mathbf{z})$ of problem (18) is strictly convex in $\mathbf{z}\in\mathbb{R}_+^{K\times 1}\otimes\mathbb{C}^{L\times 1}$ (see Lemma 1), the point $\mathbf{z}^{(n+1)}$, i.e., the solution of problem (18), is unique [24, Sec. 4.2]. Hence, for any given initial feasible point $\mathbf{z}^{(0)}$, the entries of the two sequences, $\left\{\psi\left(\mathbf{z}^{(n)}\right)\right\}$ and $\left\{\mathbf{z}^{(n)}\right\}$, have a one-to-one correspondence. As a result, the monotone convergence of the sequence $\left\{\psi\left(\mathbf{z}^{(n)}\right)\right\}$ implies the convergence of the sequence $\left\{\mathbf{z}^{(n)}\right\}$, for any initial feasible point $\mathbf{z}^{(0)}$. Let $\mathbf{z}^*\left(\mathbf{z}^{(0)}\right)$ denote the limit point of the sequence $\left\{\mathbf{z}^{(n)}\right\}$ with a feasible initialization $\mathbf{z}^{(0)}$ when the iteration number n goes to infinity, i.e., given the initial feasible point $\mathbf{z}^{(0)}$, we have

$$\mathbf{z}^*\left(\mathbf{z}^{(0)}\right)\triangleq\lim_{n\longrightarrow\infty}\mathbf{z}^{(n)}. \tag{22}$$

In general, the limit point $\mathbf{z}^*\left(\mathbf{z}^{(0)}\right)$ depends on the choice of the initial feasible point $\mathbf{z}^{(0)}$. Here, for notational simplicity, we write the limit point as $\mathbf{z}^*$, hence dropping the argument $\mathbf{z}^{(0)}$ that expresses this dependency. Regarding the limit point $\mathbf{z}^*$, we can make the following statement.

[7]Here, when we say a multivariate function is (strictly) monotone in one particular variable, we mean that, the function is (strictly) monotone in that particular variable, while all other variables are fixed.

[8]Following a similar argument, the convergence of Algorithm 2 is also guaranteed for any initial (infeasible) point.

Lemma 2 (properties of Limit Points): The limit point $\mathbf{z}^*$ of the sequence $\left\{\mathbf{z}^{(n)}\right\}$ generated by Algorithm 1, is the solution of the following convex optimization problem:

$$\min_{\mathbf{z}}\quad \psi(\mathbf{z}) \tag{23a}$$
$$\text{s.t.}\quad \beta_k(\mathbf{z})-\hat{\alpha}_k(\mathbf{z}^*,\mathbf{z})\leq 0,\ \forall k\in\mathcal{K} \tag{23b}$$
$$\mathbf{z}\in\mathbb{R}_+^{K\times 1}\otimes\mathbb{C}^{L\times 1} \tag{23c}$$

where the affine function $\hat{\alpha}_k(\mathbf{z}^*,\mathbf{z})$ is obtained by replacing $\mathbf{z}^{(n)}$ with $\mathbf{z}^*$ in (17). Moreover, the limit point $\mathbf{z}^*$ satisfies all the constraints in (23b) with *equalities*, i.e.,

$$\beta_k(\mathbf{z}^*)-\alpha_k(\mathbf{z}^*)=\beta_k(\mathbf{z}^*)-\hat{\alpha}_k(\mathbf{z}^*,\mathbf{z}^*)=0,\ \forall k\in\mathcal{K}. \tag{24}$$

Proof: Please refer to Appendix A for the proof. ■

Since the convex optimization problem in (23) is strictly feasible,[9] i.e., the Slater's condition for constraint qualifications is satisfied [24, Sec. 5.2], we know from Lemma 2 that there exist Lagrange multipliers $\{\lambda_k^*,\ \forall k\in\mathcal{K}\}$, together with the limit point $\mathbf{z}^*$, that satisfy the following Karush–Kuhn–Tucker (KKT) necessary and sufficient conditions for optimality[24, Sec. 5.5] of the convex optimization problem (23):

$$\nabla\psi(\mathbf{z}^*)+\sum_{k=1}^{K}\lambda_k^*\left(\nabla\beta_k(\mathbf{z}^*)-\nabla\alpha_k(\mathbf{z}^*)\right)=\mathbf{0} \tag{25}$$
$$\beta_k(\mathbf{z}^*)-\alpha_k(\mathbf{z}^*)=0,\ \forall k\in\mathcal{K} \tag{26}$$
$$\mathbf{z}^*\in\mathbb{R}_+^{K\times 1}\otimes\mathbb{C}^{L\times 1}. \tag{27}$$

We next prove that the limit point $\mathbf{z}^*$ is a stationary point of the DC program (14), which is then proved to be a local minimum. First of all, from observation O2) we know that, all the SINR constraints (14b) in the DC program (14) are active[10] at the (local) optimum $\mathbf{z}^\star$ of the DC program (14). Hence, the necessary KKT optimality conditions of the DC program (14) are given by the following system of equations [30, Lecture 26]:

$$\nabla\psi(\mathbf{z})+\sum_{k=1}^{K}\mu_k\left(\nabla\beta_k(\mathbf{z})-\nabla\alpha_k(\mathbf{z})\right)=\mathbf{0} \tag{28}$$
$$\beta_k(\mathbf{z})-\alpha_k(\mathbf{z})=0,\ \forall k\in\mathcal{K} \tag{29}$$
$$\mathbf{z}\in\mathbb{R}_+^{K\times 1}\otimes\mathbb{C}^{L\times 1} \tag{30}$$

where the variables $\{\mu_k,\ \forall k\in\mathcal{K}\}$ denote the Lagrange multipliers associated with the SINR constraints in (14b).

Comparing the KKT system (25)–(27), with the corresponding KKT system (28)–(30), we conclude that the limit point $\mathbf{z}^*$, together with the Lagrange multipliers $\{\lambda_k^*,\ \forall k\in\mathcal{K}\}$,

[9]By construction $\mathbf{z}^*\triangleq\left[(\mathbf{q}^*)^T,\ (\mathbf{w}^*)^T\right]^T$ is a feasible point of problem (23). Thus, we know from (13), (17), and Lemma 2 that any point $\mathbf{z}$ defined as $\mathbf{z}\triangleq\left[(\mathbf{q}^*)^T,\ (\xi\mathbf{w}^*)^T\right]^T$, with $\xi>1$, is a strictly feasible point of problem (23). Therefore, problem (23) is strictly feasible.

[10]This can be easily proved by contradiction. Specifically, suppose that the k^{th} constraint is not active, i.e., $\varphi_k(\mathbf{z}^\star)<0$, then we can scale up the variable $q_k^\star$ to make the constraint active, without violating the other SINR constraints. However, when the variable $q_k^\star$ increases, the objective function $\psi(\mathbf{z}^\star)$ decreases, which contradicts the optimality of the point $\mathbf{z}^\star$.

also satisfy the KKT conditions of the DC program (14), i.e., we can choose $\mu_k = \lambda_k^*,\ \forall k \in \mathcal{K}$. Hence, the limit point $\mathbf{z}^*$ is a stationary point of the DC program (14) [30, Lecture 26].

With the results above, we can formulate the following theorem regarding the convergence behavior of the proposed Algorithm 1.

Theorem 1 (Local Optimality): The limit point $\mathbf{z}^*\left(\mathbf{z}^{(0)}\right)$ of the sequence $\left\{\mathbf{z}^{(n)}\right\}$, generated by the proposed Algorithm 1 with an arbitrary feasible initialization $\mathbf{z}^{(0)}$, is a *local minimum* of the DC program in (14), since every stationary point is a local minimum for the DC program in (14).

Proof: Please refer to Appendix C for the proof. ■

Remark 4 (Local Optimality): From Theorem 1 we know that, from the proposed low-complexity solution in Algorithm 1, we can obtain a *local minimum* of the DC program (14), and thus a local minimum of the PADB problem in (6).

V. Numerical Results and Discussions

In this section, we present simulation results and discussions. We compare three schemes: i) the proposed low-complexity solution in Algorithm 1; ii) the alternating optimization framework of [1] for solving the PADB problem (6); and iii) the EP-DCA, which has been proposed for solving DC constrained DC programs [2] and which here has been adopted for solving the DC program (14). Both the alternating optimization method of [1] and the EP-DCA of [2] yield generally suboptimal solutions, which are not guaranteed to be KKT points of the PADB problem (6). The performance of the proposed ISFA in Algorithm 2 in finding feasible points of the DC program (14) is also compared with that of the alternating optimization method [1] and the EP-DCA [2]. We also demonstrate the convergence behavior of the proposed algorithms.

A. Simulation Settings

As in [4]–[6], it is assumed that the second-order statistics of the channel coefficients, i.e., the matrices $\mathbf{D}_k$, $\mathbf{F}_k$, and $\mathbf{Q}_{k,j},\ \forall j, k \in \mathcal{K}$, are known at the central processing node. Furthermore, we choose the same channel models used in [4]–[6] in our simulations. Specifically, channel coefficients $f_{j,l}$ and $g_{m,k} \in \mathbb{C}$, which are normalized by the noise power, can be modeled as [4]–[6]

$$f_{j,l} \triangleq \bar{f}_{j,l} + \tilde{f}_{j,l},\ \forall j \in \mathcal{K}, \forall l \in \mathcal{L} \quad (31)$$

$$g_{m,k} \triangleq \bar{g}_{m,k} + \tilde{g}_{m,k},\ \forall m \in \mathcal{L},\ \forall k \in \mathcal{K} \quad (32)$$

where $\bar{f}_{j,l} \in \mathbb{C}$ and $\bar{g}_{m,k} \in \mathbb{C}$ are the channel mean and $\tilde{f}_{j,l} \in \mathbb{C}$ and $\tilde{g}_{m,k} \in \mathbb{C}$ are zero-mean random variables, $\forall j, k \in \mathcal{K},\ \forall l, m \in \mathcal{L}$. According to [4]–[6], the channel mean $\bar{f}_{j,l}$ and $\bar{g}_{m,k}$ can be modeled, respectively, as

$$\bar{f}_{j,l} = \frac{\exp\left(\sqrt{-1}\Theta_{j,l}\right)}{\sqrt{\Upsilon_f}},\ \forall j \in \mathcal{K},\ \forall l \in \mathcal{L} \quad (33)$$

$$\bar{g}_{m,k} = \frac{\exp\left(\sqrt{-1}\Omega_{m,k}\right)}{\sqrt{\Upsilon_g}},\ \forall m \in \mathcal{L},\ \forall k \in \mathcal{K} \quad (34)$$

where the random angles $\Theta_{j,l}$ and $\Omega_{m,k}$ are chosen to be uniformly distributed on the interval $[0,\ 2\pi]$, $\forall j, k \in \mathcal{K},\ \forall l, m \in \mathcal{L}$, and Υ_f and Υ_g are positive constants, which indicate the uncertainty in the channel coefficients [4]–[6]. Furthermore, the variances of the random variables $\tilde{f}_{j,l}$ and $\tilde{g}_{m,k}$ are given by

$$\mathrm{E}\left\{\left|\tilde{f}_{j,l}\right|^2\right\} = \frac{\Upsilon_f}{1+\Upsilon_f},\ \forall j \in \mathcal{K},\ \forall l \in \mathcal{L} \quad (35)$$

$$\mathrm{E}\left\{\left|\tilde{g}_{m,k}\right|^2\right\} = \frac{\Upsilon_g}{1+\Upsilon_g},\ \forall m \in \mathcal{L},\ \forall k \in \mathcal{K}. \quad (36)$$

Similar to [4]–[6], it is assumed that the channel coefficients are mutually independent. Based on the above channel model, we write the entries of the *diagonal* matrices $\mathbf{F}_k \in \mathbb{C}^{L\times L}$ and $\mathbf{D}_k \in \mathbb{C}^{L\times L}$, as well as the entries of the matrix $\mathbf{Q}_{j,k} \in \mathbb{C}^{L\times L}$, respectively, as [4]–[6]

$$[\mathbf{F}_k]_{m,m} = \bar{f}_{k,m}\bar{f}_{k,m}^H + \frac{\Upsilon_f}{1+\Upsilon_f},\ \forall k \in \mathcal{K},\ \forall m \in \mathcal{L} \quad (37)$$

$$[\mathbf{D}_k]_{m,m} = \bar{g}_{m,k}\bar{g}_{m,k}^H + \frac{\Upsilon_g}{1+\Upsilon_g},\ \forall k \in \mathcal{K},\ \forall m \in \mathcal{L} \quad (38)$$

$$[\mathbf{Q}_{j,k}]_{l,m} = \left(\bar{f}_{j,l}\bar{f}_{j,m}^H + \frac{\Upsilon_f}{1+\Upsilon_f}\delta(l-m)\right) \times \left(\bar{g}_{l,k}\bar{g}_{m,k}^H + \frac{\Upsilon_g}{1+\Upsilon_g}\delta(l-m)\right),\ \forall j, k \in \mathcal{K},\ \forall l, m \in \mathcal{L} \quad (39)$$

$$\text{with the Delta function: } \delta(t) \triangleq \begin{cases} 1, & \text{if } t = 0, \\ 0, & \text{otherwise.} \end{cases} \quad (40)$$

Throughout our simulations, identical minimum SINR requirements are chosen for all K peers, i.e., $\gamma_k^{(\min)} = \gamma_j^{(\min)},\ \forall j \neq k,\ j,\ k \in \mathcal{K}$. The noise variances at the relays and the destinations are normalized to unity, i.e., $\sigma_\eta^2 = \sigma_\nu^2 = 1$. Similar to [4]–[6], we choose the constants Υ_f and Υ_g as: $\Upsilon_f = -10$ dB and $\Upsilon_g = -10$ dB, respectively. All the simulation results are averaged over a certain number of channel realizations according to (33) and (34), as detailed in the following.

B. Performance Comparison With Existing Schemes

In Figs. 2 and 3, the total transmitted power $\psi(\mathbf{z})$ is depicted versus the minimum SINR requirements $\gamma_k^{(\min)},\ \forall k \in \mathcal{K}$ for the MUP2P relay networks consisting of $K = 3$ and $K = 4$ source-destination pairs, and $L = 4$ and $L = 8$ relays, respectively. For both Figs. 2 and 3, the maximum number of iterations is chosen as: $N_1^{(\max)} = N_2^{(\max)} = 12$. Note that, for each channel realization, both the convex problems (18) and (21) need to be solved at most $2N_2^{(\max)}$ times. All simulation results are averaged over 300 channel realizations for which feasible points of the DC program (14) were obtained by each of the three algorithms under consideration that were all initialized with the same random points.

From Figs. 2 and 3, we observe that the proposed CCCP based solution *on average* performs better in terms of total transmitted power than the alternating optimization method of [1], which

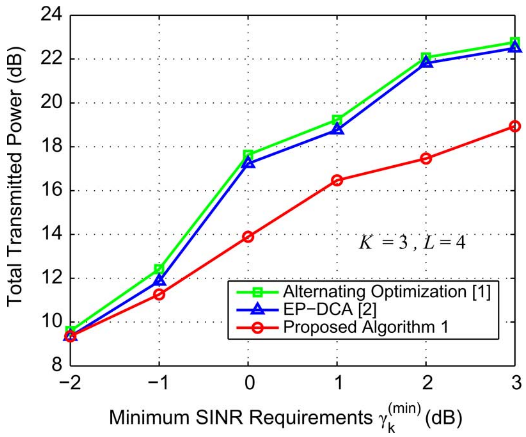


Fig. 2. Total transmitted power $\psi(\mathbf{z})$ (the y-axis) versus the minimum SINR requirements $\gamma_k^{(\min)}$, $\forall k \in \mathcal{K}$ (the x-axis), with $K = 3,\ L = 4$.

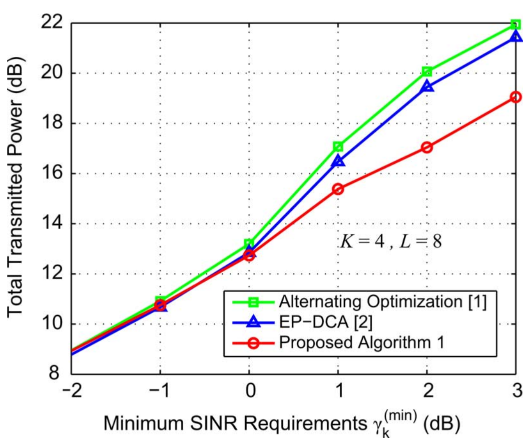


Fig. 3. Total transmitted power $\psi(\mathbf{z})$ (the y-axis) versus the minimum SINR requirements $\gamma_k^{(\min)}$, $\forall k \in \mathcal{K}$ (the x-axis), with $K = 4,\ L = 8$.

is because the source power allocations and the relay beamforming weights are not optimized *simultaneously* in the alternating optimization algorithm of [1], and the method of [1] yields solutions which are not KKT points or local minimum, while the proposed algorithm yields strictly local minima. On the other hand, the proposed CCCP based algorithm also performs better than the EP-DCA of [2], even though the penalty factor in the EP-DCA has been adjusted dynamically and gradually in all the simulations, i.e., the penalty factor Γ is chose as: $\Gamma = 3^n \Gamma_0$, with $\Gamma_0 = 20$ and n being the iterations number. The EP-DCA does not perform as well as the proposed Algorithm 1 due to the inherent difficulties in choosing the right penalty factor for which the best performance of the EP-DCA can be obtained [2]. In addition, we can observe from the figures that, when the number of relays increases, the reduction of the total power required to guarantee the predefined SINR requirement of each source-destination pair is significant. This is because as the number of relays increases, the system enjoys an increase in the number of relay antennas, as well as an increase in the degrees of freedom for optimizing the relay beamformer $\mathbf{w} \in \mathbb{C}^{L \times 1}$, and for distributing the power between the sources and the relays.

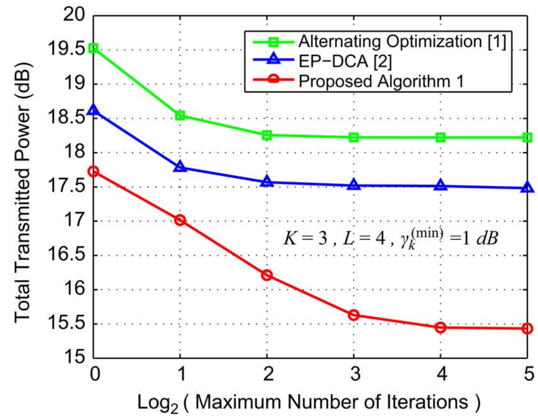


Fig. 4. Total transmitted power $\psi(\mathbf{z})$ (the y-axis) versus the number of iterations (the x-axis) in the proposed Algorithm 1, with $K = 3,\ L = 4$, and $\gamma_k^{(\min)} = 1$ dB, $\forall k \in \mathcal{K}$.

C. Convergence Illustration of the Proposed Solution

In Fig. 4, the convergence behavior of the proposed Algorithm 1, as well as the two reference algorithms, i.e., the alternating optimization scheme of [1] and the EP-DCA of [2], is illustrated for the MUP2P relay networks consisting of $K = 3$ source-destination pairs, and $L = 4$ relays. The minimum SINR requirements are set as: $\gamma_k^{(\min)} = 1$ dB, $\forall k \in \mathcal{K}$. The results are averaged over 300 channel realizations, which are obtained in the same way as in the previous subsection. Note that the main computational complexity of the proposed Algorithms 1 and 2 consists in solving the convex optimization problems (18) and (21), respectively.

As can be seen from Fig. 4, the proposed Algorithm 1 converges after approximately 16 iterations for any considered initial feasible point $\mathbf{z}^{(0)}$. In addition, it can be observed from the figure that as the maximum number of iterations increases, the performance of the algorithms under consideration does not further improve significantly and convergence is obtained after approximately 12 iterations.

D. Performance of the Proposed IFSA

In Figs. 5 and 6, the performance of the proposed IFSA in Algorithm 2 is illustrated for the MUP2P relay networks consisting of $K = 3$ and $K = 4$ source-destination pairs, and $L = 4$ and $L = 8$ relays, respectively. The maximum number of iterations is chosen as: $N_1^{(\max)} = 12$. The figure is generated in the following way. For any given value of SINR requirement $\gamma_k^{(\min)}$, the percentage of the successful cases is computed. Here and hereafter, by "successful case" we mean that, a feasible point of the DC program (14) can be found by a scheme (e.g., the EP-DCA of [2]) for a given channel realization, starting with a random *infeasible* point of the DC program (14). The results are averaged over 300 channel realizations.

As can be observed from Figs. 5 and 6, the proposed Algorithm 2 outperforms the alternating optimization method of [1] in terms of successful cases. This is because in the latter algorithm either the source power allocation $\mathbf{p}$ or the relay beamformer $\mathbf{w}$ is fixed, which results in smaller feasible sets than that

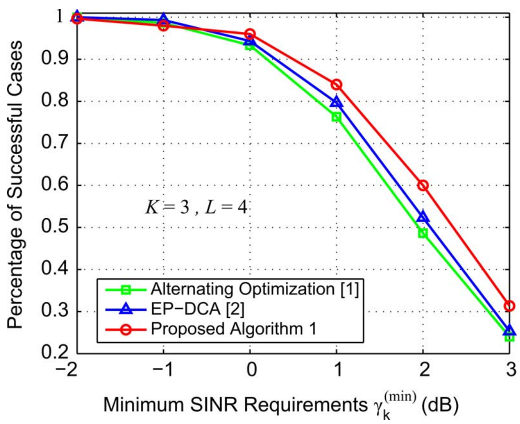


Fig. 5. Percentage of successful cases (the y-axis) versus minimum SINR requirements $\gamma_k^{(\min)}$, $\forall k \in \mathcal{K}$ (the x-axis), with $K = 3,\ L = 4$.

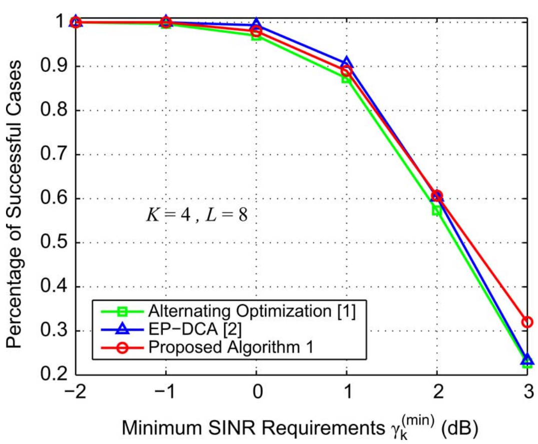


Fig. 6. Percentage of successful cases (the y-axis) versus minimum SINR requirements $\gamma_k^{(\min)}$, $\forall k \in \mathcal{K}$ (the x-axis), with $K = 4,\ L = 8$.

in the proposed Algorithm 2. Furthermore, the proposed Algorithm 2 also outperforms the EP-DCA of [2], which is because it is impossible to choose the right penalty factor that gives the best performance of the EP-DCA of [2].

E. Sensitivity of the Proposed IFSA w.r.t. Initial Points

In Fig. 7 the sensitivity of the proposed IFSA w.r.t. initial points is demonstrated for the MUP2P relay networks consisting of $K = 3$ source-destination pairs, and $L = 4$, $L = 5$, and $L = 6$ relays, respectively. The maximum number of iterations is chosen as $N_1^{(\max)} = 12$. The figure was generated as following. A number of 1000 channel realizations were drawn and 10 channel realizations (out of the 1000 channel realizations) have been selected in an arbitrary manner for which feasible points of the DC program (14) could be found by the proposed Algorithm 2. The selected channel realizations were then used to test the sensitivity of the proposed IFSA. That is, for each of the 10 selected channel realizations, the proposed IFSA in Algorithm 2 has been initialized with 300 random infeasible initial points, and the percentage of the *feasible* initializations (i.e., for which the proposed IFSA yields feasible points of the DC program (14)) out of the 300 infeasible initializations has been computed. This procedure has been carried out with each of the 10 selected channel realizations and for each SINR requirement $\gamma_k^{(\min)}$. The results have been averaged over the 10 selected channel realizations.

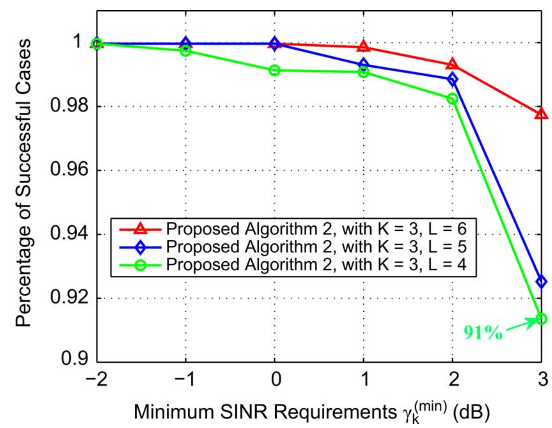


Fig. 7. Percentage of successful cases (the y-axis) versus minimum SINR requirements $\gamma_k^{(\min)}$, $\forall k \in \mathcal{K}$ (the x-axis). The results are averaged over 300 random *infeasible* initial points for an arbitrarily chosen channel realization.

It can be seen from Fig. 7 that, the proposed IFSA in Algorithm 2 is *not* sensitive w.r.t. the random infeasible initial points. In other words, as long as at least one feasible point of the DC program (14) is obtained from the proposed IFSA for a given channel realization, starting with any arbitrary infeasible initial points, the proposed IFSA can successfully obtain a feasible point of the DC program (14) in most of the times (e.g., in more than 98% of the cases for $L \geq 6$ relays) for that channel realization.

VI. Conclusion

We have investigated the QoS based PADB problem in MUP2P relay networks. By utilizing an appropriate transformation of variables, the formulated PADB problem has been converted into a DC program, and a CCCP based low-complexity solution has been proposed to solve the DC program. The proposed solution represents a novel modification of the conventional CCCP [21]. In addition, we have proposed an IFSA to find a feasible point of the DC program (14) that is used to initialize the CCCP. The proposed two-stage algorithm, Algorithm 2 and Algorithm 1, can also be applied to efficiently solve other DC programming problems. Analytic studies show that the proposed solution converges to a local minimum of the DC program (14), and thus also a local minimum of the PADB problem (6). Numerical results confirm that the proposed solution outperforms (in terms of total transmitted power) the alternating optimization method of [1] and the conventional EP-DCA [2]. Numerical results also show that the proposed IFSA outperforms the alternating optimization algorithm of [1] in finding feasible points of the DC program (14), and the proposed IFSA performs similarly to the EP-DCA of [2] regarding feasibility search.

APPENDIX A
PROOF OF LEMMA 1

The proof is carried out by examining the function $\psi(\mathbf{q}, \mathbf{w})$ term by term.

1) The term $\frac{1}{q_k}$ is a strictly convex function in $q_k \in \mathbb{R}_+$, $\forall k \in \mathcal{K}$ [24, Sec. 3.1]. Hence, the term $\sum_{k=1}^{K} \frac{1}{q_k}$ is a *strictly* convex function in $\mathbf{q} \in \mathbb{R}_+^{K\times 1}$ [24, Sec. 3.2].
2) As a direct application of Lemma 1, the term $\frac{\mathbf{w}^H \mathbf{F}_k \mathbf{w}}{q_k}$ is a convex function in $(q_k, \mathbf{w}) \in \mathbb{R}_+ \otimes \mathbb{C}^{L\times 1}$, $\forall k \in \mathcal{K}$. Hence, the term $\sum_{k=1}^{K} \frac{\mathbf{w}^H \mathbf{F}_k \mathbf{w}}{q_k}$ is a convex function in $(\mathbf{q}, \mathbf{w}) \in \mathbb{R}_+^{K\times 1} \otimes \mathbb{C}^{L\times 1}$ [24, Sec. 3.2].
3) The term $\mathbf{w}^H \mathbf{w}$ is a *strictly* convex function in $\mathbf{w} \in \mathbb{C}^{L\times 1}$ [24, Sec. 4.2].

As a result, the function $\psi(\mathbf{q}, \mathbf{w}) \triangleq \sum_{k=1}^{K} \frac{1}{q_k} + \left(\sum_{k=1}^{K} \frac{\mathbf{w}^H \mathbf{F}_k \mathbf{w}}{q_k} + \sigma_\eta^2 \mathbf{w}^H \mathbf{w}\right)$ is a *strictly* convex function in the variables $(\mathbf{q}, \mathbf{w}) \in \mathbb{R}_+^{K\times 1} \otimes \mathbb{C}^{L\times 1}$ [24, Sec. 4.2].

APPENDIX B
PROOF OF LEMMA 2

Since the point $\mathbf{z}^* \triangleq \left[(\mathbf{q}^*)^T, (\mathbf{w}^*)^T\right]^T$ is the limit point of the sequence $\{\mathbf{z}^{(n)}\}$, by definition, the point $\mathbf{z}^*$ is a feasible point for the convex optimization problem (23), and no strictly better solution exists. On the other hand, since the objective function $\psi(\mathbf{z})$ of problem (23) is strictly convex in the variable $\mathbf{z} \in \mathbb{R}_+^{K\times 1} \otimes \mathbb{C}^{L\times 1}$, the solution of problem (23) is unique [24, Sec. 4.2]. Therefore, the limit point $\mathbf{z}^*$ is the solution of problem (23).

The second part of the Lemma is proved by contradiction. Suppose that the k^{th} constraint is not active, i.e., $\hat{\varphi}_k(\mathbf{z}^*) < 0$, then from observation O3) we know that we can scale up the variable q_k^* to make the constraint active, without violating the other SINR constraints in (23b) and therefore reduce the objective function $\psi(\mathbf{z}^*)$ (see observation O1)), which contradicts the optimality of the point $\mathbf{z}^*$. Hence, we conclude that all constraints in (23b) are active at the point $\mathbf{z}^*$.

APPENDIX C
PROOF OF THEOREM 1

In this proof, we simply write $\mathbf{z}^*\left(\mathbf{z}^{(0)}\right)$ as $\mathbf{z}^*$ for brevity. It has already been shown in Section IV-C that the limit point $\mathbf{z}^*$ is a *stationary point* (also known as KKT point) of the DC program (14). In general, a stationary point could be a saddle point, a local maximum, or a local minimum of a nonlinear program [30, Lecture 26], [31, Sec. 18.2]. In the following, we prove that all stationary points are local minima for the DC program (14).

First, by the definition of saddle points,[11] all stationary points of the DC program (14) *cannot* be saddle points, since the objective function $\psi(\mathbf{z})$ of the DC program (14) is twice-continuously differentiable and is a *strictly convex* function in the variable $\mathbf{z} \in \mathbb{R}_+^{K\times 1} \otimes \mathbb{R}^{L\times 1}$ [24, Sec. 3.1].

[11] A saddle point of a smooth function is a stationary point such that the surface of the function in the neighborhood of that point is *not* entirely on any side of the *tangent space* at that point [32, Sec. 5.2].

Second, all stationary points of the DC program (14) *cannot* be local maxima, which is proved by contradiction. Assume that the limit point $\mathbf{z}^*$ is a local maximum, by definition [24, Sec. 4.1], there exists a constant $\delta > 0$, such that for any feasible point $\mathbf{z}$ of the DC program (14), which in addition satisfies $\|\mathbf{z} - \mathbf{z}^*\| \leq \delta$, we have $\psi(\mathbf{z}) \leq \psi(\mathbf{z}^*)$.

Recall that $\mathbf{z}^* \triangleq \left[(\mathbf{q}^*)^T, (\mathbf{w}^*)^T\right]^T$, and from Lemma 2 we know that $\mathbf{w}^* \neq \mathbf{0}$. Define

$$\varepsilon \triangleq \frac{\delta}{\|\mathbf{w}^*\|} > 0, \text{ and } \breve{\mathbf{w}}^* \triangleq (1+\varepsilon)\mathbf{w}^*. \tag{41}$$

Then, from (11)–(13), we can see that the point $\breve{\mathbf{z}}^* \triangleq \left[(\mathbf{q}^*)^T, (\breve{\mathbf{w}}^*)^T\right]^T$ is also a feasible point of the DC program (14), and $\|\breve{\mathbf{z}}^* - \mathbf{z}^*\| \leq \delta$. Therefore, we have

$$\psi(\breve{\mathbf{z}}^*) \leq \psi(\mathbf{z}^*). \tag{42}$$

On the other hand, from (9) we know that $\psi(\breve{\mathbf{z}}^*) > \psi(\mathbf{z}^*)$, which contradicts with (42). Hence, the point $\mathbf{z}^*$ *cannot* be a local maximum.

As a result, all stationary points must be local minima for the DC program (14). Hence, the limit point $\mathbf{z}^*$ is a local minimum.

REFERENCES

[1] Y. Jin and Y. D. Zhang, "Joint source and relay power optimization in multiuser cooperative wireless networks," in *Proc. 4th Int. Symp. Commun., Control, Signal Process.*, Limassol, Cyprus, Mar. 3–5, 2010, pp. 1–4.

[2] L. T. H. An, "D.C. programming for solving a class of global optimization problems via reformulation by exact penalty," *Global Optim. Constraint Satisfact.*, vol. 2861, pp. 87–101, 2003.

[3] Z. Ding, W. H. Chin, and K. K. Leung, "Distributed beamforming and power allocation for cooperative networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 5, pp. 1817–1822, May 2008.

[4] V. Havary-Nassab, S. Shahbazpanahi, A. Grami, and Z.-Q. Luo, "Distributed beamforming for relay networks based on second-order statistics of the channel state information," *IEEE Trans. Signal Process.*, vol. 56, no. 9, pp. 4306–4316, Sep. 2008.

[5] S. Fazeli-Dehkordy, S. Shahbazpanahi, and S. Gazor, "Multiple peer-to-peer communications using a network of relays," *IEEE Trans. Signal Process.*, vol. 57, no. 8, pp. 3053–3062, Aug. 2009.

[6] J. Li, A. P. Petropulu, and H. V. Poor, "Cooperative transmission for relay networks based on second-order statistics of channel state information," *IEEE Trans. Signal Process.*, vol. 59, no. 3, pp. 1280–1291, Mar. 2011.

[7] Y. Jing and H. Jafarkhani, "Network beamforming using relays with perfect channel information," *IEEE Trans. Inf. Theory*, vol. 55, no. 6, pp. 2499–2517, Jun. 2009.

[8] H. Chen, A. B. Gershman, and S. Shahbazpanahi, "Distributed peer-to-peer beamforming for multiuser relay networks," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Apr. 2009, pp. 2265–2268.

[9] A. H. Phan, H. D. Tuan, H. H. Kha, and H. H. Nguyen, "Nonsmooth optimization-based beamforming in multiuser wireless relay networks," in *Proc. 4th Int. Conf. Signal Process. Commun. Syst. (ICSPCS)*, Dec. 2010, pp. 1–4.

[10] N. Bornhorst, M. Pesavento, and A. B. Gershman, "Distributed beamforming for multi-group multicasting relay networks," *IEEE Trans. Signal Process.*, vol. 60, no. 1, pp. 221–232, 2011.

[11] X. Li, Y. D. Zhang, and M. G. Amin, "Joint source power scheduling and distributed relay beamforming in multiuser cooperative wireless networks," in *Proc. IEEE Global Telecommun. Conf.*, Dec. 2009, pp. 1–6.

[12] Y. D. Zhang, X. Li, and M. Amin, "Distributed beamforming in multiuser cooperative wireless networks," in *Proc. 4th Int. Conf. Commun. Netw. China*, Aug. 2009, pp. 1–5.

[13] X. Li, Y. D. Zhang, and M. G. Amin, "Joint optimization of source power allocation and relay beamforming in multiuser cooperative wireless networks," *Mobile Netw. Appl.*, vol. 16, no. 5, pp. 562–575, 2011.

[14] A. B. Gershman, N. D. Sidiropoulos, S. Shahbazpanahi, M. Bengtsson, and B. Ottersten, "Convex optimization-based beamforming: From receive to transmit and network designs," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 62–75, May 2010.

[15] K. Phan, T. Le-Ngoc, S. A. Vorobyov, and C. Tellambura, "Power allocation in wireless multi-user relay networks," *IEEE Trans. Wireless Commun.*, vol. 8, no. 5, pp. 2535–2545, May 2009.

[16] K. T. Phan, L. B. Le, S. A. Vorobyov, and T. Le-Ngoc, "Power allocation and admission control in multiuser relay networks via convex programming: centralized and distributed schemes," *EURASIP J. Wireless Commun. Netw.*, vol. 2009, no. 1, pp. 1–12, Feb. 2009.

[17] X. Gong, S. A. Vorobyov, and C. Tellambura, "Joint bandwidth and power allocation with admission control in wireless multi-user networks with and without relaying," *IEEE Trans. Signal Process.*, vol. 59, no. 4, pp. 1801–1813, Apr. 2011.

[18] P. D. Tao and L. T. H. An, "D. C. programming: Theory, algorithms and applications," *ACTA Math. Vietnamica*, vol. 22, no. 1, pp. 289–355, 1997.

[19] R. Horst and N. V. Thoai, "Dc programming: Overview," *J. Optim. Theory Appl.*, vol. 103, no. 1, pp. 1–43, Oct. 1999.

[20] A. L. Yuille and A. Rangarajan, "The concave-convex procedure," *Neural Comput.*, vol. 15, no. 4, pp. 915–936, Apr. 2003.

[21] A. J. Smola, S. V. N. Vishwanathan, and T. Hofmann, "Kernel methods for missing variables," in *Proc. 10th Int. Workshop Artif. Intell. Stat.*, Mar. 2005.

[22] B. K. Sriperumbudur and G. R. G. Lanckriet, "On the convergence of the concave-convex procedure," *Neural Inf. Process. Syst.*, pp. 1–9, 2009.

[23] Z.-Q. Luo, W.-K. Ma, A.-C. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 20–34, May 2010.

[24] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2009.

[25] T. N. Bekjan, "On joint convexity of trace functions," *Linear Algebra Appl.*, vol. 390, pp. 321–327, 2004.

[26] D. H. Brandwood, "A complex gradient operator and its application in adaptive array theory," *Proc. Inst. Electr. Eng.—Commun., Radar, Signal Process.*, vol. 130, no. 1, pp. 11–16, 1983.

[27] H. Li and T. Adali, "Complex-valued adaptive signal processing using nonlinear functions," *EURASIP J. Adv. Signal Process.*, vol. 2008, pp. 1–9, 2008.

[28] S. Haykin, *Adaptive Filter Theory*, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 1996.

[29] C. D'Ambrosio, A. Frangioni, L. Liberti, and A. Lodi, Mathematical Programming, to appear "A storm of feasibility pumps for non-convex MINLP," 2011 [Online]. Available: http://www.di.unipi.it/$\thicksim$frangio/papers/fpminlp.pdf

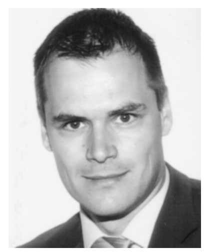
[30] A. Nedich, Operations Research Methods, 2007 [Online]. Available: https://netfiles.uiuc.edu/angelia/www/operations_research_GE330.htm

[31] H. A. Taha, *Operations Research: An Introduction*, 8th ed. Englewood Cliffs, NJ: Prentice-Hall, 2007.

[32] M. J. Osborne, Mathematical Methods for Economic Theory: A Tutorial, 2007 [Online]. Available: http://www.economics.utoronto.ca/osborne/MathTutorial/

Yong Cheng (S'09) received the B.Eng. (First Hons.) from Zhejiang University, Hangzhou, China, in 2006 and the M.Phil. degree from the Hong Kong University of Science and Technology (HKUST), Hong Kong, in 2010.

He is currently working towards the Ph.D. degree at the Communication Systems Group, the Department of Electrical Engineering and Information Technology, Darmstadt University of Technology, Darmstadt, Germany. His research interests include network optimization, transmit beamforming, and resource allocation in cooperative MIMO-OFDM systems, distributed beamforming in peer-to-peer OFDM relay networks, and mixed-integer nonlinear programming with applications in signal processing and wireless communications.



Marius Pesavento (M'00) received the Dipl.-Ing. degree from Ruhr-Universität Bochum, Germany, in 1999 and the M.Eng. degree from McMaster University, Hamilton, ON, Canada, in 2000, and the Dr.-Ing. degree in electrical engineering from Ruhr-Universität Bochum, Germany, in 2005.

Between 2005 and 2007, he was a Research Engineer at FAG Industrial Services GmbH, Aachen, Germany. From 2007 to 2009, he was the Director of the Signal Processing Section at mimoOn GmbH, Duisburg, Germany. In 2010, he became a Professor for Robust Signal Processing at the Department of Electrical Engineering and Information Technology, Darmstadt University of Technology, Darmstadt, Germany, where he is currently the Head of the Communication Systems Group. His research interests are in the area of robust signal processing and adaptive beamforming, high-resolution sensor array processing, transceiver design for cognitive radio systems, cooperative communications in relay networks, MIMO and multiantenna communications, space-time coding, multiuser and multicarrier wireless communication systems, convex optimization for signal processing and communications, statistical signal processing, spectral analysis, parameter estimation and detection theory.

Dr. Pesavento was a recipient of the 2003 ITG/VDE Best Paper Award, the 2005 Young Author Best Paper Award of the IEEE TRANSACTIONS ON SIGNAL PROCESSING, and the 2010 Best Paper Award of the CROWNCOM conference. He is a member of the Editorial board of the *EURASIP Signal Processing Journal*, an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING, and a member of the Sensor Array and Multichannel (SAM) Technical Committee of the IEEE Signal Processing Society (SPS).