Jointly Optimal Source Power Control and Relay Matrix Design in Multipoint-to-Multipoint Cooperative Communication Networks

Keyvan Zarifi, *Member, IEEE*, Ali Ghrayeb, *Senior Member, IEEE*, and Sofiène Affes, *Senior Member, IEEE*

***Abstract*—A cooperative communication network is considered wherein L sources aim to transmit to their designated destinations through the use of a multiple-antenna relay. All sources transmit to the relay in a shared channel in the first transmission phase. Then, the relay linearly processes its received signal vector using L relaying matrices and retransmits the resultant signals towards the destinations in dedicated channels in the second transmission phase. The goal is to jointly optimize the sources' transmit powers and the relaying matrices such that the worst normalized signal-to-interference-plus-noise ratio (SINR) among all L destinations is maximized while the relays' transmit powers in the dedicated channels as well as the sources' individual and total transmit powers do not exceed predetermined thresholds. It is shown that the jointly optimal sources' transmit powers and the relaying matrices are the solutions to an optimization problem with a nonconvex objective function and multiple nonconvex constraints. To solve this problem, it is first proved that all normalized SINRs are equal at the optimal point of the objective function. Then, the optimization problem is transformed through multiple stages into an equivalent problem that is amenable to an iterative solution. Finally, an efficient iterative algorithm is developed that offers the jointly optimal sources' transmit powers and the relaying matrices. An extension to the above problem is then studied in the case when the cooperative communication network acts as a cognitive system that is expected to operate such that its interfering effect on the primary users is below some admissibility thresholds. In such a case, the sources' and relay's transmit powers should further satisfy some additional constraints that compel a new technique to tackle the problem of the joint optimization of the sources' transmit powers and the relaying matrices. An iterative solution to the latter problem is also proposed and the efficiency and the high rate of convergence of the proposed iterative algorithms in both the original and the cognitive cases are verified by simulation examples.**

***Index Terms*—Cooperative communication, joint source and relay design, multi-source multi-destination network.**

Manuscript received July 07, 2010; revised January 30, 2011; accepted May 18, 2011. Date of publication June 02, 2011; date of current version August 10, 2011. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Martin Schubert. This work was supported in part by NSERC, a Canada research chair in wireless communications, Ericsson Canada, and PROMPT.

K. Zarifi was with the Institut National de la Recherche Scientifique-Énergie, Matériaux, et Télécommunications (INRS-EMT), Université du Québec, Montreal, QC, Canada. He is now with Huawei Technologies, Kanata, ON K2K 3C9, Canada (e-mail: zarifi@ieee.org).

A. Ghrayeb is with Concordia University, Montreal, QC H3G 1M8, Canada (e-mail: aghrayeb@ece.concordia.ca).

S. Affes is with INRS-EMT, Université du Québec, Montreal, QC H5A 1K6, Canada (e-mail: affes@emt.inrs.ca).

Digital Object Identifier 10.1109/TSP.2011.2158426

I. Introduction

USING relay transceivers that send properly-processed versions of their received signals through the wireless medium, cooperative communication techniques can substantially increase the link reliability, the transmission coverage, and the capacity of wireless networks [1]–[5]. While such relay-assisted schemes were originally tailored for a point-to-point communication between an isolated source-destination pair, the growing interest in the broad potential applications of wireless ad hoc networks has motivated increasing research efforts in developing efficient cooperative communication schemes for multipoint-to-multipoint wireless networks wherein multiple sources aim to transmit their signals to their designated destinations through possibly a shared channel [6]–[14]. A major challenge in multipoint-to-multipoint cooperative networks is to devise an energy-efficient communication scheme that establishes channel links between all source–destination pairs via possibly multiple relaying antennas and simultaneously provides an acceptable level of quality of service to every destination by either mitigating or preventing all cross-link interferences that may be caused by irrelevant sources. Several attempts have been made to tackle this challenge: Aiming to increase the signal-to-interference-plus-noise ratios (SINRs) at the destinations by minimizing the relays' forwarded interference and noise power, a linearly-constrained minimum-variance relay beamforming approach is developed in [7]. An energy-efficient distributed relay beamforming technique that guarantees target SINRs at the destinations is designed in [8] and is generalized in [9] to the case that the relays share some local information. To mitigate the cross-link interference effect, zero-forcing and minimum-mean-square-error-based signal processing techniques are developed in [10] and [11], and to counter the degrading effect of the imperfect channel knowledge, robust signal processing techniques are proposed in [12]. The authors in [13] and [14] assume that all network transmissions are carried out in interference-free orthogonal channels and develop efficient resource allocation techniques at the sources and/or relays to optimize some network level performance metrics.

When sources use a shared channel to transmit their data, the presence of cross-link interference makes it more challenging to achieve the desired quality of service requirements at all destinations. In such cases, the cooperative schemes that aim to optimize the communication network only at the relaying phase [7]–[9], [12] or using separate designs at the transmission and relaying stages [10], [11] may show an inadequate performance. Motivated by the above fact, we present a joint design at the transmission and relaying stages in a multipoint-to-multipoint cooperative network by jointly optimizing the transmit powers at single-antenna sources and the linear signal processing technique at the multiple-antenna relaying terminal. Despite the significant contributions to the literature of multipoint-to-multipoint cooperative communication networks, to the best of our knowledge, none of the existing works has considered the problem of joint power control at the sources and the signal processing at the relay for a multipoint-to-multipoint cooperative communication network.

In the first transmission phase, L sources transmit their signals through a shared channel and K relay antennas receive different noisy and faded mixtures of the transmitted signals. In the second phase, the relay linearly processes its K-dimensional received vector by multiplying it with L relaying matrices and then transmits the resultant signal vectors to the L intended destinations through dedicated channels.[1] As the sources use a shared channel in the first transmission phase, the cross-link interference effects degrade the performance at the destinations. Aiming to mitigate these effects while preserving fairness among all destinations, jointly optimal sources' transmit powers and relaying matrices are sought that maximize the worst normalized SINR among all destinations subject to upper-bounds on the sources' individual and total transmit powers and the relay's transmit power at all L dedicated channels. The optimal variables are shown to be the solution to a complicated optimization problem with a nonconvex objective function and multiple nonconvex constraints. As the first step to solve this optimization problem, it is shown that the optimal sources' transmit powers for any given set of relaying matrices are those that balance all normalized SINRs. This property is then used to transform the optimization problem through multiple stages into an equivalent form that lends itself to an iterative solution. An efficient algorithm is then developed that obtains the jointly optimal set of sources' transmit powers and the relaying matrices through an alternating optimization technique. The per-iteration computational complexity order of the proposed algorithm is derived and it is shown that while the studied problem here includes the problem investigated in [16] as a special case, the solution algorithm proposed in this paper enjoys a less per-iteration computational complexity order than that of [16].

The studied multipoint-to-multipoint cooperative communication network is then treated as a cognitive network that operates in the presence of a primary system. In such a case, the optimal sources' transmit powers and the relaying matrices should further satisfy additional constraints on the incurred interference to the primary users. These additional constraints pose a significant challenge to obtaining the jointly optimal transmission strategies at the sources and the relay. In particular, unlike in the original noncognitive case, the optimal relaying matrices for a given set of sources' transmit powers do not admit closed-form representations and are shown to be the solutions to a set of quasi-convex optimization problems. An efficient technique to solve the latter quasi-convex optimization problems is proposed and is then used in an iterative alternating optimization-based algorithm that jointly derives the optimal sources' transmit powers and the relaying matrices. Finally, the rapid convergence property of the proposed iterative algorithms and the validity of the analytical results are verified using a number of simulation examples. We are currently investigating related joint optimization problems in the case that the relay-destinations communication is also carried out in a shared channel. The solution to the latter problem will be disclosed in a future publication.

The rest of this paper is organized as follows. Section II presents the signal model and the problem formulation and Section III obtains the optimal sources' transmit powers for a given set of relaying matrices. Section IV derives the jointly optimal sources' transmit powers and relaying matrices. Section V tackles the above joint optimization problem in the case that the sources and relay are cognitive terminals. Section VI presents the simulation results and Section VII draws the concluding remarks.

Notation: Uppercase and lowercase bold letters denote matrices and vectors, respectively. $|\cdot|$ is the absolute value, $\|\cdot\|$ is the 2-norm of a vector, and $\mathrm{tr}(\cdot)$ is the trace of a matrix. $(\cdot)^T$, $(\cdot)^*$, and $(\cdot)^H$ denote the transpose, the conjugate, and the Hermitian transpose, respectively. $\mathrm{E}\{\cdot\}$ stands for the statistical expectation, $\Im$ is the imaginary part, and $\lceil a \rceil$ is the smallest integer larger or equal to a. $\mathbf{1}$ is the vector with all entries equal to 1 and $\mathbf{e}_i$ is the vector with 1 at the ith position and zeros elsewhere. $\mathbf{I}_K$ is the $K \times K$ identity matrix. $[\cdot]_l$ and $[\cdot]_{nl}$ stand for the lth entry of a vector and the entry at the nth row and the lth column of a matrix, respectively. $\lambda_{\max}(\cdot)$ is the maximum modulus eigenvalue, $\Omega(\cdot)$ is the eigenvector associated with the maximum modulus eigenvalue normalized such that its last entry is 1, and $\rho(\cdot)$ is the spectral radius of a matrix. $\mathbf{D}(\mathbf{a})$ is a diagonal matrix whose diagonal elements are the entries of $\mathbf{a}$. $\mathbf{a}_{\bar{l}}$ is the $(L-1)\times 1$ vector obtained by removing the lth entry of the $L \times 1$ vector $\mathbf{a}$. $\mathbf{A}_{\bullet\bar{l}}$ is the $N \times (L-1)$ matrix obtained by removing the lth column of the $N \times L$ matrix $\mathbf{A}$ and $\mathrm{vec}(\mathbf{A})$ is the $NL \times 1$ vector obtained by stacking the columns of $\mathbf{A}$ on top of one another. $\otimes$ is the Kronecker product. $\in$ denotes the membership in a set while $\mathrm{Card}(\cdot)$ is the cardinality of a set. $\cup$ and $\subseteq$ denote the union of two sets and the subset, respectively. In this paper, $>$ and $\geq$ are elementwise inequalities, that is, if $\mathbf{A} \geq \mathbf{B}$, then $[\mathbf{A}]_{nl} \geq [\mathbf{B}]_{nl}$ for all n and l. $\mathbf{a} \gneqq \mathbf{b}$ means that $\mathbf{a} \leq \mathbf{b}$ with $[\mathbf{a}]_l < [\mathbf{b}]_l$ for at least one l.

II. Signal Model and Problem Formulation

Consider the cooperative communication network depicted in Fig. 1 with L pairs of single-antenna terminals $(S_l, D_l), l = 1,\ldots,L$ where each source S_l aims to transmit its data to its designated destination D_l. Assume that all direct source-destination communication links $S_iD_j, i,j = 1,\ldots,L$ are negligibly weak due to, for instance, large distances between the sources and the destinations and the required communication links between S_l and $D_l, l = 1,\ldots,L$ are established via a K-antenna relay that employs the following two-phase amplify-and-forward (AF) communication protocol: In the first phase,

[1]The joint source power control and relay matrix design in multipoint-to-multipoint cooperative communication networks with shared relay-destinations channels is studied in [15].

all sources broadcast their signals and the relay antennas receive different noisy and faded mixtures of the transmitted signals. In the second phase, transmissions from the relay antennas to the destinations are carried out using L orthogonal channels each of which dedicated to a single destination.[2] To transmit in the lth channel, the relay multiplies its K-dimensional received signal vector by a relaying matrix $\mathbf{W}^{(l)^*}$ and forwards the result en-route to D_l. This procedure continues in all L channels. Throughout this work, it is assumed that transmitted signals from the sources as well as noises at the destinations and the relay are zero-mean and mutually statistically independent and all source-relay and relay-destination channel links are quasi-static flat fading. Let s_l denote the transmitted signal from S_l with $\mathrm{E}\{|s_l|^2\} = p_l$ and assume that we must have $p_l \leq P_l, l = 1, \ldots, L$ and $\sum_{l=1}^{L} p_l \leq P_{L+1}$ where the first L constraints may be due to the sources' limited power supply and the last constraint may be imposed by regulations. Note that $P_{L+1} < \sum_{l=1}^{L} P_l$, as otherwise the last constraint is trivially satisfied. The above power constraints can be represented in a more compact form as

$$\mathbf{u}_l^T \mathbf{p} \leq P_l \quad l = 1, \ldots, L+1 \tag{1}$$

where $\mathbf{u}_l \triangleq \mathbf{e}_l, l = 1, \ldots, L$ and $\mathbf{u}_{L+1} \triangleq \mathbf{1}$. Let g_{kl} represent the channel gain from S_l to the kth relay antenna. Introducing $\mathbf{s} \triangleq [s_1 \ldots s_L]^T$, $\mathbf{g}_l \triangleq [g_{1l} \ldots g_{Kl}]^T$, and $\mathbf{G} \triangleq [\mathbf{g}_1 \ldots \mathbf{g}_L]$, the relay's received signal vector is given by

$$\mathbf{y} = \mathbf{G}\mathbf{s} + \mathbf{v} \tag{2}$$

where $\mathbf{v} \triangleq [v_1 \ldots v_K]^T$ is the noise vector at the relay with $\mathrm{E}\{|v_k|^2\} = \sigma_{v_k}^2$. The signal transmitted from the relay in the lth channel can then be represented as

$$\mathbf{x}^{(l)} = \mathbf{W}^{(l)^*}\mathbf{y}. \tag{3}$$

Using (2) and (3) and T1.1 and T1.2 from Table I, the relay's transmit power in the lth channel may be computed as

$$\begin{aligned} p_r^{(l)} &= \mathrm{E}\left\{\mathbf{x}^{(l)^H}\mathbf{x}^{(l)}\right\} \\ &= \mathrm{tr}\left(\mathbf{W}^{(l)^*}\left(\mathbf{G}\mathbf{D}(\mathbf{p})\mathbf{G}^H + \boldsymbol{\Sigma}_v\right)\mathbf{W}^{(l)^T}\right) \\ &= \mathbf{w}^{(l)^H}\boldsymbol{\Xi}(\mathbf{p})\mathbf{w}^{(l)} \end{aligned} \tag{4}$$

where $\mathbf{w}^{(l)} \triangleq \mathrm{vec}(\mathbf{W}^{(l)})$ and $\boldsymbol{\Xi}(\mathbf{p}) \triangleq (\mathbf{G}\mathbf{D}(\mathbf{p})\mathbf{G}^H + \boldsymbol{\Sigma}_v) \otimes \mathbf{I}_K$ with $\mathbf{p} \triangleq [p_1 \ldots p_L]^T$ and $\boldsymbol{\Sigma}_v \triangleq \mathrm{E}\{\mathbf{v}\mathbf{v}^H\} = \mathbf{D}([\sigma_{v_1}^2 \ldots \sigma_{v_K}^2]^T)$. Due to regulations, it is also required that

$$p_r^{(l)} \leq P^{(l)} \quad l = 1, \ldots, L \tag{5}$$

where $P^{(l)}$ is a given upper-bound. Let $h_k^{(l)}$ denote the channel gain from the kth relay antenna to D_l in the lth channel. Introducing $\mathbf{h}^{(l)} \triangleq [h_1^{(l)} \ldots h_K^{(l)}]^T$, the received signal at D_l in its dedicated channel is

$$z_l = \mathbf{h}^{(l)^T}\mathbf{x}^{(l)} + n_l \quad l = 1, \ldots, L \tag{6}$$

where n_l is noise at D_l. Using (2) and (3) in (6), we have

$$z_l = \mathbf{h}^{(l)^T}\mathbf{W}^{(l)^*}\mathbf{G}\mathbf{s} + \mathbf{h}^{(l)^T}\mathbf{W}^{(l)^*}\mathbf{v} + n_l \quad l = 1, \ldots, L. \tag{7}$$

Let us further introduce $\mathbf{f}_k^{(l)} \triangleq \mathbf{g}_k \otimes \mathbf{h}^{(l)}$. Using T1.3 and T1.4 from Table I, (7) may be equivalently represented as

$$\begin{aligned} z_l = \mathbf{w}^{(l)^H}\mathbf{f}_l^{(l)} s_l &+ \sum_{\substack{k=1 \\ k \neq l}}^{L} \mathbf{w}^{(l)^H}\mathbf{f}_k^{(l)} s_k \\ &+ \mathbf{w}^{(l)^H}\left(\mathbf{v} \otimes \mathbf{h}^{(l)}\right) + n_l, \end{aligned} \tag{8}$$

for $l = 1, \ldots, L$. Note that the first and the second terms in (8) are, respectively, the desired signal and the interference components while the sum of the third and the fourth terms in (8) constitute the aggregate noise component of the signal received at D_l. Let $\sigma_{n_l}^2 = \mathrm{E}\{|n_l|^2\}$. It is straightforward to show from (8) and T1.5 in Table I that, for $l = 1, \ldots, L$, the SINR at D_l is

$$\begin{aligned} &\eta_l\left(\mathbf{w}^{(l)}, \mathbf{p}\right) \\ &= \frac{p_l \left|\mathbf{w}^{(l)^H}\mathbf{f}_l^{(l)}\right|^2}{\sum_{\substack{k=1 \\ k \neq l}}^{L} p_k \left|\mathbf{w}^{(l)^H}\mathbf{f}_k^{(l)}\right|^2 + \mathbf{w}^{(l)^H}\boldsymbol{\Gamma}_v^{(l)}\mathbf{w}^{(l)} + \sigma_{n_l}^2} \end{aligned} \tag{9}$$

where $\boldsymbol{\Gamma}_v^{(l)} \triangleq \boldsymbol{\Sigma}_v \otimes \mathbf{h}^{(l)}\mathbf{h}^{(l)^H}$. In this work,

$$\bar{\eta}_l\left(\mathbf{w}^{(l)}, \mathbf{p}\right) \triangleq \frac{\eta_l\left(\mathbf{w}^{(l)}, \mathbf{p}\right)}{\gamma_l} \quad l = 1, \ldots, L \tag{10}$$

is used as the links' performance measure where γ_l is a normalization factor that may be selected proportional to the target SINR at D_l. The goal is to determine the jointly optimal power vector $\mathbf{p}_o$ and relaying matrices $\mathbf{W}_o^{(l)}$ that maximize the minimum $\bar{\eta}_l(\mathbf{w}^{(l)}, \mathbf{p})$ subject to the $2L + 1$ constraints imposed by (1) and (5). A max-min (normalized) SINR optimization strategy is typically used when the intent is to preserve fairness among multiple communication links [16], [18]–[20]. The problem of interest may be more formally presented as follows:

$$\max_{\mathbf{W}, \mathbf{p}} \quad \min_{1 \leq l \leq L} \bar{\eta}_l\left(\mathbf{w}^{(l)}, \mathbf{p}\right) \tag{11a}$$

$$\text{subject to} \quad \mathbf{u}_l^T \mathbf{p} \leq P_l \quad l = 1, \ldots, L+1 \tag{11b}$$

$$\mathbf{w}^{(l)^H}\boldsymbol{\Xi}(\mathbf{p})\mathbf{w}^{(l)} \leq P^{(l)} \quad l = 1, \ldots, L \tag{11c}$$

where $\mathbf{W} \triangleq [\mathbf{W}^{(1)} \ldots \mathbf{W}^{(L)}]$. A part of the forthcoming development is easier to follow if an equivalent representation of (11c) is used that, similar to (11b), explicitly shows the linear dependency of the left-hand side (LHS) of (11c) on $\mathbf{p}$. It is straightforward to show from (4) that the lth constraint in (11c) can be rewritten as

$$\mathbf{u}_l\left(\mathbf{W}^{(l-L-1)}\right)^T \mathbf{p} \leq P_l\left(\mathbf{W}^{(l-L-1)}\right) \tag{11d}$$

[2]Depending on the duplexing mode, orthogonal relay-destination channels may be devised in time or in frequency. All results in this paper hold for both cases.

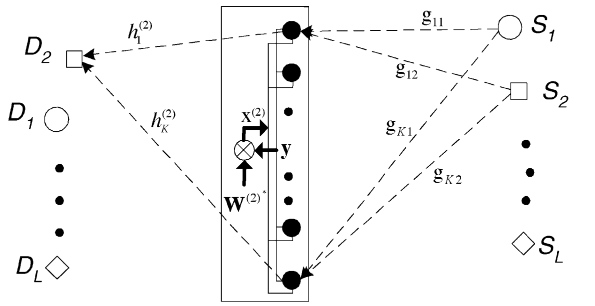

Fig. 1. Cooperative communication network with L source-destination pairs and a K-antenna relay. The depicted relay-destination channel gains and relaying matrix correspond to the second orthogonal channel.

TABLE I
SOME MATRIX OPERATIONS PROPERTIES FROM [17]

T1.1	$\mathrm{tr}\left(\mathbf{A}^T\mathbf{H}\right)$	$=$	$(\mathrm{vec}(\mathbf{A}))^T\mathrm{vec}(\mathbf{H})$
T1.2	$\mathrm{vec}(\mathbf{AS})$	$=$	$\left(\mathbf{S}^T\otimes\mathbf{I}\right)\mathrm{vec}(\mathbf{A})$
T1.3	$\mathrm{vec}\left(\mathbf{ASB}\right)$	$=$	$\left(\mathbf{B}^T\otimes\mathbf{A}\right)\mathrm{vec}(\mathbf{S})$
T1.4	$(\mathbf{A}\otimes\mathbf{B})^T$	$=$	$\left(\mathbf{A}^T\otimes\mathbf{B}^T\right)$
T1.5	$(\mathbf{A}\otimes\mathbf{B})(\mathbf{S}\otimes\mathbf{G})$	$=$	$\mathbf{AS}\otimes\mathbf{BG}$

where

$$\begin{aligned}&\mathbf{u}_l\left(\mathbf{W}^{(l-L-1)}\right)\\&\quad\triangleq\left[\left[\mathbf{G}^H\mathbf{W}^{(l-L-1)^T}\mathbf{W}^{(l-L-1)^*}\mathbf{G}\right]_{11}\cdots\right.\\&\qquad\left.\left[\mathbf{G}^H\mathbf{W}^{(l-L-1)^T}\mathbf{W}^{(l-L-1)^*}\mathbf{G}\right]_{LL}\right]^T\\&P_l\left(\mathbf{W}^{(l-L-1)}\right)\\&\quad\triangleq P^{(l-L-1)}-\mathrm{tr}\left(\mathbf{W}^{(l-L-1)^*}\mathbf{\Sigma}_v\mathbf{W}^{(l-L-1)^T}\right)\end{aligned}\tag{12}$$

and $l=L+2,\ldots,2L+1$.

The SINR expressions in (9) are decoupled in $\mathbf{w}^{(l)}$, $l=1,\ldots,L$ and, in that sense, (11) is reminiscent of the joint users' transmit powers and base-station receiver design in wireless communication networks. There is a rich literature on the joint optimization of transmit powers and receivers in which the studied problems can be typically cast as maximizing the minimum of the users' SINRs subject to a total transmit power constraint similar to the last constraint in (11b) [18], [21]–[23]. However, due to the multiple constraints on $\mathbf{p}$ and $\mathbf{w}^{(l)}$, $l=1,\ldots,L$ in (11b) and (11c) [or (11d)], these techniques cannot be directly applied to solve (11). It should be mentioned that the elegant technique proposed in [16] jointly optimizes the transmit powers and some beamforming vectors in the case that the transmit powers are constrained by several linear inequalities as in (11b). Unfortunately, the solution to (11) may not be obtained using the approach presented in [16] as (11c) (or (11d)) imposes L additional constraints whose lth one jointly depends on $\mathbf{w}^{(l)}$ and $\mathbf{p}$ and, further, is nonconvex with respect to the total design parameters $(\mathbf{w}^{(l)},\mathbf{p})$. Although (11) is more complex in comparison with the problem studied in [16], it will be shown in Section IV that the technique proposed here to solve (11) is less computationally complex than the approach introduced in [16]. Interestingly, it is straightforward to adopt the technique developed here to solve the problem studied in [16].

To determine the jointly optimal source power allocation and receiving strategy, it may be helpful to first obtain the optimal power vector for a given set of receivers and then use the acquired insight to tackle the original joint optimization problem. The next two sections adopt the above approach to obtain the jointly optimal power vector and relaying matrices. Section III derives the optimal power vector for a fixed set of relaying matrices while Section IV obtains the jointly optimal power vector and relaying matrices.

III. OPTIMAL POWER VECTOR FOR A FIXED SET OF RELAYING MATRICES

As the first step to solve (11), we consider a given set of relaying matrices $\mathbf{W}^{(1)}\ldots\mathbf{W}^{(L)}$ and derive the corresponding optimal power vector $\mathbf{p}_o(\mathbf{W})$. Due to technical reasons to be explained shortly below, we assume that

$$\left|\mathbf{w}^{(l)^H}\mathbf{f}_k^{(l)}\right|>0,\quad l,k\in\{1,\ldots,K\}.\tag{13}$$

The following reasons justify (13) in practice: 1) $\mathbf{f}_k^{(l)}$ is a stochastic vector and the probability that $\mathbf{w}^{(l)^H}\mathbf{f}_k^{(l)}=0$ is zero for any arbitrarily-selected $\mathbf{w}^{(l)}$. 2) As the aggregate noise power in (9) is larger than zero, the zero-forcing receiver is not SINR-optimal. This means that the optimal $\mathbf{w}^{(l)}$ that maximizes $\bar{\eta}_l(\mathbf{w}^{(l)},\mathbf{p})$ is not orthogonal to $\mathbf{f}_k^{(l)},k\neq l$ and, hence, satisfies (13). Following [24] and [25], (13) may be substituted by a more relaxed condition. However, such a relaxation mainly complicates the proofs of the forthcoming theorems while, unlike in [24] and [25], does not seem to correspond to scenarios of practical interest.

As the LHS of (11d) is always positive, it follows from (12) that the selected $\mathbf{W}^{(l)}$ should be such that $\mathrm{tr}(\mathbf{W}^{(l)^*}\mathbf{\Sigma}_v\mathbf{W}^{(l)^T})<P^{(l)}$ for $l=1,\ldots,L$. Assuming that the selected relaying matrices satisfy the latter conditions, (11) reduces to

$$\max_{\mathbf{p}}\quad\min_{1\le l\le L}\bar{\eta}_l\left(\mathbf{w}^{(l)},\mathbf{p}\right)\tag{14a}$$

$$\text{subject to}\quad\mathbf{u}_l^T\mathbf{p}\le P_l\quad l=1,\ldots,2L+1\tag{14b}$$

where, to simplify the notation, we have dropped the arguments of $\mathbf{u}_l(\mathbf{W}^{(l-L-1)})$ and $P_l(\mathbf{W}^{(l-L-1)})$ for $l=L+2,\ldots,2L+1$ in (14b). The following two observations are instrumental in deriving $\mathbf{p}_o(\mathbf{W})$:

Observation 1: When $\mathbf{p}=\mathbf{p}_o(\mathbf{W})$, at least one of the constraints in (14b) holds with equality. This is due to the fact that all normalized SINRs in (10) are strictly increasing functions of the power vector. Therefore, if all inequalities in (14b) are strict for $\mathbf{p}=\mathbf{p}_o(\mathbf{W})$, one can multiply $\mathbf{p}_o(\mathbf{W})$ with a factor $\alpha_1>1$ and, without violating any of the constraints in (14b), increase all $\bar{\eta}_l(\mathbf{w}^{(l)},\mathbf{p}_o(\mathbf{W}))$ to $\bar{\eta}_l(\mathbf{w}^{(l)},\alpha_1\mathbf{p}_o(\mathbf{W}))$. This contradicts the optimality of $\mathbf{p}_o(\mathbf{W})$.

Let $\mathcal{E}_o(\mathbf{W})$ denote the set of indexes of the constraints in (14b) that hold with equality when $\mathbf{p}=\mathbf{p}_o(\mathbf{W})$. We have

$$\mathbf{u}_l^T\mathbf{p}_o(\mathbf{W})=P_l\quad l\in\mathcal{E}_o(\mathbf{W})\tag{15}$$

$$\mathbf{u}_l^T\mathbf{p}_o(\mathbf{W})<P_l\quad l\notin\mathcal{E}_o(\mathbf{W}).\tag{16}$$

Observation 2: $\mathbf{p}_o(\mathbf{W})$ balances all normalized SINRs, that is,

$$\bar{\eta}(\mathbf{W}) = \bar{\eta}_l\left(\mathbf{w}^{(l)}, \mathbf{p}_o(\mathbf{W})\right) \quad \text{for all} \quad l = 1, \ldots, L. \tag{17}$$

It is due to (13), which implies that the lth normalized SINR is a strictly decreasing function of the powers of all sources but the lth one. Therefore, if one normalized SINR, say, $\bar{\eta}_{\tilde{l}}(\mathbf{w}^{(\tilde{l})}, \mathbf{p}_o(\mathbf{W}))$ is larger than the others, a decrease in the corresponding power $[\mathbf{p}_o(\mathbf{W})]_{\tilde{l}}$ results in an increase in other normalized SINRs, and, consequently, the objective function. This contradicts the optimality of $\mathbf{p}_o(\mathbf{W})$.

Note that the SINR balancing property is the core to many power control techniques [16], [18], [21]–[23], [26]. Let us introduce

$$[\boldsymbol{\Psi}(\mathbf{W})]_{lk} \triangleq \begin{cases} \left|\mathbf{w}^{(l)^H}\mathbf{f}_k^{(l)}\right|^2 & l \neq k \\ 0 & l = k \end{cases}$$
$$\boldsymbol{\Omega}(\mathbf{W}) \triangleq \mathbf{D}\left(\left[\frac{\gamma_1}{\left|\mathbf{w}^{(1)^H}\mathbf{f}_1^{(1)}\right|^2} \cdots \frac{\gamma_L}{\left|\mathbf{w}^{(L)^H}\mathbf{f}_L^{(L)}\right|^2}\right]^T\right)$$
$$[\boldsymbol{\sigma}(\mathbf{W})]_l \triangleq \mathbf{w}^{(l)^H}\boldsymbol{\Gamma}_v^{(l)}\mathbf{w}^{(l)} + \sigma_{n_l}^2. \tag{18}$$

Then, it is direct to show from (9) and (10) that $\mathbf{p}_o(\mathbf{W})$ in (17) satisfies

$$\boldsymbol{\Omega}(\mathbf{W})\boldsymbol{\Psi}(\mathbf{W})\mathbf{p} + \boldsymbol{\Omega}(\mathbf{W})\boldsymbol{\sigma}(\mathbf{W}) = \frac{1}{\bar{\eta}(\mathbf{W})}\mathbf{p}. \tag{19}$$

Using (19) in (15), it follows that $\mathbf{p}_o(\mathbf{W})$ is also a solution to

$$\frac{1}{P_l}\mathbf{u}_l^T\boldsymbol{\Omega}(\mathbf{W})\boldsymbol{\Psi}(\mathbf{W})\mathbf{p} + \frac{1}{P_l}\mathbf{u}_l^T\boldsymbol{\Omega}(\mathbf{W})\boldsymbol{\sigma}(\mathbf{W}) = \frac{1}{\bar{\eta}(\mathbf{W})} \tag{20}$$

for $l \in \mathcal{E}_o(\mathbf{W})$. Let

$$\boldsymbol{\Lambda}_l(\mathbf{W}) \triangleq \begin{bmatrix} \boldsymbol{\Omega}(\mathbf{W})\boldsymbol{\Psi}(\mathbf{W}) & \boldsymbol{\Omega}(\mathbf{W})\boldsymbol{\sigma}(\mathbf{W}) \\ \frac{1}{P_l}\mathbf{u}_l^T\boldsymbol{\Omega}(\mathbf{W})\boldsymbol{\Psi}(\mathbf{W}) & \frac{1}{P_l}\mathbf{u}_l^T\boldsymbol{\Omega}(\mathbf{W})\boldsymbol{\sigma}(\mathbf{W}) \end{bmatrix}, \tag{21}$$

for $l = 1, \ldots, 2L+1$. We have from (19) and (20) that $[\mathbf{p}_o(\mathbf{W})^T 1]^T$ satisfies

$$\boldsymbol{\Lambda}_l(\mathbf{W})\begin{bmatrix}\mathbf{p}\\1\end{bmatrix} = \frac{1}{\bar{\eta}(\mathbf{W})}\begin{bmatrix}\mathbf{p}\\1\end{bmatrix} \quad l \in \mathcal{E}_o(\mathbf{W}). \tag{22}$$

As $\bar{\eta}(\mathbf{W})$ and the transmit powers are positive scalars, (22) shows that $1/\bar{\eta}(\mathbf{W})$ and $[\mathbf{p}_o(\mathbf{W})^T 1]^T$ should constitute a jointly positive eigenpair of $\boldsymbol{\Lambda}_l(\mathbf{W})$ for all $l \in \mathcal{E}_o(\mathbf{W})$. To continue our developments, we need the following theorem whose proof is given in Appendix A.

Theorem 1: $\boldsymbol{\Lambda}_l(\mathbf{W})$ is a nonnegative primitive[3] matrix for $l = 1, \ldots, 2L+1$.

An interesting result of Theorem 1 is that the Perron's Theorem that holds for strictly positive matrices is also applicable to $\boldsymbol{\Lambda}_l(\mathbf{W}), l = 1, \ldots, 2L+1$ (see [27, Ch. 8]). In particular, we have the following.

Corollary 1: Among all eigenvalues of $\boldsymbol{\Lambda}_l(\mathbf{W})$, $\lambda_{\max}(\boldsymbol{\Lambda}_l(\mathbf{W}))$ is the unique eigenvalue of maximum modulus. It is real and positive, algebraically simple, and associated with a positive eigenvector that is unique up to an arbitrary scalar. $\boldsymbol{\Lambda}_l(\mathbf{W})$ does not have any other eigenpair that is jointly positive.

Note that $\lambda_{\max}(\boldsymbol{\Lambda}_l(\mathbf{W}))$ is usually called the Perron root of $\boldsymbol{\Lambda}_l(\mathbf{W})$. Let $\Omega(\boldsymbol{\Lambda}_l(\mathbf{W})) \triangleq [\mathbf{p}_l(\mathbf{W})^T 1]^T$ for $l = 1, \ldots, 2L+1$. Due to the structure of $\boldsymbol{\Lambda}_l(\mathbf{W})$ in (22) we have

$$\mathbf{u}_l^T\mathbf{p}_l(\mathbf{W}) = P_l \quad l = 1, \ldots, 2L+1. \tag{23}$$

Moreover, it is straightforward from Corollary 1 that

$$\bar{\eta}(\mathbf{W}) = \frac{1}{\lambda_{\max}(\boldsymbol{\Lambda}_l(\mathbf{W}))} \quad l \in \mathcal{E}_o(\mathbf{W}) \tag{24a}$$
$$\begin{bmatrix}\mathbf{p}_o(\mathbf{W})\\1\end{bmatrix} = \begin{bmatrix}\mathbf{p}_l(\mathbf{W})\\1\end{bmatrix} = \Omega(\boldsymbol{\Lambda}_l(\mathbf{W})) \quad l \in \mathcal{E}_o(\mathbf{W}). \tag{24b}$$

Equation (24a) implies that when $\mathrm{Card}(\mathcal{E}_o(\mathbf{W})) > 1$ we must have $\lambda_{\max}(\boldsymbol{\Lambda}_m(\mathbf{W})) = \lambda_{\max}(\boldsymbol{\Lambda}_n(\mathbf{W}))$ for $m, n \in \mathcal{E}_o(\mathbf{W})$. Moreover, (24b) suggests that $\mathbf{p}_o(\mathbf{W})$ is not unique unless $\mathbf{p}_m(\mathbf{W}) = \mathbf{p}_n(\mathbf{W})$ for $m, n \in \mathcal{E}_o(\mathbf{W})$. In what follows, we obtain $\mathcal{E}_o(\mathbf{W})$ and prove that $\lambda_{\max}(\boldsymbol{\Lambda}_m(\mathbf{W})) = \lambda_{\max}(\boldsymbol{\Lambda}_n(\mathbf{W}))$ and $\mathbf{p}_m(\mathbf{W}) = \mathbf{p}_n(\mathbf{W})$ for $m, n \in \mathcal{E}_o(\mathbf{W})$. First, we require the following theorem.

Theorem 2: For any two distinct m and $n \in \{1, \ldots, 2L+1\}$ we have

$$\mathbf{u}_m^T\mathbf{p}_n(\mathbf{W}) < P_m \tag{25}$$

if and only if

$$\lambda_{\max}(\boldsymbol{\Lambda}_m(\mathbf{W})) < \lambda_{\max}(\boldsymbol{\Lambda}_n(\mathbf{W})). \tag{26}$$

Moreover, the following four statements are equivalent:

$$\lambda_{\max}(\boldsymbol{\Lambda}_m(\mathbf{W})) = \lambda_{\max}(\boldsymbol{\Lambda}_n(\mathbf{W})) \tag{27a}$$
$$\mathbf{p}_m(\mathbf{W}) = \mathbf{p}_n(\mathbf{W}) \tag{27b}$$
$$\mathbf{u}_m^T\mathbf{p}_n(\mathbf{W}) = P_m \tag{27c}$$
$$\mathbf{u}_n^T\mathbf{p}_m(\mathbf{W}) = P_n. \tag{27d}$$

Proof: See Appendix B.

Now, we have all the machinery to prove the main result of this section.

Theorem 3: The following four statements are equivalent:

$$\begin{aligned}\lambda_{\max}\left(\boldsymbol{\Lambda}_{l_1}(\mathbf{W})\right) &= \cdots = \lambda_{\max}\left(\boldsymbol{\Lambda}_{l_M}(\mathbf{W})\right) \\ &> \lambda_{\max}\left(\boldsymbol{\Lambda}_l(\mathbf{W})\right) \quad l \notin \{l_1, \ldots, l_M\} \end{aligned}\tag{28}$$
$$\mathcal{E}_o(\mathbf{W}) = \{l_1, \ldots, l_M\} \tag{29}$$
$$\begin{aligned}\frac{1}{\bar{\eta}(\mathbf{W})} &= \lambda_{\max}\left(\boldsymbol{\Lambda}_{l_1}(\mathbf{W})\right) = \cdots \\ &= \lambda_{\max}\left(\boldsymbol{\Lambda}_{l_M}(\mathbf{W})\right) \\ &\neq \lambda_{\max}\left(\boldsymbol{\Lambda}_l(\mathbf{W})\right) \quad l \notin \{l_1, \ldots, l_M\}\end{aligned} \tag{30}$$
$$\begin{aligned}\mathbf{p}_o(\mathbf{W}) &= \mathbf{p}_{l_1}(\mathbf{W}) = \cdots = \mathbf{p}_{l_M}(\mathbf{W}) \\ &\neq \mathbf{p}_l(\mathbf{W}) \quad l \notin \{l_1, \ldots, l_M\}.\end{aligned} \tag{31}$$

Proof: See Appendix C.

Note that (28) states that the l_1th, the l_2th, …, and the l_Mth elements in the sequence $\lambda_{\max}(\boldsymbol{\Lambda}_l(\mathbf{W})), \quad l = 1, \ldots, 2L+1$ are equal to one another and are strictly larger than other elements. One may conjecture that it is unlikely to have $M = \mathrm{Card}(\mathcal{E}_o(\mathbf{W})) > 1$. However, it will be shown in

[3]A nonnegative matrix is primitive if it is irreducible and only has one eigenvalue of maximum modulus [27].

Section IV that when the optimal set of relaying matrices $\mathbf{W}_o \triangleq [\mathbf{W}_o^{(1)} \dots \mathbf{W}_o^{(L)}]$ is used, $M_o \triangleq \mathrm{Card}(\mathcal{E}_o(\mathbf{W}_o)) \geq L+1$.

Using the developments in this section, the procedure to solve (14) for a given $\mathbf{W} = [\mathbf{W}^{(1)} \dots \mathbf{W}^{(L)}]$ can be summarized as follows:

1) Form $\mathbf{\Lambda}_l(\mathbf{W})$ in (21) and compute $\lambda_{\max}(\mathbf{\Lambda}_l(\mathbf{W}))$ for $l = 1, \dots, 2L+1$.
2) Sort the so-obtained $\lambda_{\max}(\mathbf{\Lambda}_l(\mathbf{W}))$, $\quad l = 1, \dots, 2L+1$ in a non-increasing order. There may be $M \geq 1$ eigenvalues that are equal to one another but larger than the others. Denote the indexes of the corresponding matrices as $l_1, \dots, l_M$ and find $\bar{\eta}(\mathbf{W})$, the optimal value of the objective function in (14), from (30).
3) Form $\mathcal{E}_o(\mathbf{W})$ as in (29). Knowing that $\Omega(\mathbf{\Lambda}_{l_m}(\mathbf{W})) = [\mathbf{p}_{l_m}(\mathbf{W})^T 1]^T = [\mathbf{p}_{l_n}(\mathbf{W})^T 1]^T = \Omega(\mathbf{\Lambda}_{l_n}(\mathbf{W}))$ for any two distinct l_n and $l_m \in \mathcal{E}_o(\mathbf{W})$, choose an $l_m \in \mathcal{E}_o(\mathbf{W})$ and compute $\Omega(\mathbf{\Lambda}_{l_m}(\mathbf{W}))$. Obtain the unique $\mathbf{p}_o(\mathbf{W})$ from (31).

Note also from Theorem 3 that when $\mathbf{p} = \mathbf{p}_o(\mathbf{W})$, the l_1th, the l_2th, …, the l_Mth constraints in (14b) hold with equality and all other constraints hold with a strict inequality.

IV. Jointly Optimal Power Vector and Relaying Matrices

Building on the results in Section III, a four-step constructive approach is used in this section to derive a necessary and sufficient condition for the optimal pair of $\mathbf{W}_o$ and $\mathbf{p}_o \triangleq \mathbf{p}_o(\mathbf{W}_o)$ that solve (11). In the first step, an initial representation of this condition is given. The next two steps are intermediate stages towards an equivalent representation of the so-obtained condition that can be used to develop an efficient iterative algorithm to obtain $\mathbf{W}_o$ and $\mathbf{p}_o$. The final form of the necessary and sufficient optimality condition along with the aforementioned iterative algorithm are presented in the fourth step.

Step 1: Theorem 3 states that $\mathbf{p}_o$ can be uniquely computed as $[\mathbf{p}_o^T 1]^T = \Omega(\mathbf{\Lambda}_l(\mathbf{W}_o)), l \in \mathcal{E}_o(\mathbf{W}_o)$. To characterize $\mathbf{W}_o$, note from Theorem 3 that when a set of relaying matrices $\mathbf{W}$ is used, the balanced normalized SINR $\bar{\eta}(\mathbf{W})$ is equal to $1/\lambda_{\max}(\mathbf{\Lambda}_l(\mathbf{W})), l \in \mathcal{E}_o(\mathbf{W})$. Therefore, it should hold that $\lambda_{\max}(\mathbf{\Lambda}_l(\mathbf{W}_o)) \leq \lambda_{\max}(\mathbf{\Lambda}_l(\mathbf{W}))$ where $l \in \mathcal{E}_o(\mathbf{W}_o)$ and $\mathbf{W}$ is any arbitrary set of relaying matrices. In other words, $\mathbf{W}_o$ minimizes $\lambda_{\max}(\mathbf{\Lambda}_l(\mathbf{W}))$ for $l \in \mathcal{E}_o(\mathbf{W}_o)$. However, the latter minimization should be performed over the feasible set of relaying matrices that satisfy (11d). It should be stressed that the constraints in (11b) do not depend on relaying matrices and, therefore, do not affect the aforementioned feasible set. The above argument shows that $\mathbf{W}_o$ and $\mathbf{p}_o$ are jointly optimal if and only if

$$\mathbf{W}_o = \underset{\mathbf{W}}{\operatorname{argmin}}\, \lambda_{\max}(\mathbf{\Lambda}_l(\mathbf{W})) \quad l \in \mathcal{E}_o(\mathbf{W}_o), \tag{32a}$$

$$\text{subject to } \mathbf{u}_l\left(\mathbf{W}^{(l-L-1)}\right)^T \mathbf{p}_o \leq P_l\left(\mathbf{W}^{(l-L-1)}\right), \tag{32b}$$

for $l = L+2, \dots, 2L+1$, where

$$\begin{bmatrix} \mathbf{p}_o \\ 1 \end{bmatrix} = \Omega(\mathbf{\Lambda}_l(\mathbf{W}_o)) \quad l \in \mathcal{E}_o(\mathbf{W}_o). \tag{32c}$$

The optimization problem (32) has a complicated structure. Some challenges to solve (32) are as follows:

1) The aim in (32a) is to jointly minimize the common maximum-modulus eigenvalue of multiple non-symmetric matrices $\lambda_{\max}(\mathbf{\Lambda}_l(\mathbf{W})), l \in \mathcal{E}_o(\mathbf{W}_o)$. It seems that there is no systematic approach to solve such a problem in general.
2) This minimization should be performed over a feasible set that satisfies (32b). The feasible set in (32b) depends on $\mathbf{p}_o$ which, according to (32c), itself is a complicated function of the to-be-obtained $\mathbf{W}_o$.
3) As $\mathbf{W}_o$ is unknown, so is $\mathcal{E}_o(\mathbf{W}_o)$ in (32a) and (32c).
4) In general, $\mathcal{E}_o(\mathbf{W}_o)$ may have a nonempty intersection with the set of constraints indexes $\{L+2, \dots, 2L+1\}$ in (32b).

Therefore, some $\mathbf{u}_l(\mathbf{W}^{(l-L-1)})$ and $P_l(\mathbf{W}^{(l-L-1)})$ that determine the feasible set of relaying matrices in (32b) may also have an indirect effect on this feasible set through $\mathbf{p}_o$ in (32c).

Step 2: The following observations help to represent (32) in a more benign form.

Observation 3: $\mathbf{W}_o$ and $\mathbf{p}_o$ satisfy all constraints in (11d) with equality, that is,

$$\mathbf{u}_l\left(\mathbf{W}_o^{(l-L-1)}\right)^T \mathbf{p}_o = P_l\left(\mathbf{W}_o^{(l-L-1)}\right), \tag{33}$$

for $l = L+2, \dots, 2L+1$.

Equation (33) can be proved as follows. First, assume that $P_{\tilde{l}}(\mathbf{W}_o^{(\tilde{l}-L-1)})/\mathbf{u}_{\tilde{l}}(\mathbf{W}_o^{(\tilde{l}-L-1)})^T \mathbf{p}_o = \alpha_2 > 1$ for some $L+2 \leq \tilde{l} \leq 2L+1$. Then, without violating any constraint in (11b) or (11d), one can multiply $\mathbf{w}_o^{(\tilde{l}-L-1)}$ with $\sqrt{\alpha_2}$ and increase $\bar{\eta}_{\tilde{l}-L-1}(\mathbf{w}_o^{(\tilde{l}-L-1)}, \mathbf{p}_o)$ to $\bar{\eta}_{\tilde{l}-L-1}(\sqrt{\alpha_2}\mathbf{w}_o^{(\tilde{l}-L-1)}, \mathbf{p}_o)$. A part of the achieved gain in $\bar{\eta}_{\tilde{l}-L-1}(\sqrt{\alpha_2}\mathbf{w}_o^{(\tilde{l}-L-1)}, \mathbf{p}_o)$ can be offset by decreasing $[\mathbf{p}_o]_{\tilde{l}-L-1}$. Such a decrease in $[\mathbf{p}_o]_{\tilde{l}-L-1}$ results in an increase in $\bar{\eta}_l(\mathbf{w}_o^{(l)}, \mathbf{p}_o)$ for all $1 \leq l \neq \tilde{l}-L-1 \leq L$. This shows that if (33) does not hold for all $l = L+2, \dots, 2L+1$, one can increase the objective function in (11) without violating any constraints. This is in contradiction with the joint optimality of $\mathbf{W}_o$ and $\mathbf{p}_o$.

Observation 4: When $\mathbf{W}_o$ and $\mathbf{p}_o$ are jointly used, at least one of the inequalities in (11b) holds with equality, that is,

$$\mathbf{u}_l^T \mathbf{p}_o = P_l, \tag{34}$$

for at least one $l \in \{1, \dots, L+1\}$.

This can also be proved by contradiction. Assume that all inequalities in (11b) are strict at optimum. Then, $\min_{1 \leq l \leq L+1} P_l/\mathbf{u}_l^T \mathbf{p}_o \triangleq \alpha_3 > 1$ and $\alpha_3 \mathbf{p}_o$ does not violate any of the constraints in (11b). However, it follows from (33) that if $\alpha_3 \mathbf{p}_o$ is used along with $\mathbf{W}_o$, then all constraints in (11d) are violated. To avoid this, the relaying matrices $\mathbf{W}_o^{(l-L-1)}$ can be replaced by $\sqrt{\beta_{l-L-1}}\mathbf{W}_o^{(l-L-1)}$ where

$$\beta_{l-L-1} \triangleq P^{(l-L-1)}\left(\alpha_3 P^{(l-L-1)} - (\alpha_3 - 1)\mathrm{tr}\left({\mathbf{W}_o}^{(l-L-1)^*} \mathbf{\Sigma}_v {\mathbf{W}_o}^{(l-L-1)^T}\right)\right)^{-1}, \tag{35}$$

for $l = L+2, \ldots, 2L+1$. It is straightforward to verify that

$$\alpha_3 \mathbf{u}_l \left(\sqrt{\beta_{l-L-1}}\mathbf{W}_o^{(l-L-1)}\right)^T \mathbf{p}_o = P_l\left(\sqrt{\beta_{l-L-1}}\mathbf{W}_o^{(l-L-1)}\right), \quad (36)$$

for $l = L+2, \ldots, 2L+1$. Therefore, the set of relaying matrices $[\sqrt{\beta_1}\mathbf{W}_o^{(1)} \ldots \sqrt{\beta_L}\mathbf{W}_o^{(L)}]$ and the transmit power vector $\alpha_3\mathbf{p}_o$ satisfy all constraints in (11b) and (11d). It follows from (9) that

$$\eta_l\left(\sqrt{\beta_l}\mathbf{w}_o^{(l)}, \alpha_3\mathbf{p}_o\right) = \frac{[\mathbf{p}_o]_l \left|\mathbf{w}_o^{(l)^H}\mathbf{f}_l^{(l)}\right|^2}{\sum_{\substack{k=1\\k\neq l}}^{L}[\mathbf{p}_o]_k\left|\mathbf{w}_o^{(l)^H}\mathbf{f}_k^{(l)}\right|^2 + \frac{1}{\alpha_3}\cdot\mathbf{w}_o^{(l)^H}\mathbf{\Gamma}_v^{(l)}\mathbf{w}_o^{(l)} + \frac{1}{\alpha_3\beta_l}\cdot\sigma_{n_l}^2} \quad (37)$$

where

$$\frac{1}{\alpha_3\beta_l} = 1 - \frac{\alpha_3 - 1}{\alpha_3}\cdot\frac{\mathrm{tr}\left(\mathbf{W}_o{}^{(l)^*}\mathbf{\Sigma}_v\mathbf{W}_o{}^{(l)^T}\right)}{P^{(l)}} < 1, \quad (38)$$

for $l = 1, \ldots, L$. Using (38) along with the fact that $1/\alpha_3 < 1$, it can be obtained from (9), (10), and (37) that

$$\bar{\eta}_l\left(\sqrt{\beta_l}\mathbf{w}_o^{(l)}, \alpha_3\mathbf{p}_o\right) > \bar{\eta}_l\left(\mathbf{w}_o^{(l)}, \mathbf{p}_o\right) \quad l = 1, \ldots, L. \quad (39)$$

This contradicts the joint optimality of $\mathbf{W}_o$ and $\mathbf{p}_o$.

When $\mathbf{W}_o$ and $\mathbf{p}_o$ are jointly used, let $\mathcal{A}_o(\mathbf{W}_o) \triangleq \{l \in \{1, \ldots, L+1\} | \mathbf{u}_l^T\mathbf{p}_o = P_l\}$. Then, it directly follows from Observations 3 and 4 that $\mathcal{E}_o(\mathbf{W}_o) = \mathcal{A}_o(\mathbf{W}_o) \cup \{L+2, \ldots, 2L+1\}$ and the optimization problem (32) can be reformulated as

$$\mathbf{W}_o = \underset{\mathbf{W}}{\mathrm{argmin}}\, \lambda_{\max}(\mathbf{\Lambda}_l(\mathbf{W})) \quad l \in \mathcal{E}_o(\mathbf{W}_o), \quad (40a)$$

$$\text{subject to} \quad \mathbf{u}_l\left(\mathbf{W}^{(l-L-1)}\right)^T\mathbf{p}_o = P_l\left(\mathbf{W}^{(l-L-1)}\right), \quad (40b)$$

for $l = L+2, \ldots, 2L+1$, where

$$\begin{bmatrix}\mathbf{p}_o\\1\end{bmatrix} = \Omega\left(\mathbf{\Lambda}_l(\mathbf{W}_o)\right)$$
$$l \in \mathcal{A}_o(\mathbf{W}_o) \cup \{L+2, \ldots, 2L+1\}. \quad (40c)$$

As $\mathbf{p}_o$ is the optimal power vector associated with $\mathbf{W}_o$ and the equality constraints in (40b) hold for $\mathbf{W} = \mathbf{W}_o$, (40b) implies by its own that $\{L+2, \ldots, 2L+1\} \subseteq \mathcal{E}_o(\mathbf{W}_o)$ (see the definition of $\mathcal{E}_o(\mathbf{W}_o)$ in Section III). From the equivalence of (29) and (31), it follows that $\{L+2, \ldots, 2L+1\} \subseteq \mathcal{E}_o(\mathbf{W}_o)$ is an alternative representation of $[\mathbf{p}_o^T 1]^T = \Omega(\mathbf{\Lambda}_{L+2}(\mathbf{W}_o)) = \cdots = \Omega(\mathbf{\Lambda}_{2L+1}(\mathbf{W}_o))$. Therefore, (40c) is partly enforced by (40b). This redundancy can be avoided by replacing $l \in \mathcal{A}_o(\mathbf{W}_o) \cup \{L+2, \ldots, 2L+1\}$ in (40c) by $l \in \mathcal{A}_o(\mathbf{W}_o)$. Further, as $\mathcal{A}_o(\mathbf{W}_o) \subseteq \{1, \ldots, L+1\}$ is also a nonempty subset of $\mathcal{E}_o(\mathbf{W}_o)$, the equivalence of (28) and (29) establishes the fact that $l^\star \in \mathcal{A}_o(\mathbf{W}_o)$ if and only if $l^\star = \arg\max_{1\le l\le L+1}\lambda_{\max}(\mathbf{\Lambda}_l(\mathbf{W}_o))$. The optimization problem (40) can now be equivalently expressed as

$$\mathbf{W}_o = \underset{\mathbf{W}}{\mathrm{argmin}}\, \lambda_{\max}(\mathbf{\Lambda}_l(\mathbf{W})) \quad l \in \mathcal{E}_o(\mathbf{W}_o), \quad (41a)$$

$$\text{subject to } \mathbf{u}_l\left(\mathbf{W}^{(l-L-1)}\right)^T\mathbf{p}_o = P_l\left(\mathbf{W}^{(l-L-1)}\right), \quad (41b)$$

for $l = L+2, \ldots, 2L+1$, where

$$\begin{bmatrix}\mathbf{p}_o\\1\end{bmatrix} = \Omega\left(\mathbf{\Lambda}_{l^\star}(\mathbf{W}_o)\right) \quad l^\star = \underset{1\le l\le L+1}{\mathrm{argmax}}\, \lambda_{\max}\left(\mathbf{\Lambda}_l(\mathbf{W}_o)\right). \quad (41c)$$

Note that l and $l^\star$ in (41b) and (41c) belong, respectively, to the disjoint sets of $\{L+2, \ldots, 2L+1\}$ and $\{1, \ldots, L+1\}$. Moreover, according to (1), the set of $\mathbf{\Lambda}_l(\mathbf{W}_o)$ in (41c) depends on the pairs of $\mathbf{u}_l$ and P_l that are independent from the relaying matrices. These facts are useful in developing a simple iterative algorithm that alternately optimizes $\mathbf{p}$ and $\mathbf{W}$ to obtain $\mathbf{p}_o$ and $\mathbf{W}_o$. We would like to stress again the fact that even if $l^\star$ in (41c) is not unique, the equivalence of (27a) and (27b) guarantees the uniqueness of $\mathbf{p}_o$.

Step 3: The structural properties of $\mathbf{\Lambda}_l(\mathbf{W})$ can be used to represent (41) in a simpler form. First, as $\mathbf{\Lambda}_l(\mathbf{W}), l = 1, \ldots, 2L+1$ is a nonnegative primitive matrix, we have [27, Corollary 8.1.31]

$$\lambda_{\max}(\mathbf{\Lambda}_l(\mathbf{W})) = \max_{\mathbf{x}>\mathbf{0}}\min_{1\le i\le L+1}\frac{[\mathbf{\Lambda}_l(\mathbf{W})\mathbf{x}]_i}{[\mathbf{x}]_i} = \min_{\mathbf{x}>\mathbf{0}}\max_{1\le i\le L+1}\frac{[\mathbf{\Lambda}_l(\mathbf{W})\mathbf{x}]_i}{[\mathbf{x}]_i}, \quad (42)$$

for $l = 1, \ldots, 2L+1$. Moreover, the following lemma holds.

Lemma 1: If an $L\times 1$ vector $\mathbf{q} > \mathbf{0}$ satisfies the lth constraint in (14b) with equality, then

$$\max_{1\le i\le L+1}\frac{\left[\mathbf{\Lambda}_l(\mathbf{W})\begin{bmatrix}\mathbf{q}\\1\end{bmatrix}\right]_i}{\begin{bmatrix}\mathbf{q}\\1\end{bmatrix}_i} = \max_{1\le i\le L}\frac{1}{\bar{\eta}_i\left(\mathbf{w}^{(i)}, \mathbf{q}\right)},$$
$$\min_{1\le i\le L+1}\frac{\left[\mathbf{\Lambda}_l(\mathbf{W})\begin{bmatrix}\mathbf{q}\\1\end{bmatrix}\right]_i}{\begin{bmatrix}\mathbf{q}\\1\end{bmatrix}_i} = \min_{1\le i\le L}\frac{1}{\bar{\eta}_i\left(\mathbf{w}^{(i)}, \mathbf{q}\right)}. \quad (43)$$

Proof: See Appendix D.

Using a different representation, a special case of (43) has been shown in [18, Lemma 2] when only $\sum_{i=1}^{L}[\mathbf{q}]_i$ is constrained. An important property of the equations in (43) is that their right-hand sides (RHSs) are not functions of the index of $\mathbf{\Lambda}_l(\mathbf{W})$. Therefore, as long as $\mathbf{u}_l^T\mathbf{q} = P_l$, the extremal values at the LHSs of (43) are independent from l. The following theorem uses (42) and (43) to derive an equivalent form of (41).

Theorem 4: The optimization problem (41) can be equivalently represented as

$$\mathbf{w}_o^{(l)} = \underset{\mathbf{w}^{(l)}}{\operatorname{argmax}}\, \bar{\eta}_l \left(\mathbf{w}^{(l)}, \mathbf{p}_o\right) \quad l = 1, \ldots, L \quad \text{(44a)}$$

$$\text{subject to } {\mathbf{w}^{(l)}}^H \boldsymbol{\Xi}(\mathbf{p}_o)\mathbf{w}^{(l)} = P^{(l)} \quad l = 1, \ldots, L \quad \text{(44b)}$$

where

$$\begin{bmatrix} \mathbf{p}_o \\ 1 \end{bmatrix} = \Omega(\boldsymbol{\Lambda}_{l^\star}(\mathbf{W}_o)) \quad l^\star = \underset{1 \le l \le L+1}{\operatorname{argmax}}\, \lambda_{\max}(\boldsymbol{\Lambda}_l(\mathbf{W}_o)). \quad \text{(44c)}$$

Proof: See Appendix D.

Step 4: Note in Theorem 4 that the lth normalized SINR depends only on $\mathbf{w}^{(l)}$ and $\mathbf{w}^{(l)}$ appears only in the lth constraint in (44b). Therefore, if $\mathbf{p}_o$ is known, the subproblem (44a)–(44b) can be solved by independently maximizing $\bar{\eta}_l(\mathbf{w}^{(l)}, \mathbf{p}_o)$ subject to the lth constraint in (44b) for $l = 1, \ldots, L$. This property is used in the following theorem that describes $\mathbf{W}_o$ and $\mathbf{p}_o$ as explicit functions of one another.

Theorem 5: Let

$$\zeta_l(\mathbf{p}) \triangleq \xi_l \sqrt{P^{(l)}} \left({\mathbf{f}_l^{(l)}}^H \left(P^{(l)} \boldsymbol{\Upsilon}_l(\mathbf{p}) + \sigma_{n_l}^2 \boldsymbol{\Xi}(\mathbf{p}) \right)^{-1} \boldsymbol{\Xi}(\mathbf{p}) \times \left(P^{(l)} \boldsymbol{\Upsilon}_l(\mathbf{p}) + \sigma_{n_l}^2 \boldsymbol{\Xi}(\mathbf{p}) \right)^{-1} \mathbf{f}_l^{(l)} \right)^{-1/2} \quad (45)$$

where ξ_l can be selected as any arbitrary unit-norm scalar and

$$\boldsymbol{\Upsilon}_l(\mathbf{p}) \triangleq \mathbf{F}_{\bullet \bar{l}}^{(l)} \mathbf{D}(\mathbf{p}_{\bar{l}}) {\mathbf{F}_{\bullet \bar{l}}^{(l)}}^H + \boldsymbol{\Gamma}_v^{(l)} \quad (46)$$

with $\mathbf{F}^{(l)} \triangleq [\mathbf{f}_1^{(l)} \cdots \mathbf{f}_L^{(l)}]$. Then, $\mathbf{W}_o$ and $\mathbf{p}_o$ are jointly optimal if and only if

$$\mathbf{w}_o^{(l)} = \zeta_l(\mathbf{p}_o) \left(P^{(l)} \boldsymbol{\Upsilon}_l(\mathbf{p}_o) + \sigma_{n_l}^2 \boldsymbol{\Xi}(\mathbf{p}_o) \right)^{-1} \mathbf{f}_l^{(l)} \quad l = 1, \ldots, L \quad \text{(47a)}$$

$$\begin{bmatrix} \mathbf{p}_o \\ 1 \end{bmatrix} = \Omega(\boldsymbol{\Lambda}_{l^\star}(\mathbf{W}_o)) \quad l^\star = \underset{1 \le l \le L+1}{\operatorname{argmax}}\, \lambda_{\max}(\boldsymbol{\Lambda}_l(\mathbf{W}_o)). \text{(47b)}$$

Proof: See Appendix E.

The fact that $\mathbf{W}_o$ and $\mathbf{p}_o$ are explicit functions of one another in (47a) and (47b) gives rise to Algorithm I in Table II that obtains $\mathbf{W}_o$ and $\mathbf{p}_o$ with an arbitrary accuracy by alternately optimizing $\mathbf{W}_{[n]}$ for the given $\mathbf{p}_{[n-1]}$ and then $\mathbf{p}_{[n]}$ for the given $\mathbf{W}_{[n]}$. The convergence of $\mathbf{W}_{[n]}$ and $\mathbf{p}_{[n]}$ to $\mathbf{W}_o$ and $\mathbf{p}_o$ for a growing n is also shown in Appendix E. The convergence criterion in Algorithm I may be selected as $\|\mathbf{p}_{[n]} - \mathbf{p}_{[n-1]}\| < \epsilon_1$ or, considering that all normalized SINRs are balanced at optimum, as $\max_{1 \le l \le L} \bar{\eta}_l(\mathbf{w}_{[n]}^{(l)}, \mathbf{p}_{[n-1]}) - \min_{1 \le l \le L} \bar{\eta}_l(\mathbf{w}_{[n]}^{(l)}, \mathbf{p}_{[n-1]}) < \epsilon_2$ where ϵ_1 and ϵ_2 are small constants.

As will be verified by numerical results in Section VI, Algorithm I enjoys a very rapid convergence and, in practice, only a few (less than 5) iterations are typically required to balance all normalized SINRs to $\bar{\eta}(\mathbf{W}_o)$. The following remark concludes this section.

TABLE II
ALGORITHM I: SOLUTION TO JOINT POWER CONTROL AND RELAYING MATRICES DESIGN

1: `Initialize:` $n = 0$; $\mathbf{p}_{[0]} \ge \mathbf{0}$
2: `Repeat`
3: $n = n + 1$
4: $\mathbf{w}_{[n]}^{(l)} = \zeta_l\left(\mathbf{p}_{[n-1]}\right) \left(P^{(l)} \boldsymbol{\Upsilon}_l(\mathbf{p}_{[n-1]}) + \sigma_{n_l}^2 \boldsymbol{\Xi}(\mathbf{p}_{[n-1]}) \right)^{-1} \mathbf{f}_l^{(l)}$
5: $\mathbf{W}_{[n]} = \left[\mathbf{W}_{[n]}^{(1)} \cdots \mathbf{W}_{[n]}^{(L)} \right]$
6: $l_{[n]}^\star = \underset{1 \le l \le L+1}{\operatorname{argmax}}\, \lambda_{\max}\left(\boldsymbol{\Lambda}_l\left(\mathbf{W}_{[n]}\right)\right)$
7: $\begin{bmatrix} \mathbf{p}_{[n]} \\ 1 \end{bmatrix} = \Omega\left(\boldsymbol{\Lambda}_{l_{[n]}^\star}\left(\mathbf{W}_{[n]}\right)\right)$
8: `until the prescribed convergence criterion is satisfied`

Remark 1: The most computationally demanding steps in each iteration of Algorithm I are Steps 4 and 6 where it is required to compute the inverse of a $K^2 \times K^2$ matrix and eigendecompose $L + 1$ matrices of size $L + 1 \times L + 1$. Therefore, the computational complexity of each iteration of Algorithm I is $\mathcal{O}(K^6 + L^4)$. A slightly modified version of Algorithm I can be used to solve the joint beamforming and power allocation problem considered in [16] which may be recast as a subproblem of (11) with the objective function (11a) and the $L+1$ constraints in (11b). The approach used to solve (11a)–(11b) in [16] is to decouple it into $L+1$ parallel problems each of which aiming to optimize the objective function (11a) subject to one of the constraints in (11b). It has been shown in [16] that one and only one of the solutions to the so-obtained $L + 1$ decoupled problems is also the solution to (11a)–(11b). Scaling the problem in [16] such that it has the same dimensionality as our investigated problem, the approach used to solve (11a)–(11b) in [16] requires to iteratively solve up to $L+1$ decoupled problems with each iteration having the complexity of $\mathcal{O}(K^6 + L^3)$. This results in a total computational complexity of $\mathcal{O}(LK^6 + L^4)$ per iteration to solve (11a)–(11b). Therefore, the per-iteration computational complexity of Algorithm I is less than that of the algorithm in [16]. Both algorithms seem to have similar convergence rate and, hence, the smaller per-iteration computational complexity of Algorithm I may be translated into its overall computational advantage to the technique proposed in [16]. Recalling the discussion at the end of Section II, it should be further stressed that the algorithm in [16] may not be used to solve (11) in its entirety due to the presence of L additional nonconvex constraints in (11c).

V. JOINTLY OPTIMAL DESIGN IN A COGNITIVE NETWORK

A. Problem Formulation and Solution Description

In this section, the jointly optimal power vector and relaying matrices are derived in the case that the studied multipoint-to-multipoint cooperative system acts as a cognitive network that shares the spectrum with M receiving terminals of a primary system (see Fig. 2). In such a scenario, the cognitive network should not disrupt the communication between the primary terminals. When the cognitive transmitters and the primary receivers concurrently operate in the same radio channels, the latter requirement necessitates maintaining the interfering effect of the cognitive transmitters on the primary receivers below certain tolerance thresholds [14], [16], [28], [29]. Let t_{ml} denote the channel gain from S_l to the mth primary receiver U_m and P_{U_m} stand for the interference tolerance threshold of U_m in the first transmission phase of the cognitive network. Then, it should hold that

$$\mathbf{t}_m^T \mathbf{p} \leq P_{U_m} \quad m = 1, \ldots, M \tag{48}$$

where $\mathbf{t}_m = [|t_{m1}|^2 \ldots |t_{mL}|^2]^T$. In the second phase, the interfering effect on the primary network is due to the relay's L orthogonal transmissions. Let $r_{mk}^{(l)}$ denote the channel gain from the kth relay antenna to the mth primary receiver in the lth orthogonal channel. Then, the interference inflicted on the mth primary receiver in the lth channel is

$$I_m^{(l)} = \mathbf{r}_m^{(l)^T} \mathbf{x}^{(l)} = \mathbf{w}^{(l)^H} \left(\mathbf{y} \otimes \mathbf{r}_m^{(l)} \right) \quad m = 1, \ldots, M \tag{49}$$

where $\mathbf{r}_m^{(l)} = [r_{m1}^{(l)} \ldots r_{mK}^{(l)}]^T$ and the second equality is a direct result of (3). Assume that the interference tolerance threshold of U_m in the lth channel is $P_{U_m}^{(l)}$ which, in general, may be different from P_{U_m}. Then, it should hold that

$$\begin{aligned} \mathrm{E}\left\{ \left| I_m^{(l)} \right|^2 \right\} &= \mathbf{w}^{(l)^H} \left(\mathrm{E}\{\mathbf{y}\mathbf{y}^H\} \otimes \mathbf{r}_m^{(l)} \mathbf{r}_m^{(l)^H} \right) \mathbf{w}^{(l)} \\ &= \mathbf{w}^{(l)^H} \mathbf{\Xi}_m^{(l)}(\mathbf{p}) \mathbf{w}^{(l)} \\ &\leq P_{U_m}^{(l)} \quad m = 1, \ldots, M, \quad l = 1, \ldots, L \end{aligned} \tag{50}$$

where $\mathbf{\Xi}_m^{(l)}(\mathbf{p}) \triangleq (\mathbf{G}\mathbf{D}(\mathbf{p})\mathbf{G}^H + \mathbf{\Sigma}_v) \otimes \mathbf{r}_m^{(l)} \mathbf{r}_m^{(l)^H}$. Note that if the operating frequency band of U_m does not overlap with that of $S_1, \ldots, S_L$ or some of the L orthogonal channels, then the corresponding constraints in (48) and (50) should be ignored or, alternatively, the corresponding interference tolerance thresholds at the RHSs of (48) and (50) may be set to infinity.

Following a similar discussion prior to (11), the jointly optimal power vector $\mathbf{p}_{c,o}$ and the set of relaying matrices $\mathbf{W}_{c,o} \triangleq [\mathbf{W}_{c,o}^{(1)} \ldots \mathbf{W}_{c,o}^{(L)}]$ can be computed as the solution to

$$\begin{aligned} &\max_{\mathbf{W},\mathbf{p}} \quad \min_{1 \leq l \leq L} \bar{\eta}_l \left(\mathbf{w}^{(l)}, \mathbf{p} \right) \\ &\text{subject to} \quad \text{(11b), (11c), (48), (50).} \end{aligned} \tag{51}$$

While the objective function (11a) and the set of constraints (11b) and (11c) are shared in (11) and (51), the latter optimization problem is further constrained by M inequalities in (48) and ML inequalities in (50). Fortunately, (48) and (50), respectively, have the same structures as (11b) and (11c) [or (11d)] and, therefore, a procedure similar to that in Sections III and IV may be adopted to solve (51). However, one should take note of the following remarks when trying to adopt Observations 3 and 4 in the current context:

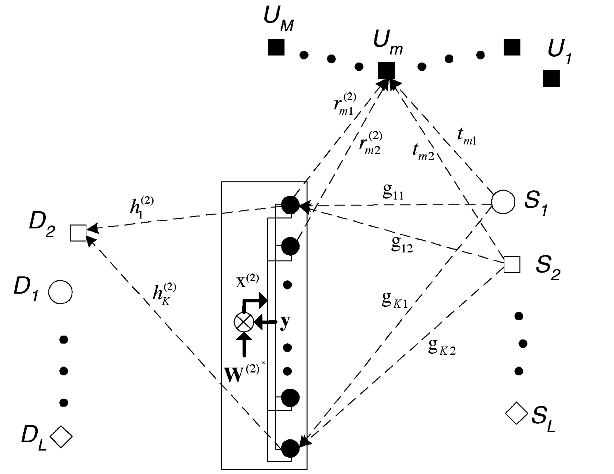

Fig. 2. The cooperative communication network as a cognitive system. The depicted relay channel gains and relaying matrix correspond to the second orthogonal channel.

Remark 2: In (11), $\mathbf{w}^{(l)}$ appears only in the lth constraint in (11c) [or (11d)]. Due to this one-to-one correspondence between $\mathbf{w}^{(l)}, l = 1, \ldots, L$ and the constraints in (11d), all those constraints hold with equality at the optimal point of (11). In contrast, in (51), $\mathbf{w}^{(l)}$ appears not only in the lth constraint in (11c) [or (11d)], but also in M constraints in (50). In general, not all above $M + 1$ constraints can simultaneously hold with equality at the optimal point of (51).

Remark 3: It has been shown in Observation 4 that at least one of the $L + 1$ inequalities in (11b) holds with equality at the optimal point of (11). Using a similar approach as in the proof of (34), it can be verified that at least one of the $L + 1 + M$ inequalities in (11b) and (48) hold with equality at the optimal point of (51).

Following a procedure similar to that in Sections III and IV prior to Theorem 4 while taking into account Remarks 2 and 3, the solution to (51) can be characterized in the theorem below.

Theorem 6: Let $\mathbf{w}_{c,o}^{(l)} \triangleq \mathrm{vec}(\mathbf{W}_{c,o}^{(l)})$ and $\mathbf{\Lambda}_{c,l}(\mathbf{W})$ as in (52) at the bottom of the page. Then, it holds for $l = 1, \ldots, L$ that

$$\mathbf{w}_{c,o}^{(l)} = \underset{\mathbf{w}^{(l)}}{\operatorname{argmax}}\, \bar{\eta}_l \left(\mathbf{w}^{(l)}, \mathbf{p}_{c,o} \right) \tag{53a}$$

$$\text{subject to} \quad \mathbf{w}^{(l)^H} \mathbf{\Xi}(\mathbf{p}_{c,o}) \mathbf{w}^{(l)} \leq P^{(l)} \tag{53b}$$

$$\mathbf{w}^{(l)^H} \mathbf{\Xi}_m^{(l)}(\mathbf{p}_{c,o}) \mathbf{w}^{(l)} \leq P_{U_m}^{(l)} \quad m = 1, \ldots, M \tag{53c}$$

$$\mathbf{\Lambda}_{c,l}(\mathbf{W}) \triangleq \begin{cases} \mathbf{\Lambda}_l(\mathbf{W}), & l = 1, \ldots, L+1 \\ \begin{bmatrix} \mathbf{\Omega}(\mathbf{W})\mathbf{\Psi}(\mathbf{W}) & \mathbf{\Omega}(\mathbf{W})\boldsymbol{\sigma}(\mathbf{W}) \\ \frac{1}{P_{U_{l-L-1}}} \mathbf{t}_{l-L-1}^T \mathbf{\Omega}(\mathbf{W})\mathbf{\Psi}(\mathbf{W}) & \frac{1}{P_{U_{l-L-1}}} \mathbf{t}_{l-L-1}^T \mathbf{\Omega}(\mathbf{W})\boldsymbol{\sigma}(\mathbf{W}) \end{bmatrix}, & l = L+2, \ldots, M+L+1 \end{cases}. \tag{52}$$

where

$$\begin{bmatrix} \mathbf{p}_{c,o} \\ 1 \end{bmatrix} = \Omega(\mathbf{\Lambda}_{c,l^\star}(\mathbf{W}_{c,o}))$$
$$l^\star = \underset{1\le l\le M+L+1}{\operatorname{argmax}} \lambda_{\max}(\mathbf{\Lambda}_{c,l}(\mathbf{W}_{c,o})). \quad (53d)$$

The following comparison between (44) and (53) is instructive: If $\mathbf{p}_o$ is known, then $\mathbf{w}_o^{(l)}, l = 1,\dots,L$ in (44) can be derived by independently solving the subproblem (44a)–(44b) for each l. Similarly, it can be observed from (53) that if $\mathbf{p}_{c,o}$ is known, then $\mathbf{w}_{c,o}^{(l)}, l = 1,\dots,L$ can be obtained by independently solving the subproblem (53a)–(53c) for each l. In other words, if $\mathbf{p} = \mathbf{p}_{c,o}$, then $\mathbf{w}_{c,o}^{(l)}$ is the solution to the following problem:

$$\max_{\mathbf{w}^{(l)}} \bar{\eta}_l\left(\mathbf{w}^{(l)}, \mathbf{p}\right) \quad (54a)$$
$$\text{subject to} \quad {\mathbf{w}^{(l)}}^H \mathbf{\Xi}(\mathbf{p})\mathbf{w}^{(l)} \le P^{(l)} \quad (54b)$$
$${\mathbf{w}^{(l)}}^H \mathbf{\Xi}_m^{(l)}(\mathbf{p})\mathbf{w}^{(l)} \le P_{U_m}^{(l)} \quad m = 1,\dots,M. \quad (54c)$$

In contrast to (44a)–(44b), (54) does not admit a closed-form solution. In fact, (54) is not even convex in $\mathbf{w}^{(l)}$ and, hence, may not lend itself in general to an efficient numerical solution technique. Despite the above fact, it is shown in Section V-B that (54) can be recast as a quasi-convex optimization problem and, hence, be efficiently solved through examining a sequence of convex feasibility problems [30].

B. Optimal Relaying Matrices for a Fixed Power Vector

First, it is direct to show from (9) that

$$\bar{\eta}_l\left(\mathbf{w}^{(l)}, \mathbf{p}\right) = \frac{([\mathbf{p}]_l/\gamma_l)\left|{\mathbf{w}^{(l)}}^H \mathbf{f}_l^{(l)}\right|^2}{{\mathbf{w}^{(l)}}^H \mathbf{\Upsilon}_l(\mathbf{p})\mathbf{w}^{(l)} + \sigma_{n_l}^2}. \quad (55)$$

Due to (55), the optimal solution vectors in (54) and the following problem are identical:

$$\max_{\mathbf{w}^{(l)}} \frac{\left|{\mathbf{w}^{(l)}}^H \mathbf{f}_l^{(l)}\right|^2}{{\mathbf{w}^{(l)}}^H \mathbf{\Upsilon}_l(\mathbf{p})\mathbf{w}^{(l)} + \sigma_{n_l}^2} \quad \text{subject to (54b), (54c).} \quad (56)$$

Next, note that neither the objective function nor the $M+1$ constraints in (56) change with an arbitrary rotation of $\mathbf{w}^{(l)}$. Therefore, without any loss of generality, $\mathbf{w}^{(l)}$ can be rotated such that $\Im\{{\mathbf{w}^{(l)}}^H \mathbf{f}_l^{(l)}\} = 0$. It is now direct to observe that the optimal solution vectors in (56) and the following problem are the same up to an immaterial unit-norm scalar:

$$\max_{\mathbf{w}^{(l)}} \frac{{\mathbf{w}^{(l)}}^H \mathbf{f}_l^{(l)}}{\left\| \begin{bmatrix} \mathbf{\Upsilon}_l(\mathbf{p})^{1/2} & \mathbf{0} \\ \mathbf{0} & \sigma_{n_l} \end{bmatrix} \begin{bmatrix} \mathbf{w}^{(l)} \\ 1 \end{bmatrix} \right\|}$$
$$\text{subject to} \quad \text{(54b), (54c)}$$
$$\Im\left\{{\mathbf{w}^{(l)}}^H \mathbf{f}_l^{(l)}\right\} = 0. \quad (57)$$

The to-be-maximized objective function, the $M+1$ inequality constraints, and the single equality constraint in (57) are quasi-concave, convex, and linear, respectively. Therefore, (57) is a quasi-convex optimization problem [30]. It can be equivalently represented as

$$\max_{\mathbf{w}^{(l)},\vartheta} \vartheta$$
$$\text{subject to} \quad \text{(54b), (54c)}$$
$$\Im\left\{{\mathbf{w}^{(l)}}^H \mathbf{f}_l^{(l)}\right\} = 0$$
$$\vartheta \left\| \begin{bmatrix} \mathbf{\Upsilon}_l(\mathbf{p})^{1/2} & \mathbf{0} \\ \mathbf{0} & \sigma_{n_l} \end{bmatrix} \begin{bmatrix} \mathbf{w}^{(l)} \\ 1 \end{bmatrix} \right\| \le {\mathbf{w}^{(l)}}^H \mathbf{f}_l^{(l)}. \quad (58)$$

Note that the last constraint in (58) is convex for any fixed ϑ. Let $\vartheta^\star$ denote the optimal value of the objective function obtained by solving (58). Fix ϑ and consider the following associated convex feasibility problem

$$\text{Find} \quad \mathbf{w}^{(l)}$$
$$\text{such that} \quad \text{(54b), (54c)}$$
$$\Im\left\{{\mathbf{w}^{(l)}}^H \mathbf{f}_l^{(l)}\right\} = 0$$
$$\vartheta \left\| \begin{bmatrix} \mathbf{\Upsilon}_l(\mathbf{p})^{1/2} & \mathbf{0} \\ \mathbf{0} & \sigma_{n_l} \end{bmatrix} \begin{bmatrix} \mathbf{w}^{(l)} \\ 1 \end{bmatrix} \right\| \le {\mathbf{w}^{(l)}}^H \mathbf{f}_l^{(l)}. \quad (59)$$

It is straightforward to verify that $\vartheta^\star \ge \vartheta$ if and only if (59) is feasible (see also [30]). This property can be exploited to develop an efficient algorithm that uses a simple bisection technique to solve the quasi-convex optimization problem (58). First, select an interval $[\vartheta_L \vartheta_U]$ that is known to contain $\vartheta^\star$. For instance, one can choose $\vartheta_L = 0$. To select ϑ_U, note from (58) that

$${\vartheta^\star}^2 \le \max_{\mathbf{w}^{(l)}} \frac{\left|{\mathbf{w}^{(l)}}^H \mathbf{f}_l^{(l)}\right|^2}{{\mathbf{w}^{(l)}}^H \mathbf{\Upsilon}_l(\mathbf{p})\mathbf{w}^{(l)} + \sigma_{n_l}^2} \quad \text{subject to (54b).} \quad (60)$$

Therefore, $\vartheta_U > 0$ can be selected such that

$${\vartheta_U}^2 = \max_{\mathbf{w}^{(l)}} \frac{\left|{\mathbf{w}^{(l)}}^H \mathbf{f}_l^{(l)}\right|^2}{{\mathbf{w}^{(l)}}^H \mathbf{\Upsilon}_l(\mathbf{p})\mathbf{w}^{(l)} + \sigma_{n_l}^2} \quad \text{subject to (54b).} \quad (61)$$

The optimization problem (61) has a closed-form solution for $\mathbf{w}^{(l)}$. Substituting the optimal $\mathbf{w}^{(l)}$ into the objective function of (61) yields $\vartheta_U = (P^{(l)} \cdot {\mathbf{f}_l^{(l)}}^H (P^{(l)}\mathbf{\Upsilon}_l(\mathbf{p}) + \sigma_{n_l}^2 \mathbf{\Xi}(\mathbf{p}))^{-1}\mathbf{f}_l^{(l)})^{1/2}$. After determining ϑ_L and ϑ_U, the convex feasibility problem (59) can be solved for the midpoint $\vartheta = (\vartheta_L + \vartheta_U)/2$ to determine whether $\vartheta^\star$ is in the lower or the upper half of the interval. If the problem is feasible, set $\vartheta_L = \vartheta$ and if not, set $\vartheta_U = \vartheta$. Continue the procedure until $\vartheta_U - \vartheta_L \le \epsilon_\vartheta$ where ϵ_ϑ is the required accuracy. It is guaranteed that $\vartheta^\star \in [\vartheta_L \vartheta_U]$ at each iteration of the above algorithm. As the length of $[\vartheta_L \vartheta_U]$ is halved at each iteration, exactly $\lceil \log_2(D_\vartheta/\epsilon_\vartheta) \rceil$ iterations are required to guarantee that ϑ is in the ϵ_ϑ-vicinity of $\vartheta^\star$ where D_ϑ is the length of the original interval. The last $\mathbf{w}^{(l)}$ obtained by successfully solving the convex feasibility problem (59) is then the approximate solution to (57) and, hence, to (54). Note that the approximation

TABLE III
ALGORITHM II: SOLUTION TO JOINT POWER CONTROL AND RELAYING MATRICES DESIGN IN A COGNITIVE NETWORK

```
1: Initialize: n = 0; p_[0] ≥ 0
2: Repeat
3: n = n + 1
4: Obtain w_[n]^(l), l = 1,...,L by solving (54) for p = p_[n−1]
5: W_[n] = [W_[n]^(1) ··· W_[n]^(L)]
6: l*_[n] = argmax_{1≤l≤M+L+1} λ_max(Λ_{c,l}(W_[n]))
7: [p_[n]; 1] = Ω(Λ_{c,l*_[n]}(W_[n]))
8: until the prescribed convergence criterion
   is satisfied
```

error can be made arbitrarily small by properly selecting ϵ_ϑ. In this paper, CVX [31], [32], a free package for specifying and solving convex problems, is used to solve (59). For other feasibility examining-based techniques to solve quasi-convex problems see, for instance, [19] and [33].

C. Derivation of the Jointly Optimal Solution

The technique developed in Section V-B to solve (54) obtains the optimal relaying matrices for a given power vector. This, along with Theorem 6 that expresses $\mathbf{W}_{c,o}$ and $\mathbf{p}_{c,o}$ as explicit functions of one another, give rise to Algorithm II in Table III that derives $\mathbf{W}_{c,o}$ and $\mathbf{p}_{c,o}$ with an arbitrary accuracy by alternately optimizing $\mathbf{W}_{[n]}$ for the given $\mathbf{p}_{[n-1]}$ and then $\mathbf{p}_{[n]}$ for the given $\mathbf{W}_{[n]}$. Note that although Algorithms I and II have similar structures, they have two essential differences: 1) In Algorithm I, $\mathbf{w}_{[n]}^{(l)}$ is a closed-form function of $\mathbf{p}_{[n-1]}$ while, in Algorithm II, it should be obtained using the bisection technique developed in Section V-B; 2) in Algorithm I, $l^\star_{[n]}$ is the argument for which $\lambda_{\max}(\mathbf{\Lambda}_l(\mathbf{W}_{[n]})), l = 1,\ldots,L+1$ is maximal, while in Algorithm II, $l^\star_{[n]}$ is the argument for which $\lambda_{\max}(\mathbf{\Lambda}_{c,l}(\mathbf{W}_{[n]})),\quad l = 1,\ldots,M+L+1$ is maximal.

The following remark is in order.

Remark 4: Since $\mathbf{\Lambda}_{c,l}(\mathbf{W}_{[n]}), l = 1,\ldots,M+L+1$ are nonnegative primitive matrices, if $\mathbf{\Lambda}_{c,l_1}(\mathbf{W}_{[n]}) \le \mathbf{\Lambda}_{c,l_2}(\mathbf{W}_{[n]})$, then $\lambda_{\max}(\mathbf{\Lambda}_{c,l_1}(\mathbf{W}_{[n]})) \le \lambda_{\max}(\mathbf{\Lambda}_{c,l_2}(\mathbf{W}_{[n]}))$ [27, Corollary 8.1.19]. This property may be useful in reducing the computational complexity of Algorithm II. First, note from (21) and (52) that, for any two distinct $l_1, l_2 \in \{1,\ldots,L+M+1\}$, $\lambda_{\max}(\mathbf{\Lambda}_{c,l_1}(\mathbf{W}_{[n]}))$ and $\lambda_{\max}(\mathbf{\Lambda}_{c,l_2}(\mathbf{W}_{[n]}))$ only differ in their last rows. Therefore, if, for instance, the last row of $\mathbf{\Lambda}_{c,l_1}(\mathbf{W}_{[n]})$ is element-wise less than or equal to that of $\mathbf{\Lambda}_{c,l_2}(\mathbf{W}_{[n]})$, then $l^\star_{[n]} \neq l_1$, and, hence, there is no need to compute $\lambda_{\max}(\mathbf{\Lambda}_{c,l_1}(\mathbf{W}_{[n]}))$. This decreases the required number of maximum eigenvalue computations in step 6 of Algorithm II and, consequently, reduces the overall computational complexity of the algorithm. Finally, note from (52) that, for any two distinct $l_1, l_2 \in \{L+2,\ldots,L+M+1\}$, the last line of $\mathbf{\Lambda}_{c,l_1}(\mathbf{W}_{[n]})$ is element-wise less than or equal to that of $\mathbf{\Lambda}_{c,l_2}(\mathbf{W}_{[n]})$ if $\frac{1}{P_{U_{l_1-L-1}}}\mathbf{t}_{l_1-L-1} \le \frac{1}{P_{U_{l_2-L-1}}}\mathbf{t}_{l_2-L-1}$.

VI. SIMULATIONS

Simulation examples are used to verify the derived analytical results. In all examples, $\sigma^2_{v_k} = \sigma^2_{n_l} = \sigma^2$ for $k = 1,\ldots,K$ and $l = 1,\ldots,L$ and $P^{(l)} = 10\sigma^2$ for $l = 1,\ldots,L$. The target SINRs are set to $\gamma_l = 10$ (dB) for $l = 1,\ldots,L$. All source-relay and relay-destination channel gains are randomly and independently drawn from a zero-mean unit-variance circular complex Gaussian distribution and in all but the first example remain fixed during the simulation.

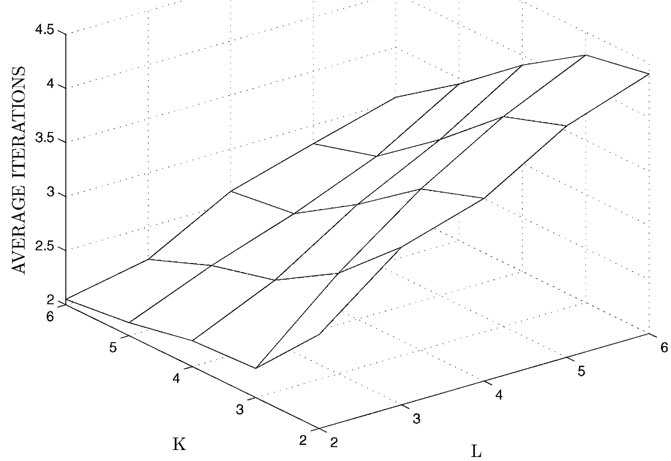

Fig. 3. Average number of iterations required for the convergence of Algorithm I versus L and K.

Fig. 3 shows the average number of iterations required to satisfy the convergence criterion $\epsilon_2 = 0.02$ in Algorithm I for different L and K and 100 independent realizations of the channel gains. Upper bounds on the sources' individual transmit power are set to $P_l = 10\sigma^2$ for $l = 1,\ldots,L$ while the sources' total transmit power is constrained by $P_{L+1} = \varphi_T \cdot L \cdot 10\sigma^2$ with $\varphi_T = 0.8$. Note that since $P_{L+1} < \sum_{l=1}^{L} P_l$, the total transmit power constraint is not trivially satisfied. As can be observed from Fig. 3, the average number of required iterations increases with L. However, even in the most severe case when only $K = 2$ relay antennas are available to balance $L = 6$ normalized SINRs, the average number of required iterations does not exceed 4.5. This verifies the excellent convergence rate of the algorithm.

Throughout the rest of simulations, $L = 3$ and $K = 4$ are considered. Fig. 4 shows $\bar{\eta}_l(\mathbf{w}_{[n]}^{(l)}, \mathbf{p}_{[n-1]})$ versus the iteration index n in Algorithm I with $P_l = 10\sigma^2$ for $l = 1,2,3$ and $P_4 = \varphi_T \cdot 3 \cdot 10\sigma^2$ with $\varphi_T = 0.8$. A very rapid convergence of all normalized SINRs to the common optimal value can be observed from the figure.

In the next example, it is assumed that $P_l = \varphi \cdot p$ for $l = 1,2,3$ and $P_4 = 3 \cdot p$. Then, the optimal balanced normalized SINRs $\bar{\eta}_l(\mathbf{w}_o^{(l)}, \mathbf{p}_o)$ are obtained from Algorithm I. Fig. 5 displays $\bar{\eta}_l(\mathbf{w}_o^{(l)}, \mathbf{p}_o)$ versus p/σ^2 for several φ. For the sake of comparison, the minimum of the normalized SINRs when only the relaying matrices are optimized and all sources' transmit powers are equal to p is also shown. In the latter case, the sources' total transmit power is P_4, and, hence, is always larger than or equal to the sources' total transmit power in the case when the sources' transmit powers and the relaying matrices are jointly optimized. Despite the above fact, Fig. 5

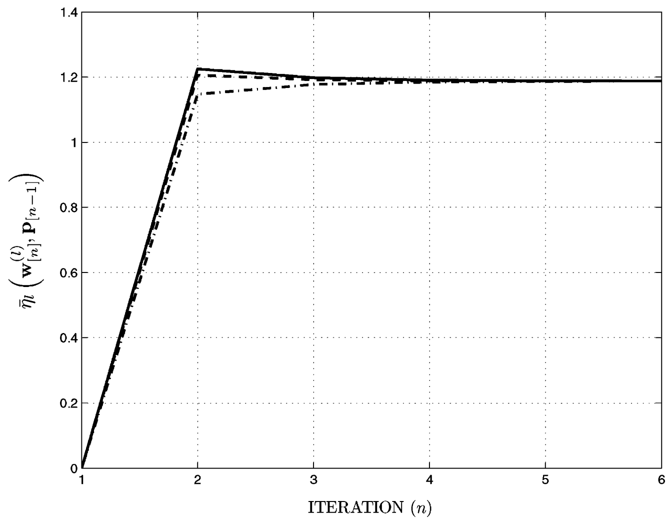

Fig. 4. $\bar{\eta}_l(\mathbf{w}_{[n]}^{(l)}, \mathbf{p}_{[n-1]})$ versus the iteration index n in Algorithm I.

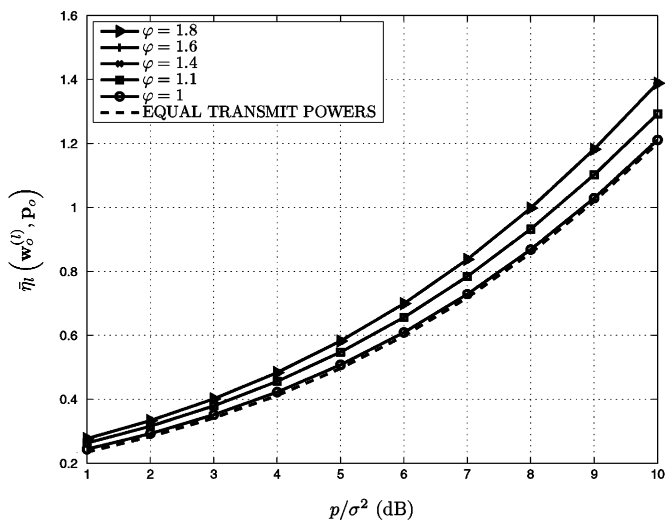

Fig. 5. $\bar{\eta}_l(\mathbf{w}_o^{(l)}, \mathbf{p}_o)$ versus p/σ^2.

shows that the proposed joint optimization approach always performs better than the case when the sources transmit with equal powers. Note that as φ increases, the upper-bound on the sources' individual transmit power gets larger and the sources' total transmit power constraint is expected to act as the only active constraint[4] in (11b). If it is the case, the optimization problems corresponding to larger values of φ share exactly the same set of active constraints and, hence, have exactly the same solution. This property explains the observation that the curves corresponding to $\varphi = 1.6$ and $\varphi = 1.8$ in Fig. 5 are indistinguishable from each other. Finally, note that as the total transmit power constraint seems to be active when $\varphi = 1.6$ and $\varphi = 1.8$, the sum of the sources' transmit powers should be equal to that in the simulated equal-power case. Compared to the latter curve, the curves corresponding to $\varphi = 1.6$ and $\varphi = 1.8$ show close to 20% increase in the minimum of the normalized SINRs.

[4]The active constraint is the one that holds with the equality.

The next example investigates the case when the upper-bounds on the sources' individual transmit powers are $P_1 = \alpha \cdot 10\sigma^2$ and $P_2 = P_3 = 10\sigma^2$ while the sources' total transmit power is constrained by $P_4 = \varphi_T \cdot (P_1 + P_2 + P_3)$. Fig. 6 displays $\bar{\eta}_l(\mathbf{w}_o^{(l)}, \mathbf{p}_o)$ versus α for $\varphi_T = 0.5$, $\varphi_T = 0.8$, and $\varphi_T = 1.1$. For the sake of comparison, a plot associated with each of the three $\bar{\eta}_l(\mathbf{w}_o^{(l)}, \mathbf{p}_o)$ curves is also shown that displays $\tilde{\eta}$, the minimum of the normalized SINRs when only the relaying matrices are optimized and all sources' transmit powers are equal to $\iota = ([\mathbf{p}_o]_1 + [\mathbf{p}_o]_2 + [\mathbf{p}_o]_3)/3$. Note that the sources' total transmit powers in each of the associated $\tilde{\eta}$ plots is equal to the sources' total transmit powers in its corresponding $\bar{\eta}_l(\mathbf{w}_o^{(l)}, \mathbf{p}_o)$ curve. It can be observed from Fig. 6 that the minimum of the normalized SINRs obtained by the proposed joint optimization approach is substantially higher than its counterpart when sources' transmit powers are equal.

It is known from Observation 4 that when the sources' transmit powers and the relaying matrices are jointly optimized, at least one of the inequalities in (11b) holds with equality. Fig. 7 shows the index of the active constraint in (11b) versus α for the three $\bar{\eta}_l(\mathbf{w}_o^{(l)}, \mathbf{p}_o)$ curves shown in Fig. 6. As expected, when α is small and, hence, the first source's transmit power has a small upper-bound, the first constraint in (11b) is active. As α grows, the fourth constraint that corresponds to the total transmit power becomes active first for $\varphi_T = 0.5$ and then for $\varphi_T = 0.8$. Note that when $\varphi_T = 1.1$, we have $P_4 > P_1 + P_2 + P_3$ and, therefore, the total transmit power constraint should never be active. This is corroborated in Fig. 7.

Figs. 8 and 9 investigate the case when the studied multipoint-to-multipoint cooperative system is a cognitive network that shares the spectrum with M receiving terminals of a primary system. It is assumed in both examples that $P_l = 10\sigma^2$ for $l = 1, 2, 3$ and $P_4 = \varphi_T \cdot 3 \cdot 10\sigma^2$ with $\varphi_T = 0.8$. Moreover, all source-primary user and relay-primary user channel gains are randomly and independently drawn from a zero-mean circular complex Gaussian distribution with the variance 0.2 and $P_{U_m} = P_{U_m}^{(l)} = \tilde{P}$ for $m = 1, \ldots, M$ and $l = 1, \ldots, L$ where $\tilde{P}/\sigma^2 = \varrho$.

Fig. 8 shows the optimal balanced normalized SINRs $\bar{\eta}_l(\mathbf{w}_{c,o}^{(l)}, \mathbf{p}_{c,o})$ obtained from Algorithm II versus the number of primary users M for $\varrho = 2$, $\varrho = 5$, and $\varrho = 10$ (dB). When $\varrho = 10$ (dB), the interference tolerance threshold of the primary users is so high that the constraints in (48) and (50) are very unlikely to be active. As can be observed from Fig. 8, in such a case the performance of the cognitive network is not sensitive to M. However, as $\varrho = 10$ decreases, the constraints in (48) and (50) can impose an increasingly degrading effect on the performance of the cognitive network. This effect should be in general exacerbated as M and, consequently, the number of constraints in (48) and (50) increase. Fig. 8 verifies the above conjecture by showing that $\bar{\eta}_l(\mathbf{w}_{c,o}^{(l)}, \mathbf{p}_{c,o})$ is a decreasing function of M for $\varrho = 2$ and $\varrho = 5$.

Fig. 9 displays $\bar{\eta}_l(\mathbf{w}_{[n]}^{(l)}, \mathbf{p}_{[n-1]})$ versus the iteration index n in Algorithm II for $M = 3$ and $\varrho = 2$, $\varrho = 5$, and $\varrho = 10$ (dB). A rapid convergence to the optimal values can be observed in all cases.

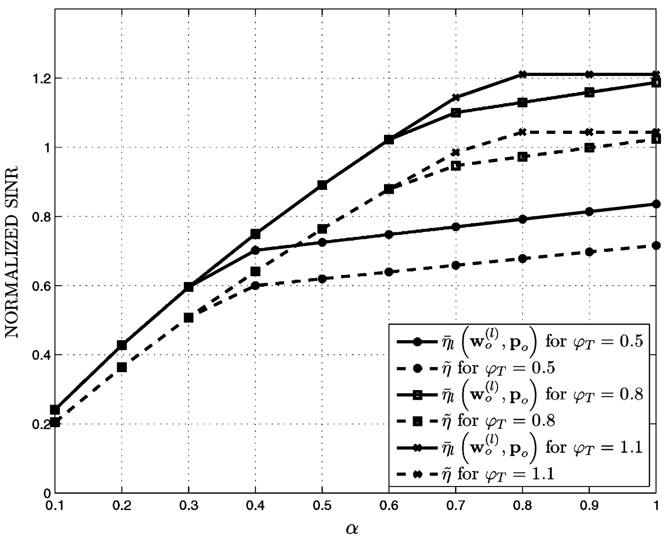

Fig. 6. $\bar{\eta}_l(\mathbf{w}_o^{(l)}, \mathbf{p}_o)$ along with $\tilde{\eta}$ versus α.

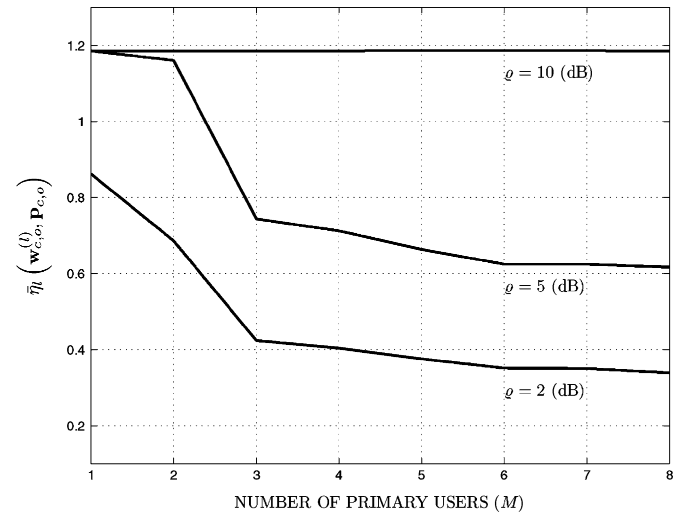

Fig. 8. $\bar{\eta}_l(\mathbf{w}_{c,o}^{(l)}, \mathbf{p}_{c,o})$ obtained from Algorithm II versus M.

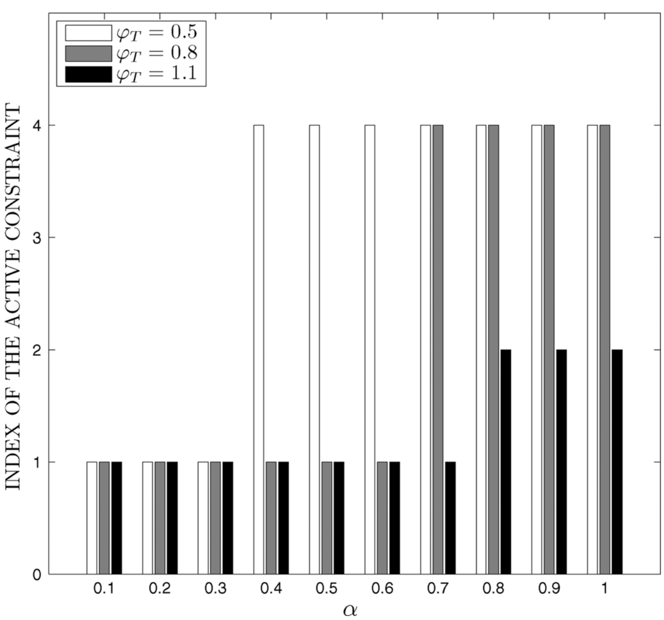

Fig. 7. Index of the active constraint in (11b) versus α for the $\bar{\eta}_l(\mathbf{w}_o^{(l)}, \mathbf{p}_o)$ curves shown in Fig. 6.

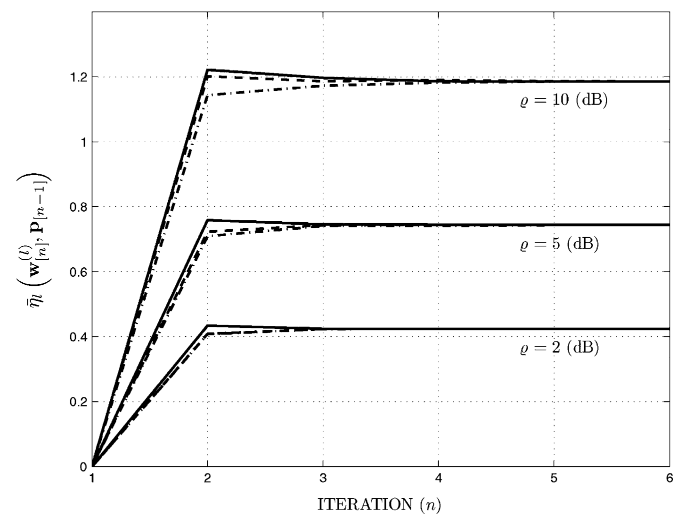

Fig. 9. $\bar{\eta}_l(\mathbf{w}_{[n]}^{(l)}, \mathbf{p}_{[n-1]})$ versus the iteration index n in Algorithm II for $M = 3$.

VII. CONCLUSION

A multipoint-to-multipoint cooperative communication network has been studied wherein multiple sources transmit their signals to a multiple-antenna relay in a shared channel and then the relay retransmits linearly processed versions of its received signal to the designated destinations in orthogonal channels. The jointly optimal sources' transmit powers and the relay's linear processing matrices have been sought so as to maximize the worst normalized SINR at the destinations subject to the relay's as well as both sources' individual and total transmit power constraints. The optimization problem has a nonconvex objective function with multiple nonconvex constraints. It has been shown that the optimal sources' transmit powers and the relaying matrices balance all normalized SINRs and an efficient iterative alternating optimization-based algorithm has been developed to obtain those optimal values. The per-iteration computational complexity order of the proposed algorithm has been derived and it has been shown that although the problem considered in [16] can be recast as a simple special case of the problem investigated in this paper, the proposed algorithm has a less per-iteration computation complexity order than that of the algorithm in [16]. An extension to the studied joint optimization problem has then been tackled in the case that the sources and relay are cognitive terminals whose transmit powers are further constrained by maximal admissible levels of interference induced on the primary users. Although these additional constraints necessitate using a completely new approach to find the optimal relaying matrices for a given set of the sources' transmit powers, the developed solution algorithm in the cognitive scenario has a minimal structural difference with its counterpart in the original noncognitive case. Simulation results have been used to verify the efficiency and the rapid convergence of the proposed algorithms.

An interesting research direction currently under investigation is to solve related problems in the case that the relay-destinations communication is also carried out in a shared channel. The solution to the latter problem will be disclosed in a future publication.

Appendix A
Proof of Theorem 1

The fact that $\boldsymbol{\Lambda}_l(\mathbf{W}) \geq \mathbf{0}$ is obvious from (18). To prove that $\boldsymbol{\Lambda}_l(\mathbf{W})$ is primitive, we need to show that there is an $m \geq 1$ for which $\boldsymbol{\Lambda}_l(\mathbf{W})^m > \mathbf{0}$ (see [27, Theorem 8.5.2]). It can be observed from (18) that each of the first L rows of $\boldsymbol{\Lambda}_l(\mathbf{W})$ has only one 0. It is also straightforward to show that the last row of $\boldsymbol{\Lambda}_l(\mathbf{W})$ has at most one 0 for any $\mathbf{u}_l \gneqq \mathbf{0}$. Now, note that $[\boldsymbol{\Lambda}_l(\mathbf{W})^2]_{ij} = \sum_{k=1}^{L+1}[\boldsymbol{\Lambda}_l(\mathbf{W})]_{ik}[\boldsymbol{\Lambda}_l(\mathbf{W})]_{kj}$ is comprised of $L+1$ summands. Following our above discussion, at most two of these summands are zero and the rest are positive. Therefore, $\boldsymbol{\Lambda}_l(\mathbf{W})^2 > \mathbf{0}$ for $L > 1$. Obviously, the single-user case of $L = 1$ is of no interest to us. This completes the proof of Theorem 1.

Appendix B
Proof of Theorem 2

We require the following two lemmas to prove Theorem 2.

Lemma 2: If $\boldsymbol{\Lambda}$ is a nonnegative primitive matrix, there is no $\mathbf{y}_1$ and $\mathbf{y}_2$ such that

$$\boldsymbol{\Lambda}\mathbf{y}_1 \lneqq \lambda_{\max}(\boldsymbol{\Lambda})\mathbf{y}_1 \tag{62}$$

$$\boldsymbol{\Lambda}\mathbf{y}_2 \gneqq \lambda_{\max}(\boldsymbol{\Lambda})\mathbf{y}_2. \tag{63}$$

Proof: Since $\boldsymbol{\Lambda}$ is a nonnegative primitive matrix, $\lambda_{\max}(\boldsymbol{\Lambda})$ is real and positive. Moreover, $\boldsymbol{\Lambda}^T$ is also a nonnegative primitive matrix and, hence, there exists a unique (up to a scaling factor) $\mathbf{q} > \mathbf{0}$ such that $\boldsymbol{\Lambda}^T\mathbf{q} = \lambda_{\max}(\boldsymbol{\Lambda})\mathbf{q}$ [27, Ch. 8]. Now, assume that (62) holds. As at least one entry of $\boldsymbol{\Lambda}\mathbf{y}_1$ is strictly less than the corresponding entry of $\lambda_{\max}(\boldsymbol{\Lambda})\mathbf{y}_1$, we must have $\mathbf{q}^T\boldsymbol{\Lambda}\mathbf{y}_1 < \lambda_{\max}(\boldsymbol{\Lambda})\mathbf{q}^T\mathbf{y}_1$, and, therefore, $\lambda_{\max}(\boldsymbol{\Lambda})\mathbf{q}^T\mathbf{y}_1 < \lambda_{\max}(\boldsymbol{\Lambda})\mathbf{q}^T\mathbf{y}_1$. This contradiction shows that there is no $\mathbf{y}_1$ that satisfies (62). The proof of (63) is similar to that of (62), and we skip it.

Lemma 3: The matrix $\lambda_{\max}(\boldsymbol{\Lambda}_n(\mathbf{W}))\mathbf{I} - \boldsymbol{\Omega}(\mathbf{W})\boldsymbol{\Psi}(\mathbf{W})$ is invertible for $n = 1, \ldots, 2L+1$.

Proof: As $[\mathbf{p}_n(\mathbf{W})^T 1]^T = \Omega(\boldsymbol{\Lambda}_n(\mathbf{W}))$, we have

$$\boldsymbol{\Omega}(\mathbf{W})\boldsymbol{\Psi}(\mathbf{W})\mathbf{p}_n(\mathbf{W}) + \boldsymbol{\Omega}(\mathbf{W})\boldsymbol{\sigma}(\mathbf{W}) = \lambda_{\max}(\boldsymbol{\Lambda}_n(\mathbf{W}))\mathbf{p}_n(\mathbf{W}). \tag{64}$$

Using (64) and the fact that $\boldsymbol{\Omega}(\mathbf{W})\boldsymbol{\sigma}(\mathbf{W}) > \mathbf{0}$, we obtain

$$\boldsymbol{\Omega}(\mathbf{W})\boldsymbol{\Psi}(\mathbf{W})\mathbf{p}_n(\mathbf{W}) < \lambda_{\max}(\boldsymbol{\Lambda}_n(\mathbf{W}))\mathbf{p}_n(\mathbf{W}). \tag{65}$$

As $\boldsymbol{\Omega}(\mathbf{W})\boldsymbol{\Psi}(\mathbf{W}) \geq \mathbf{0}$ and $\mathbf{p}_n(\mathbf{W}) > \mathbf{0}$, (65) implies that $\rho(\boldsymbol{\Omega}(\mathbf{W})\boldsymbol{\Psi}(\mathbf{W})) < \lambda_{\max}(\boldsymbol{\Lambda}_n(\mathbf{W}))$ [27, Corollary 8.1.20], and, hence, $\lambda_{\max}(\boldsymbol{\Lambda}_n(\mathbf{W}))\mathbf{I} - \boldsymbol{\Omega}(\mathbf{W})\boldsymbol{\Psi}(\mathbf{W})$ is invertible. This completes the proof.

Proof of Theorem 2: First, we assume that (25) holds and establish (26). Left-multiplying both sides of (64) by $\mathbf{u}_m^T$ and using (25) in the resulting equation, we obtain

$$\frac{1}{P_m}\mathbf{u}_m^T\boldsymbol{\Omega}(\mathbf{W})\boldsymbol{\Psi}(\mathbf{W})\mathbf{p}_n(\mathbf{W}) + \frac{1}{P_m}\mathbf{u}_m^T\boldsymbol{\Omega}(\mathbf{W})\boldsymbol{\sigma}(\mathbf{W}) < \lambda_{\max}(\boldsymbol{\Lambda}_n(\mathbf{W})). \tag{66}$$

A straightforward result of (64) and (66) is that

$$\boldsymbol{\Lambda}_m(\mathbf{W})\begin{bmatrix}\mathbf{p}_n(\mathbf{W})\\1\end{bmatrix} \lneqq \lambda_{\max}(\boldsymbol{\Lambda}_n(\mathbf{W}))\begin{bmatrix}\mathbf{p}_n(\mathbf{W})\\1\end{bmatrix}. \tag{67}$$

If (26) does not hold, we have $\lambda_{\max}(\boldsymbol{\Lambda}_m(\mathbf{W})) \geq \lambda_{\max}(\boldsymbol{\Lambda}_n(\mathbf{W}))$. Using the latter inequality in (67), it follows that

$$\boldsymbol{\Lambda}_m(\mathbf{W})\begin{bmatrix}\mathbf{p}_n(\mathbf{W})\\1\end{bmatrix} \lneqq \lambda_{\max}(\boldsymbol{\Lambda}_m(\mathbf{W}))\begin{bmatrix}\mathbf{p}_n(\mathbf{W})\\1\end{bmatrix}. \tag{68}$$

However, Lemma 2 states that (68) is impossible. Therefore, (26) holds. Now, assume that (26) holds and prove (25). From (26) and (64) it follows that

$$\boldsymbol{\Omega}(\mathbf{W})\boldsymbol{\Psi}(\mathbf{W})\mathbf{p}_n(\mathbf{W}) + \boldsymbol{\Omega}(\mathbf{W})\boldsymbol{\sigma}(\mathbf{W}) > \lambda_{\max}(\boldsymbol{\Lambda}_m(\mathbf{W}))\mathbf{p}_n(\mathbf{W}). \tag{69}$$

If (25) does not hold, then $\mathbf{u}_m^T\mathbf{p}_n(\mathbf{W}) \geq P_m$. Left-multiplying both sides of (69) by $\mathbf{u}_m^T$ and using the latter inequality yields

$$\frac{1}{P_m}\mathbf{u}_m^T\boldsymbol{\Omega}(\mathbf{W})\boldsymbol{\Psi}(\mathbf{W})\mathbf{p}_n(\mathbf{W}) + \frac{1}{P_m}\mathbf{u}_m^T\boldsymbol{\Omega}(\mathbf{W})\boldsymbol{\sigma}(\mathbf{W}) > \lambda_{\max}(\boldsymbol{\Lambda}_m(\mathbf{W})). \tag{70}$$

Combining (69) and (70), we obtain

$$\boldsymbol{\Lambda}_m(\mathbf{W})\begin{bmatrix}\mathbf{p}_n(\mathbf{W})\\1\end{bmatrix} > \lambda_{\max}(\boldsymbol{\Lambda}_m(\mathbf{W}))\begin{bmatrix}\mathbf{p}_n(\mathbf{W})\\1\end{bmatrix}. \tag{71}$$

Inequality (71) is in contradiction with Lemma 2. Therefore, (25) is correct. It remains to prove the equivalence of (27a)–(27d). First, let us show that (27a) implies (27b). From (64) and Lemma 3, we have that

$$\mathbf{p}_n(\mathbf{W}) = (\lambda_{\max}(\boldsymbol{\Lambda}_n(\mathbf{W}))\mathbf{I} - \boldsymbol{\Omega}(\mathbf{W})\boldsymbol{\Psi}(\mathbf{W}))^{-1} \times\boldsymbol{\Omega}(\mathbf{W})\boldsymbol{\sigma}(\mathbf{W}). \tag{72}$$

It is also known that

$$\boldsymbol{\Omega}(\mathbf{W})\boldsymbol{\Psi}(\mathbf{W})\mathbf{p}_m(\mathbf{W}) + \boldsymbol{\Omega}(\mathbf{W})\sigma(\mathbf{W}) = \lambda_{\max}(\boldsymbol{\Lambda}_m(\mathbf{W}))\mathbf{p}_m(\mathbf{W}) \tag{73}$$

which implies that

$$\mathbf{p}_m(\mathbf{W}) = (\lambda_{\max}(\boldsymbol{\Lambda}_m(\mathbf{W}))\mathbf{I} - \boldsymbol{\Omega}(\mathbf{W})\boldsymbol{\Psi}(\mathbf{W}))^{-1} \times\boldsymbol{\Omega}(\mathbf{W})\boldsymbol{\sigma}(\mathbf{W}). \tag{74}$$

Therefore, when (27a) holds, the RHSs of (72) and (74) are equal and (27b) immediately follows. To obtain (27a) from (27b), just note that when (27b) holds, the LHSs of (64) and (73) are equal and, therefore, $\lambda_{\max}(\boldsymbol{\Lambda}_n(\mathbf{W}))\mathbf{p}_n(\mathbf{W}) = \lambda_{\max}(\boldsymbol{\Lambda}_m(\mathbf{W}))\mathbf{p}_m(\mathbf{W}) = \lambda_{\max}(\boldsymbol{\Lambda}_m(\mathbf{W}))\mathbf{p}_n(\mathbf{W})$. As $\mathbf{p}_n(\mathbf{W}) \neq \mathbf{0}$, (27a) follows. This establishes the equivalence of (27a) and (27b). Now, let us prove the equivalence of (27b) and

(27c). It is immediate from (23) that (27c) follows from (27b). The fact that (27c) yields (27b) can be proved as follows. If (27c) holds, then (64) yields

$$\frac{1}{P_m}\mathbf{u}_m^T\mathbf{\Omega}(\mathbf{W})\mathbf{\Psi}(\mathbf{W})\mathbf{p}_n(\mathbf{W})+\frac{1}{P_m}\mathbf{u}_m^T\mathbf{\Omega}(\mathbf{W})\boldsymbol{\sigma}(\mathbf{W}) = \lambda_{\max}(\mathbf{\Lambda}_n(\mathbf{W})). \quad (75)$$

From (64) and (75), it holds that

$$\mathbf{\Lambda}_m(\mathbf{W})\begin{bmatrix}\mathbf{p}_n(\mathbf{W})\\ 1\end{bmatrix} = \lambda_{\max}(\mathbf{\Lambda}_m(\mathbf{W}))\begin{bmatrix}\mathbf{p}_n(\mathbf{W})\\ 1\end{bmatrix}. \quad (76)$$

However, $\mathbf{\Lambda}_m(\mathbf{W})$ is a nonnegative primitive matrix and has one and only one positive eigenpair. Therefore, $[\mathbf{p}_n(\mathbf{W})^T 1]^T = \Omega(\mathbf{\Lambda}_m(\mathbf{W})) = [\mathbf{p}_m(\mathbf{W})^T 1]^T$. This establishes (27b). The equivalence of (27d) and (27b) can be proved using a similar approach as in the proof of the equivalence of (27c) and (27b).

APPENDIX C
PROOF OF THEOREM 3

Let us first prove the equivalence of (28) and (31). Assume (28). According to Theorem 2, (28) holds if and only if

$$\mathbf{u}_{l_m}^T\mathbf{p}_{l_n}(\mathbf{W}) = P_{l_m} \quad l_m, l_n \in \{l_1,\ldots,l_M\} \quad (77)$$

and

$$\mathbf{u}_l^T\mathbf{p}_{l_n}(\mathbf{W}) < P_l \quad l_n \in \{l_1,\ldots,l_M\} \text{ and } l \notin \{l_1,\ldots,l_M\}. \quad (78)$$

Using the equivalence of (27b)–(27d), we have from (77) and (78) that $\mathbf{p}_{l_1}(\mathbf{W}) = \cdots = \mathbf{p}_{l_M}(\mathbf{W}) \neq \mathbf{p}_l(\mathbf{W})$ for $l \notin \{l_1,\ldots,l_M\}$. We know that $\mathbf{p}_o(\mathbf{W}) \in \{\mathbf{p}_1(\mathbf{W}),\ldots,\mathbf{p}_{2L+1}(\mathbf{W})\}$. However, $\mathbf{p}_o(\mathbf{W})$ cannot be equal to $\mathbf{p}_{\tilde{l}}$ for an $\tilde{l} \notin \{l_1,\ldots,l_M\}$. This is due to the fact that otherwise we have $\mathbf{u}_{l_n}^T\mathbf{p}_{\tilde{l}}(\mathbf{W}) \leq P_{l_n}$ and, hence, the contradictory relation of $\lambda_{\max}(\mathbf{\Lambda}_{\tilde{l}}(\mathbf{W})) \geq \lambda_{\max}(\mathbf{\Lambda}_{l_n}(\mathbf{W}))$ for $l_n \in \{l_1,\ldots,l_M\}$. This establishes (31). To prove the reverse direction assume (31). Then, (23) yields (77) and the equivalence of (27b)–(27d) along with the fact that $\mathbf{p}_o(\mathbf{W})$ satisfies all constraints in (14b) yield (78). We already know that (77) and (78) together are equivalent to (28). This completes the proof of the equivalence of (28) and (31). The equivalence of (29) and (31) can be shown as follows. Assume (29). From (24b) we have that $\mathbf{p}_o(\mathbf{W}) = \mathbf{p}_{l_n}(\mathbf{W})$ for $l_n \in \mathcal{E}_o(\mathbf{W})$. However, to obtain (31) we need to prove that $\mathbf{p}_o(\mathbf{W})$ is unique, that is, $\mathbf{p}_o(\mathbf{W}) = \mathbf{p}_{l_1}(\mathbf{W}) = \cdots = \mathbf{p}_{l_M}(\mathbf{W})$ and, further, $\mathbf{p}_o(\mathbf{W}) \neq \mathbf{p}_l(\mathbf{W})$ for $l \notin \{l_1,\ldots,l_M\}$. Using the fact that $\mathbf{p}_o(\mathbf{W}) = \mathbf{p}_{l_n}(\mathbf{W})$ for $l_n \in \mathcal{E}_o(\mathbf{W})$ in (15) and (16) and applying the equivalence of (27b)–(27d), (31) follows. In turn, when (31) holds, (29) directly follows from (23). Now, let us prove the equivalence of (30) and (31). From the equivalence of (27a) and (27b), it directly follows that $\mathbf{p}_{l_1}(\mathbf{W}) = \cdots = \mathbf{p}_{l_M}(\mathbf{W}) \neq \mathbf{p}_l(\mathbf{W})$ if and only if $\lambda_{\max}(\mathbf{\Lambda}_{l_1}(\mathbf{W})) = \cdots = \lambda_{\max}(\mathbf{\Lambda}_{l_M}(\mathbf{W})) \neq \lambda_{\max}(\mathbf{\Lambda}_l(\mathbf{W}))$ for $l \notin \{l_1,\ldots,l_M\}$. Now, recall that $1/\bar{\eta}(\mathbf{W})$ and $[\mathbf{p}_o(\mathbf{W})^T 1]^T$ constitute the jointly positive eigenpair of $\mathbf{\Lambda}_l(\mathbf{W})$ for some $l \in \{1,\ldots,2L+1\}$. Therefore, if $\mathbf{p}_o(\mathbf{W}) = \mathbf{p}_{l_1}(\mathbf{W}) = \cdots = \mathbf{p}_{l_M}(\mathbf{W}) \neq \mathbf{p}_l(\mathbf{W})$ for $l \notin \{l_1,\ldots,l_M\}$, then $1/\bar{\eta}(\mathbf{W})$ cannot be equal to $\lambda_{\max}(\mathbf{\Lambda}_{\tilde{l}}(\mathbf{W}))$ for an $\tilde{l} \notin \{l_1,\ldots,l_M\}$ for otherwise we have the contradictory relation of $\mathbf{p}_o(\mathbf{W}) = \mathbf{p}_{\tilde{l}}(\mathbf{W})$. A similar argument when $1/\bar{\eta}(\mathbf{W}) = \lambda_{\max}(\mathbf{\Lambda}_{l_1}(\mathbf{W})) = \cdots = \lambda_{\max}(\mathbf{\Lambda}_{l_M}(\mathbf{W})) \neq \lambda_{\max}(\mathbf{\Lambda}_l(\mathbf{W}))$ for $l \notin \{l_1,\ldots,l_M\}$ shows that $\mathbf{p}_o(\mathbf{W})$ cannot be equal to $\mathbf{p}_{\tilde{l}}(\mathbf{W})$ for an $\tilde{l} \notin \{l_1,\ldots,l_M\}$. This establishes the equivalence of (30) and (31) and completes the proof.

APPENDIX D
PROOFS OF LEMMA 1 AND THEOREM 4

Proof of Lemma 1: For any $L \times 1$ vector $\mathbf{q} > \mathbf{0}$ and $l = 1,\ldots,2L+1$, it holds that

$$\frac{\left[\mathbf{\Lambda}_l(\mathbf{W})\begin{bmatrix}\mathbf{q}\\ 1\end{bmatrix}\right]_i}{\begin{bmatrix}\mathbf{q}\\ 1\end{bmatrix}_i} = \begin{cases}\dfrac{1}{\bar{\eta}_i\left(\mathbf{w}^{(i)},\mathbf{q}\right)}, & i = 1,\ldots,L\\ \dfrac{1}{P_l}\cdot\displaystyle\sum_{m=1}^{L}\frac{[\mathbf{u}_l]_m[\mathbf{q}]_m}{\bar{\eta}_m\left(\mathbf{w}^{(m)},\mathbf{q}\right)}, & i = L+1.\end{cases} \quad (79)$$

Note that (79) directly follows from the definition of $\bar{\eta}_i(\mathbf{w}^{(i)},\mathbf{p})$ in (10) and that of $\mathbf{\Lambda}_l(\mathbf{W})$ in (21). If $\mathbf{q}$ satisfies the lth constraint in (14b) with equality, we have

$$\begin{aligned}\min_{1\leq i\leq L}\frac{1}{\bar{\eta}_i\left(\mathbf{w}^{(i)},\mathbf{q}\right)} &= \left(\min_{1\leq i\leq L}\frac{1}{\bar{\eta}_i\left(\mathbf{w}^{(i)},\mathbf{q}\right)}\right)\cdot\frac{\mathbf{u}_l^T\mathbf{q}}{P_l}\\ &\leq \frac{1}{P_l}\cdot\sum_{m=1}^{L}\frac{[\mathbf{u}_l]_m[\mathbf{q}]_m}{\bar{\eta}_m\left(\mathbf{w}^{(m)},\mathbf{q}\right)}\\ &\leq \left(\max_{1\leq i\leq L}\frac{1}{\bar{\eta}_i\left(\mathbf{w}^{(i)},\mathbf{q}\right)}\right)\cdot\frac{\mathbf{u}_l^T\mathbf{q}}{P_l}\\ &= \max_{1\leq i\leq L}\frac{1}{\bar{\eta}_i\left(\mathbf{w}^{(i)},\mathbf{q}\right)}.\end{aligned} \quad (80)$$

Using (80) in (79), the equations in (14b) follow.

Proof of Theorem 4: First, note that as (11c) and (11d) are equivalent, so are (41b) and (44b). As such, (41b) and (41c) are in fact repeated, respectively, in (44b) and (44c). Therefore, we need to show that if (41b) and (41c) hold, then (41a) yields (44a) and vice versa. To show this, we use an approach that is inspired by the proof of [18, Theorem 2]. Let us first assume (41a). For $l \in \mathcal{E}_o(\mathbf{W}_o)$ we have

$$\begin{aligned}\lambda_{\max}(\mathbf{\Lambda}_l(\mathbf{W}_o)) &\leq \lambda_{\max}(\mathbf{\Lambda}_l(\mathbf{W}))\\ &= \min_{\mathbf{x}>\mathbf{0}}\max_{1\leq i\leq L+1}\frac{[\mathbf{\Lambda}_l(\mathbf{W})\mathbf{x}]_i}{[\mathbf{x}]_i}\\ &\leq \max_{1\leq i\leq L+1}\frac{\left[\mathbf{\Lambda}_l(\mathbf{W})\begin{bmatrix}\mathbf{p}_o\\ 1\end{bmatrix}\right]_i}{\begin{bmatrix}\mathbf{p}_o\\ 1\end{bmatrix}_i}\\ &= \max_{1\leq i\leq L}\frac{1}{\bar{\eta}_i\left(\mathbf{w}^{(i)},\mathbf{p}_o\right)}\end{aligned} \quad (81)$$

where the first equality is due to (42) and the second equality follows from the fact that $\mathbf{p}_o$ satisfies the lth constraint in (14b) with equality for all $l \in \mathcal{E}_o(\mathbf{W}_o)$ and all matrices $\mathbf{W}$ that satisfy (41b). Therefore, for all the latter matrices $\mathbf{W}$ and all $l \in \mathcal{E}_o(\mathbf{W}_o)$ it holds that

$$\min_{\mathbf{W}} \max_{1\le i\le L} \frac{1}{\bar{\eta}_i\left(\mathbf{w}^{(i)},\mathbf{p}_o\right)} \ge \lambda_{\max}(\mathbf{\Lambda}_l(\mathbf{W}_o)). \tag{82}$$

As $[\mathbf{p}_o^T 1]^T = \Omega(\mathbf{\Lambda}_l(\mathbf{W}_o))$ for $l \in \mathcal{E}_o(\mathbf{W}_o)$, we have

$$\begin{aligned}\lambda_{\max}(\mathbf{\Lambda}_l(\mathbf{W}_o)) &= \frac{1}{\bar{\eta}(\mathbf{W}_o)} \\ &= \frac{1}{\bar{\eta}_1\left(\mathbf{w}_o^{(1)},\mathbf{p}_o\right)} = \cdots = \frac{1}{\bar{\eta}_L\left(\mathbf{w}_o^{(L)},\mathbf{p}_o\right)} \\ &= \max_{1\le i\le L} \frac{1}{\bar{\eta}_i\left(\mathbf{w}_o^{(i)},\mathbf{p}_o\right)}.\end{aligned} \tag{83}$$

It follows from (82) and (83) that

$$\begin{aligned}\min_{\mathbf{W}} \max_{1\le i\le L} \frac{1}{\bar{\eta}_i\left(\mathbf{w}^{(i)},\mathbf{p}_o\right)} &\ge \max_{1\le i\le L} \frac{1}{\bar{\eta}_i\left(\mathbf{w}_o^{(i)},\mathbf{p}_o\right)} \\ &= \frac{1}{\bar{\eta}(\mathbf{W}_o)} \\ &= \frac{1}{\bar{\eta}_1\left(\mathbf{w}_o^{(1)},\mathbf{p}_o\right)} = \cdots \\ &= \frac{1}{\bar{\eta}_L\left(\mathbf{w}_o^{(L)},\mathbf{p}_o\right)}.\end{aligned} \tag{84}$$

Obviously, the inequality in (84) can only hold with equality. Using the latter fact in (84), we have

$$\begin{aligned}\max_{\mathbf{W}} \min_{1\le i\le L} \bar{\eta}_i\left(\mathbf{w}^{(i)},\mathbf{p}_o\right) &= \bar{\eta}(\mathbf{W}_o) \\ &= \bar{\eta}_1\left(\mathbf{w}_o^{(1)},\mathbf{p}_o\right) = \cdots \\ &= \bar{\eta}_L\left(\mathbf{w}_o^{(L)},\mathbf{p}_o\right).\end{aligned} \tag{85}$$

Recall that $\mathbf{W}_o$ and $\mathbf{p}_o$ constitute the unique pair of solutions to (41) and $\bar{\eta}(\mathbf{W}_o)$ is the maximum possible value among all minimum normalized SINRs. Therefore, a necessary condition to achieve $\bar{\eta}(\mathbf{W}_o)$ is that all normalized SINRs are equal. Using this result in (85), we have

$$\begin{aligned}\max_{\mathbf{W}} \bar{\eta}_i\left(\mathbf{w}^{(i)},\mathbf{p}_o\right) &= \bar{\eta}(\mathbf{W}_o) \\ &= \bar{\eta}_1\left(\mathbf{w}_o^{(1)},\mathbf{p}_o\right) = \cdots \\ &= \bar{\eta}_L\left(\mathbf{w}_o^{(L)},\mathbf{p}_o\right) \quad i = 1,\ldots,L.\end{aligned} \tag{86}$$

Finally, note that $\bar{\eta}_i(\mathbf{w}^{(i)},\mathbf{p}_o)$ only depends on $\mathbf{w}^{(i)}$ and the maximum of $\bar{\eta}_i(\mathbf{w}^{(i)},\mathbf{p}_o)$ over $\mathbf{W}$ is in fact the maximum of $\bar{\eta}_i(\mathbf{w}^{(i)},\mathbf{p}_o)$ over $\mathbf{w}^{(i)}$ for $i = 1,\ldots,L$. This establishes (44a).

Let us now assume (44a) and prove (41a). We have for $l \in \mathcal{E}_o(\mathbf{W}_o)$ that

$$\begin{aligned}\min_{\mathbf{W}} \lambda_{\max}(\mathbf{\Lambda}_l(\mathbf{W})) &= \min_{\mathbf{W}} \max_{\mathbf{x}>\mathbf{0}} \min_{1\le i\le L+1} \frac{[\mathbf{\Lambda}_l(\mathbf{W})\mathbf{x}]_i}{[\mathbf{x}]_i} \\ &\ge \min_{\mathbf{W}} \min_{1\le i\le L+1} \frac{\left[\mathbf{\Lambda}_l(\mathbf{W})\begin{bmatrix}\mathbf{p}_o\\1\end{bmatrix}\right]_i}{\begin{bmatrix}\mathbf{p}_o\\1\end{bmatrix}_i} \\ &= \min_{\mathbf{W}} \min_{1\le i\le L} \frac{1}{\bar{\eta}_i\left(\mathbf{w}^{(i)},\mathbf{p}_o\right)} \\ &= \min_{1\le i\le L} \frac{1}{\bar{\eta}_i\left(\mathbf{w}_o^{(i)},\mathbf{p}_o\right)} \\ &= \min_{1\le i\le L+1} \frac{\left[\mathbf{\Lambda}_l(\mathbf{W}_o)\begin{bmatrix}\mathbf{p}_o\\1\end{bmatrix}\right]_i}{\begin{bmatrix}\mathbf{p}_o\\1\end{bmatrix}_i} \\ &= \lambda_{\max}\left(\mathbf{\Lambda}_l(\mathbf{W}_o)\right)\end{aligned} \tag{87}$$

where the second and the fourth equalities follow from the fact that $\mathbf{p}_o$ satisfies the lth constraint in (14b) with equality for all $l \in \mathcal{E}_o(\mathbf{W}_o)$ and all set of relaying matrices that satisfy (41b) and the third equality is due to (44a). Comparing the first and the last expressions in (87), it also follows that the inequality in (87) can only hold with equality. This establishes (41a) and completes the proof.

APPENDIX E
PROOFS OF THEOREM 5 AND THE CONVERGENCE OF ALGORITHM I

Proof of Theorem 5: Note that (44) is an alternative representation of the necessary and sufficient joint optimality condition (32). As (47b) repeats (44c), we only need to show that $\mathbf{w}_o^{(l)}$ in (47a) is the unique solution (up to an arbitrarily selected unit-norm scalar) to the subproblem (44a)–(44b). Let $\tilde{\mathbf{w}}^{(l)} \triangleq (1/\sqrt{P^{(l)}}) \cdot \mathbf{\Xi}(\mathbf{p}_o)^{1/2}\mathbf{w}^{(l)}$ and $\tilde{\mathbf{w}}_o^{(l)} \triangleq (1/\sqrt{P^{(l)}}) \cdot \mathbf{\Xi}(\mathbf{p}_o)^{1/2}\mathbf{w}_o^{(l)}$. Using (55) for $\mathbf{p} = \mathbf{p}_o$, the lth optimization problem in (44a)–(44b) may be represented as (88) at the bottom of the page. It follows from the unit-norm constraint on $\tilde{\mathbf{w}}^{(l)}$ that $\sigma_{n_l}^2 = \sigma_{n_l}^2 \tilde{\mathbf{w}}^{(l)^H}\tilde{\mathbf{w}}^{(l)}$ and (88) may be expressed as (89) at the top of the next page. The unique solution (up to an arbitrarily selected unit-norm scalar) to (89) is given by

$$\begin{aligned}\tilde{\mathbf{w}}_o^{(l)} = \frac{\zeta_l(\mathbf{p}_o)}{\sqrt{P^{(l)}}} \cdot \Big(\mathbf{\Xi}(\mathbf{p}_o)^{-1/2}\Big(&P^{(l)}\mathbf{\Upsilon}_l(\mathbf{p}_o) \\ &+ \sigma_{n_l}^2\mathbf{\Xi}(\mathbf{p}_o)\Big)\mathbf{\Xi}(\mathbf{p}_o)^{-1/2}\Big)^{-1}\mathbf{\Xi}(\mathbf{p}_o)^{-1/2}\mathbf{f}_l^{(l)}\end{aligned}$$

$$\tilde{\mathbf{w}}_o^{(l)} = \underset{\tilde{\mathbf{w}}^{(l)}}{\operatorname{argmax}} \frac{([\mathbf{p}_o]_l/\gamma_l)P^{(l)}\left|\tilde{\mathbf{w}}^{(l)^H}\mathbf{\Xi}(\mathbf{p}_o)^{-1/2}\mathbf{f}_l^{(l)}\right|^2}{P^{(l)}\tilde{\mathbf{w}}^{(l)^H}\mathbf{\Xi}(\mathbf{p}_o)^{-1/2}\mathbf{\Upsilon}_l(\mathbf{p}_o)\mathbf{\Xi}(\mathbf{p}_o)^{-1/2}\tilde{\mathbf{w}}^{(l)} + \sigma_{n_l}^2} \quad \text{subject to } \tilde{\mathbf{w}}^{(l)^H}\tilde{\mathbf{w}}^{(l)} = 1. \tag{88}$$

$$\tilde{\mathbf{w}}_o^{(l)} = \underset{\tilde{\mathbf{w}}^{(l)}}{\operatorname{argmax}} \frac{([\mathbf{p}_o]_l/\gamma_l)P^{(l)} \left|\tilde{\mathbf{w}}^{(l)^H}\boldsymbol{\Xi}(\mathbf{p}_o)^{-1/2}\mathbf{f}_l^{(l)}\right|^2}{\tilde{\mathbf{w}}^{(l)^H}\boldsymbol{\Xi}(\mathbf{p}_o)^{-1/2}\left(P^{(l)}\boldsymbol{\Upsilon}_l(\mathbf{p}_o)+\sigma_{n_l}^2\boldsymbol{\Xi}(\mathbf{p}_o)\right)\boldsymbol{\Xi}(\mathbf{p}_o)^{-1/2}\tilde{\mathbf{w}}^{(l)}} \quad \text{subject to } \tilde{\mathbf{w}}^{(l)^H}\tilde{\mathbf{w}}^{(l)} = 1. \tag{89}$$

$$= \frac{\zeta_l(\mathbf{p}_o)}{\sqrt{P^{(l)}}} \cdot \boldsymbol{\Xi}(\mathbf{p}_o)^{1/2}\left(P^{(l)}\boldsymbol{\Upsilon}_l(\mathbf{p}_o)+\sigma_{n_l}^2\boldsymbol{\Xi}(\mathbf{p}_o)\right)^{-1}\mathbf{f}_l^{(l)} \tag{90}$$

where the scaling factor $\zeta_l(\mathbf{p}_o)/\sqrt{P^{(l)}}$ in (90) is to guarantee $\tilde{\mathbf{w}}_o^{(l)^H}\tilde{\mathbf{w}}_o^{(l)} = 1$. Equation (47a) directly follows from (90). This completes the proof of the theorem.

Convergence of Algorithm I: The vector $\mathbf{p}_{[n]}$ in Algorithm I satisfies all constraints in (1) for all n. Therefore, $\|\mathbf{p}_{[n]}\|$ is bounded by a scalar that is independent from n and, hence, there exists a $\mathbf{p}_\infty$ such that $\lim_{n\to\infty}\mathbf{p}_{[n]} = \mathbf{p}_\infty$ [18], [34]. Note that a detailed proof of the convergence of $\mathbf{p}_{[n]}$ to a vector $\mathbf{p}_\infty$ may be presented along with the same lines as in the proof of ([35], Theorem 1). It is direct to show that $P^{(l)}\boldsymbol{\Upsilon}_l(\mathbf{p}_{[n-1]}) + \sigma_{n_l}^2\boldsymbol{\Xi}(\mathbf{p}_{[n-1]})$ is full-rank and $\zeta_l(\mathbf{p}_{[n-1]})$ is upper-bounded and both are continuous functions of $\mathbf{p}_{[n-1]}$ for $l = 1,\ldots,L+1$. Hence, it follows from Step 4 in Algorithm I that $\mathbf{w}_{[n]}^{(l)}$ is a continuous and bounded function of $\mathbf{p}_{[n-1]}$ for $l = 1,\ldots,L+1$. As such, $\lim_{n\to\infty}\mathbf{w}_{[n]}^{(l)} = \mathbf{w}_\infty^{(l)} \triangleq \zeta_l(\mathbf{p}_\infty)(P^{(l)}\boldsymbol{\Upsilon}_l(\mathbf{p}_\infty)+\sigma_{n_l}^2\boldsymbol{\Xi}(\mathbf{p}_\infty))^{-1}\mathbf{f}_l^{(l)}, \quad l = 1,\ldots,L+1$. Next, $\lambda_{\max}(\boldsymbol{\Lambda}_l(\mathbf{W}_{[n]}))$ is a continuous and bounded function of $\mathbf{W}_{[n]}$ for $l = 1,\ldots,L+1$. This shows that $\lambda_{\max}(\boldsymbol{\Lambda}_l(\mathbf{W}_{[n]}))$, $n = 1,2,\ldots$ is a convergent sequence to $\lambda_{\max}(\boldsymbol{\Lambda}_l(\mathbf{W}_\infty))$ for $l = 1,\ldots,L+1$. Let $l_\infty^\star \triangleq \operatorname{argmax}_{1\le l\le L+1}\lambda_{\max}(\boldsymbol{\Lambda}_l(\mathbf{W}_\infty))$. Note from the equivalence of (27a) and (27b) that even if $l_\infty^\star$ is not unique, $\Omega(\boldsymbol{\Lambda}_{l_\infty^\star}(\mathbf{W}_\infty))$ is unique. The vector $\Omega(\boldsymbol{\Lambda}_l(\mathbf{W}_{[n]}))$ is also a continuous function of $\mathbf{W}_{[n]}$ for $l = 1,\ldots,L+1$. Therefore, $\lim_{n\to\infty}\Omega(\boldsymbol{\Lambda}_l(\mathbf{W}_{[n]})) = \Omega(\boldsymbol{\Lambda}_l(\mathbf{W}_\infty)), l = 1,\ldots,L+1$. In particular, $\lim_{n\to\infty}\Omega(\boldsymbol{\Lambda}_{l_{[n]}^\star}(\mathbf{W}_{[n]})) = \Omega(\boldsymbol{\Lambda}_{l_\infty^\star}(\mathbf{W}_\infty))$. Considering both sides of Step 7 in Algorithm I when $n\to\infty$, we have $[\mathbf{p}_\infty^T 1]^T = \Omega(\boldsymbol{\Lambda}_{l_\infty^\star}(\mathbf{W}_\infty))$. Hence, $\mathbf{W}_\infty$ and $\mathbf{p}_\infty$ jointly satisfy the necessary and sufficient optimality condition in (47a)–(47b) and are optimal.

REFERENCES

[1] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation disversity—Part I: System description," *IEEE Trans. Commun.*, vol. 51, pp. 1927–1938, Nov. 2003.

[2] J. N. Laneman and G. W. Wornell, "Distributed spacetime-coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Trans. Inf. Theory*, vol. 49, pp. 2415–2425, Oct. 2003.

[3] T. Hunter, S. Sanayei, and A. Nosratinia, "Outage analysis of coded cooperation," *IEEE Trans. Inf. Theory*, vol. 52, pp. 375–391, Feb. 2006.

[4] H. Bölcskei, R. U. Nabar, Ö. Oyman, and A. J. Paulraj, "Capacity scaling laws in MIMO relay networks," *IEEE Trans. Wireless Commun.*, vol. 5, pp. 1433–1444, Jun. 2006.

[5] K. J. R. Liu, A. K. Sadek, W. Su, and A. Kwasinski, *Cooperative Communications and Networking*. New York: Cambridge Univ. Press, 2009.

[6] Ö. Oyman and A. J. Paulraj, "Power-bandwidth tradeoff in dense multi-antenna relay networks," *IEEE Trans. Wireless Commun.*, vol. 6, pp. 2282–2293, Jun. 2007.

[7] A. El-Keyi and B. Champagne, "Adaptive linearly constrained minimum variance beamforming for multiuser cooperative relaying using the Kalman filter," *IEEE Trans. Wireless Commun.*, vol. 9, pp. 641–651, Feb. 2010.

[8] S. Fazeli-Dehkordy, S. Shahbazpanahi, and S. Gazor, "Multiple peer-to-peer communications using a network of relays," *IEEE Trans. Signal Process.*, vol. 57, pp. 3053–3062, Aug. 2009.

[9] R. Krishna, K. Cumanan, X. Zhilan, and S. Lambotharan, "A novel cooperative relaying strategy for wireless networks with signal quantization," *IEEE Trans. Veh. Technol.*, vol. 59, pp. 485–489, Jan. 2010.

[10] R. H. Y. Louie, Y. Li, and B. Vucetic, "Zero forcing in general two-hop relay networks," *IEEE Trans. Veh. Technol.*, vol. 59, pp. 191–202, Jan. 2010.

[11] 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, pp. 1833–1846, Mar. 2010.

[12] B. K. Chalise and L. Vandendorpe, "MIMO relay design for multipoint-to-multipoint communications with imperfect channel state information," *IEEE Trans. Signal Process.*, vol. 57, pp. 2785–2796, Jul. 2009.

[13] S. Serbetli and A. Yener, "Relay assisted F/TDMA ad hoc networks: Node classification, power allocation and relaying strategies," *IEEE Trans. Commun.*, vol. 56, pp. 937–947, Jun. 2008.

[14] K. T. Phan, T. Le-Ngoc, S. A. Vorobyov, and C. Tellambura, "Power allocation in wireless multi-user relay networks," *IEEE Trans. Wireless Commun.*, vol. 8, pp. 2535–2545, May 2009.

[15] K. Zarifi, S. Affes, and A. Ghrayeb, "Joint source power control and relay beamforming in amplify-and-forward cognitive networks with multiple source-destination pairs," presented at the IEEE Int. Conf. Commun., Kyoto, Japan, Jun. 2011.

[16] L. Zhang, Y.-C. Liang, and Y. Xin, "Joint beamforming and power allocation for multiple access channels in cognitive radio networks," *IEEE J. Sel. Areas Commun.*, vol. 26, p. 3851, Jan. 2008.

[17] J. W. Brewer, "Kronecker products and matrix calculus in system theory," *IEEE Trans. Circuits Syst.*, vol. 25, pp. 772–781, Sep. 1978.

[18] M. Schubert and H. Boche, "Solution of the multi-user downlink beamforming problem with individual SINR constraints," *IEEE Trans. Veh. Technol.*, vol. 53, pp. 18–28, Jan. 2004.

[19] E. Karipidis, N. D. Sidiropoulos, and Z.-Q. Luo, "Quality of service and max-min fair transmit beamforming to multiple co-channel multicast groups," *IEEE Trans. Signal Process.*, vol. 56, pp. 1268–1279, Mar. 2008.

[20] K. T. Phan, S. A. Vorobyov, N. D. Sidiropoulos, and C. Tellambura, "Spectrum sharing in wireless networks via QoS-aware secondary multicast beamforming," *IEEE Trans. Signal Process.*, vol. 57, pp. 2323–2335, Jun. 2009.

[21] F. Rashid-Farrokhi, L. Tassiulas, and K. J. R. Liu, "Joint optimal power control and beamforming in wireless networks using antenna arrays," *IEEE Trans. Commun.*, vol. 46, pp. 1313–1324, Oct. 1998.

[22] A. Wiesel, Y. C. Eldar, and S. Shamai, "Linear precoding via conic optimization for fixed MIMO receivers," *IEEE Trans. Signal Process.*, vol. 54, pp. 161–176, Jan. 2006.

[23] M. Schubert and H. Boche, "A generic approach to QoS-based transceiver optimization," *IEEE Trans. Commun.*, vol. 55, pp. 1557–1566, Aug. 2007.

[24] H. Boche and M. Schubert, "A general theory for SIR balancing," *EURASIP J. Wireless Commun. Netw.*, p. 18, 2006, article ID 60681.

[25] H. Boche and M. Schubert, "On the structure of the multiuser QoS region," *IEEE Trans. Signal Process.*, vol. 44, pp. 3484–3495, Jul. 2007.

[26] S. A. Grandhi, R. Vijayan, D. J. Goodman, and J. Zander, "Centralized power control in cellular radio systems," *IEEE Trans. Veh. Technol.*, vol. 42, pp. 466–468, Nov. 1993.

[27] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1990.

[28] A. Goldsmith, S. A. Jafar, I. Marić, and S. Srinivasa, "Breaking spectrum gridlock with cognitive radios: An information theoretic perspective," *Proc. IEEE*, vol. 97, pp. 894–914, May 2009.

[29] N. Devroye, M. Vu, and V. Tarokh, "Cognitive radio networks," *IEEE Signal Process. Mag.*, vol. 25, pp. 12–23, Nov. 2008.

[30] S. Boyd and L. Vandenberghe*, Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[31] M. Grant and S. Boyd, CVX: Matlab Software for Disciplined Convex Programming. ver. Version 1.21, Apr. 2010 [Online]. Available: http://cvxr.com/cvx

[32] M. Grant and S. Boyd, "Graph implementations for nonsmooth convex programs," in *Recent Advances in Learning and Control (A Tribute to M. Vidyasagar)*, ser. Lecture Notes in Control and Information Sciences, V. Blondel, S. Boyd, and H. Kimura, Eds. New York: Springer, 2008, pp. 95–110 [Online]. Available: http://stanford.edu/boyd/graph_dcp.html

[33] H. Chen, A. B. Gershman, and S. Shahbazpanahi, "Filter-and-forward distributed beamforming in relay networks with frequency selective fading," *IEEE Trans. Signal Process.*, vol. 58, pp. 1251–1262, Mar. 2010.

[34] H. Boche and M. Schubert, "SIR balancing for multiuser downlink beamforming—Convergence analysis," in *Proc. IEEE Int. Conf. Commun. (ICC)*, New York, Apr. 2002.

[35] H. Boche and M. Schubert, "Perron-root minimization for interference-coupled systems with adaptive receive strategies," *IEEE Trans. Commun.*, vol. 57, pp. 3164–3173, Oct. 2009.

Keyvan Zarifi (S'04–M'08) received the Ph.D. degree (with the highest hons.) in electrical and computer engineering from Darmstadt University of Technology, Darmstadt, Germany, in 2007.

He has held research appointments in the Department of Communication Systems, University of Duisburg-Essen, Duisburg, Germany, in the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada, and in École Supérieure d'Électricité (Supélec), Gif-sur-Yvette, France. From September 2007 to May 2011, he was jointly with the Institut National de la Recherche Scientifique-Énergie, Matériaux, et Télécommunications (INRS-EMT), Université du Québec, and Concordia University, Montreal, QC, Canada, as a Postdoctoral Fellow. He is now a Senior Engineer in Huawei Technologies, Kanata, ON, Canada. His research interests include statistical signal processing, wireless sensor networks, MIMO and cooperative communications, and blind estimation and detection techniques.

Dr. Zarifi has received a Postdoctoral Fellowship from the Natural Sciences and Engineering Research Council of Canada (NSERC) in 2008,

Ali Ghrayeb (S'97–M'00–SM'06) received the Ph.D. degree in electrical engineering from the University of Arizona, Tucson, in 2000.

He is currently a Professor with the Department of Electrical and Computer Engineering, Concordia University, Montreal, QC, Canada. He holds a Concordia University Research Chair in Wireless Communications. He is the coauthor of the book *Coding for MIMO Communication Systems* (Wiley, 2008). His research interests include wireless and mobile communications, error correcting coding, MIMO systems, wireless cooperative networks, and cognitive radio systems.

Dr. Ghrayeb is a co-recipient of the IEEE GLOBECOM. He has instructed/co-instructed technical tutorials at several major IEEE conferences. He serves as an Associate Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS and *Physical Communications* (Elsevier). He served as an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING, the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, and the Wiley *Wireless Communications and Mobile Computing Journal*.

Sofiène Affes (S'94–M'95–SM'04) received the Diplôme d'Ingénieur degree in telecommunications in 1992 and the Ph.D. degree with honors in signal processing in 1995, both from the École Nationale Supérieure des Télécommunications (ENST), Paris, France.

He has been since with INRS-EMT, University of Quebec, Montreal, QC, Canada, as a Research Associate from 1995 until 1997, as an Assistant Professor until 2000, and then as an Associate Professor until 2009. Currently, he is a Full Professor in the Wireless Communications Group. His research interests are in wireless communications, statistical signal and array processing, adaptive space-time processing, and MIMO. From 1998 to 2002, he has been leading the radio design and signal processing activities of the Bell/Nortel/NSERC Industrial Research Chair in Personal Communications at INRS-EMT, Montreal, QC, Canada. Since 2004, he has been actively involved in major projects in wireless of PROMPT (Partnerships for Research on Microelectronics, Photonics and Telecommunications).

Prof. Affes was the corecipient of the 2002 Prize for Research Excellence of INRS. He currently holds a Canada Research Chair in Wireless Communications and a Discovery Accelerator Supplement Award from NSERC (Natural Sciences and Engineering Research Council of Canada). In 2006, he served as a General Co-Chair of the IEEE Vehicular Technology Conference (VTC) 2006-Fall, Montreal, QC, Canada. In 2008, he received from the IEEE Vehicular Technology Society the IEEE VTC Chair Recognition Award for exemplary contributions to the success of IEEE VTC. He currently acts as a member of the Editorial Board of the IEEE TRANSACTIONS ON SIGNAL PROCESSING, the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, and the Wiley *Journal on Wireless Communications & Mobile Computing*.