

# Optimal Resource Allocation for MIMO Ad Hoc Cognitive Radio Networks

Seung-Jun Kim, *Member, IEEE*, and Georgios B. Giannakis, *Fellow, IEEE*

**Abstract**—Maximization of the weighted sum-rate of secondary users (SUs) possibly equipped with multi-antenna transmitters and receivers is considered in the context of cognitive radio (CR) networks with coexisting primary users (PUs). The total interference power received at the primary receiver is constrained to maintain reliable communication for the PU. An interference channel configuration is considered for ad hoc networking, where the receivers treat the interference from undesired transmitters as noise. Without the CR constraint, a convergent distributed algorithm is developed to obtain (at least) a locally optimal solution. With the CR constraint, a semidistributed algorithm is introduced. An alternative centralized algorithm based on geometric programming and network duality is also developed. Numerical results show the efficacy of the proposed algorithms. The novel approach is flexible to accommodate modifications aiming at interference alignment. However, the stand-alone weighted sum-rate optimal schemes proposed here have merits over interference-alignment alternatives especially for practical SNR values.

**Index Terms**—Ad hoc network, cognitive radio, interference network, MIMO, optimization.

## I. INTRODUCTION

A COGNITIVE radio (CR) is an intelligent radio that senses the environment and adapts its transmission parameters to efficiently utilize the scarce radio spectrum [17], [19]. One of the promising applications of the CR technology is hierarchical spectrum sharing, where primary users (PUs), which are licensed to use certain spectrum bands, allow secondary users (SUs) to access the spectrum as long as the interference does not degrade the communication quality of the PU links [36].

Hierarchical sharing is well motivated in various settings. In the so-called *overlay* scenario, the bands not used by the PUs are identified in temporal and/or geographical (spatial) domains so that the CRs may utilize the available spectrum opportunistically. In the so-called *underlay* scenario, the interference power perceived at the PUs must be maintained under a tolerable level. For example, the ultra-wideband radios can spread their transmission power over a swath of bandwidth so that the interference is essentially buried in the background noise. Alternatively,

Manuscript received April 28, 2010; revised January 18, 2011; accepted January 21, 2011. Date of current version April 20, 2011. This work was supported by the National Science Foundation under Grants CCF 0830480, CON 014658, ECCS 0824007, and ECCS 1002180. Part of this work was presented at the 46th Annual Allerton Conference on Communication, Control, and Computing, Monticello, IL, 2008.

The authors are with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455 USA (e-mail: seungjun@umn.edu; georgios@umn.edu).

Communicated by H. El Gamal, Associate Editor for the special issue on "Interference Networks".

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

Digital Object Identifier 10.1109/TIT.2011.2120270

the CRs can perform active power control to meet the interference requirement. If transmitters are equipped with multi-antenna arrays, the interference can also be steered away from the PUs by beam-steering techniques.

The focus of the present work is on the *underlay* scenario. Specifically, an ad hoc network of CRs is considered where each node has a multi-antenna array. The multiple antennas are exploited to serve a dual objective: avoid the interference (toward both the unintended CR as well as the PU receivers), and enhance throughput via multi-input multi-output (MIMO) transmission toward the intended CR receiver. To this end, a network-wide weighted sum-rate maximization problem is formulated to optimize transmit powers, and the linear transmit filters of the CR transmitters.

It is known that even the power control problem alone (without considering MIMO transmission) for maximizing the weighted sum-rate in an interfering network is a difficult task. The challenge is that the relevant optimization problem often turns out to be nonconvex. Successive solution of geometric programming (GP) problems was proposed in [4] with promising results. In the special case of two interfering links, a simple on/off power control strategy was shown to be optimal [7]. When multicarrier transmission techniques are adopted, the dual decomposition technique was shown to yield the optimal power allocation as the number of subcarriers tends to infinity [15]. In [11], a locally optimal power allocation was obtained using a game-theoretic distributed algorithm with interference pricing.

In the case of MIMO ad hoc networks, a distributed optimization algorithm was pursued in [2] under a *network-wide* (as opposed to per-node) transmit-power constraint using the concept of network duality. A distributed algorithm was derived in [12] based on noncooperative game theory and network duality, again under a *network-wide* power constraint. Studies concerning resource optimization of MIMO CRs were also reported recently. Point-to-point MIMO links were studied in [35] under CR interference constraints, and a similar setup was considered in [34]. Notably, a distributed algorithm in the MIMO ad hoc CR network setup was proposed in [24], where nulls were placed toward the directions of the PU receivers, while each link selfishly maximized its own rate. However, due to the selfish nature of the per-link optimization metric, the algorithm cannot capture a socially efficient operating point in general. Moreover, it is not straightforward to extend the algorithm to the case where unequal weights are assigned to the individual links.

The contributions of the present work can be summarized as follows.

C1) A *distributed* weighted sum-rate maximization algorithm for MIMO ad hoc networks under per-node transmit-power constraints, attaining a Karush-Kuhn-Tucker (KKT) point of the

corresponding team optimization problem, is developed and its convergence proved. Each link in the network needs to solve a convex optimization problem per iteration, requiring only local feedback from the interfering links.

C2) A *semidistributed* algorithm is developed based on primal decomposition when the CR interference constraint is additionally present to protect PU links. The only global coordination necessary is adjustment of the permissible interference contributions of the CR links to the PU receiver, and the CR interference condition is enforced all the time during the iterative process.

C3) An alternative *centralized* algorithm for weighted sum-rate maximization for MIMO ad hoc CR networks is developed. For this, the network duality established in the literature for the MIMO multiple access and broadcast channel configurations is extended to the MIMO ad hoc network topology with per-node power constraints and CR interference constraints. Then, it is shown that iterative application of minimum-variance distortionless response (MVDR) receive-beamforming, GP-based power control, and network duality-based optimal transmit-beamforming achieves the desired KKT optimality.

Recent advances in the degrees-of-freedom characterization of MIMO interference channels (ICs) motivated development of optimization algorithms based on interference alignment [9], [20]. Although proper performance comparison is beyond the scope of the present work, it is emphasized that the weighted sum-rate maximization formulation studied here does not preclude interference alignment whenever it is feasible, and, more importantly, leads to desirable solutions in practical SNR regimes.

The paper is organized as follows. In Section II, the optimization problem is formulated. Section III presents a distributed solution without the CR interference constraints, followed by Section IV, which incorporates the CR constraints, and introduces a solution based on primal decomposition. Section V develops an alternative centralized algorithm based on network duality, and the successive GP technique. Numerical results are provided in Section VI, and conclusions are drawn in Section VII.

## II. PROBLEM STATEMENT

Consider a CR network, where  $I$  secondary users (SUs) share the spectrum with a primary user (PU). The extension to the case of more than one PUs is straightforward. The SUs form a MIMO ad hoc network, where the wireless any-to-any MIMO links interfere with each other. The transmitter of the  $i$ -th link, where  $i \in \{1, 2, \dots, I\}$ , is equipped with  $M_i$  transmit antennas, while the receiver of the same link with  $N_i$  receive antennas. At the receiver of the  $i$ -th link, the received complex baseband-equivalent signal vector can be represented as

$$\mathbf{y}_i = \mathbf{H}_{i,i}\mathbf{x}_i + \sum_{j \neq i} \mathbf{H}_{i,j}\mathbf{x}_j + \mathbf{z}_i, \quad i = 1, \dots, I \quad (1)$$

where  $\mathbf{H}_{i,j} \in \mathbb{C}^{N_i \times M_j}$  is the complex quasi-static flat fading channel matrix from the transmitter of the  $j$ -th link to the receiver of the  $i$ -th link,  $\mathbf{x}_i \in \mathbb{C}^{M_i}$  denotes the transmitted signal

vector of the  $i$ -th link, and  $\mathbf{z}_i \in \mathbb{C}^{N_i}$  stands for circularly symmetric, zero-mean, complex Gaussian noise vector with identity covariance matrix  $\mathbf{I}$ . The noise vector captures the background noise as well as the interference from outside the CR network including that from the PU transmitter, possibly after prewhitening.

In the coexistence CR network considered, SUs must be capable of preserving the performance of the incumbent PUs by dynamically adapting their transmission parameters to the channel conditions. Here, we specifically adopt the spectrum underlay architecture, where the interference power at the PU is regulated [34]. Note that this is a set-up different from the ones used in information-theoretic analyses of CR networks involving noncausal sharing of messages [13]. Aiming at reduced-complexity receivers, we further assume that joint decoding of the interfering signals is not an option; thus, the interference is treated as noise at the receivers.

The problem of interest is to maximize over the transmit-covariance matrices the weighted total achievable rate of the CR network. Define the transmit-covariance matrix of link  $i$  as  $\mathbf{Q}_i \triangleq E\{\mathbf{x}_i\mathbf{x}_i^H\}$ , where  $\mathcal{H}$  represents the Hermitian transpose. Let  $\lambda_i$  denote the nonnegative priority weight of link  $i$ , and  $\bar{P}_i$  the maximum transmit-power of the  $i$ -th link. Also, let  $\mathbf{H}_{0,i} \in \mathbb{C}^{N_0 \times M_i}$  denote the flat fading channel matrix from the transmitter of the  $i$ -th SU link to the receiver of the PU link. Then, the optimization problem is formulated as

$$(P1) \quad \max_{\{\mathbf{Q}_i \geq \mathbf{0}\}} \sum_{i=1}^I \lambda_i \log |\mathbf{I} + \mathbf{R}_i^{-1} \mathbf{H}_{i,i} \mathbf{Q}_i \mathbf{H}_{i,i}^H| \quad (2a)$$

$$\text{subject to } \text{tr}\{\mathbf{Q}_i\} \leq \bar{P}_i, \quad i \in \{1, 2, \dots, I\} \quad (2b)$$

$$\sum_{i=1}^I \text{tr}\{\mathbf{H}_{0,i} \mathbf{Q}_i \mathbf{H}_{0,i}^H\} \leq \bar{P}_0 \quad (2c)$$

where  $\mathbf{R}_i$  is the interference-plus-noise covariance matrix at the receiver of link  $i$  defined by

$$\mathbf{R}_i \triangleq \sum_{j \neq i} \mathbf{H}_{i,j} \mathbf{Q}_j \mathbf{H}_{i,j}^H + \mathbf{I} \quad (3)$$

Equation (2b) represents the per-node transmit-power constraints, and (2c) captures the condition that the total interference from the SU transmitters summed over all  $N_0$  antennas at the receiver of the PU link be capped to a certain maximum tolerable level  $\bar{P}_0$ .

Although the focus of this paper is on any-to-any CR networks, problems similar to (P1) arise naturally when applying dual decomposition to a cross-layer network optimization problem, where the weights  $\{\lambda_i\}$  play the role of Lagrange multipliers associated with the physical layer link capacity constraints; see e.g., [29].

Problem (P1) is clearly nonconvex in  $\{\mathbf{Q}_i\}$  due to the interference present. Therefore, finding the global optimum is challenging. The goal of this paper is two-fold: a) develop algorithms that yield (at least) a locally optimum solution of (P1); and b) devise distributed schemes that neither require a central

unit nor central collection of channel state information across the network.

### III. DISTRIBUTED SOLUTION

For ease of exposition, consider first the counterpart of (P1) without the CR constraint (to be revisited in Section IV):

$$(P2) \quad \max_{\{\mathbf{Q}_i \geq 0\}} \sum_{i=1}^I \lambda_i \log |\mathbf{I} + \mathbf{R}_i^{-1} \mathbf{H}_{i,i} \mathbf{Q}_i \mathbf{H}_{i,i}^H| \quad (4a)$$

$$\text{subject to } \text{tr}\{\mathbf{Q}_i\} \leq \bar{P}_i, \quad i \in \{1, 2, \dots, I\}. \quad (4b)$$

To isolate the interference which renders (P2) nonconvex, define the weighted sum-rate of the links other than the  $i$ -th as  $f_i(\mathbf{Q}_i, \mathbf{Q}_{-i}) \triangleq \sum_{j \neq i} \lambda_j \log |\mathbf{I} + \mathbf{R}_j^{-1} \mathbf{H}_{j,j} \mathbf{Q}_j \mathbf{H}_{j,j}^H|$ , where  $\mathbf{Q}_{-i}$  refers to the set of transmit-covariances of all links except the  $i$ -th; i.e.,  $\mathbf{Q}_{-i} \triangleq \{\mathbf{Q}_1, \dots, \mathbf{Q}_{i-1}, \mathbf{Q}_{i+1}, \dots, \mathbf{Q}_I\}$ . The negative partial derivative of  $f_i(\cdot)$  w.r.t.  $\mathbf{Q}_i$ , evaluated at  $\mathbf{Q}_i = \bar{\mathbf{Q}}_i$ , is then given by (5), at the bottom of the page. Retaining only the linear term in the Taylor's expansion of  $f_i(\cdot)$  around  $\bar{\mathbf{Q}}_i$ , it is possible to approximate (P2) by a set of  $I$  per-link problems given for  $i \in \{1, 2, \dots, I\}$  by [cf. (4a) and (5)]

$$(P3) \quad \max_{\mathbf{Q}_i \geq 0} \lambda_i \log |\mathbf{I} + \mathbf{R}_i^{-1} \mathbf{H}_{i,i} \mathbf{Q}_i \mathbf{H}_{i,i}^H| - \text{tr}\{\mathbf{A}_i^H \mathbf{Q}_i\} \quad (6a)$$

$$\text{subject to } \text{tr}\{\mathbf{Q}_i\} \leq \bar{P}_i. \quad (6b)$$

Unlike (P2), (P3) is convex in  $\mathbf{Q}_i$  and can be efficiently solved by numerical iterative algorithms ( $\bar{\mathbf{Q}}_i$  needed for  $\mathbf{A}_i$  is thus available from the previous iteration). Note that (P3) essentially maximizes the same objective function as (P2) with respect to  $\mathbf{Q}_i$ , except that the weighted sum-rate of the other links is approximated to the first order at the point  $\bar{\mathbf{Q}}_i$ . The trace term in (6a) plays the role of interference tax, discouraging selfish behavior of link  $i$ , which would otherwise just want to maximize its own rate. This is in the spirit of [12], [11], [33], where similar techniques were employed in different contexts.

The MIMO optimization identifies the optimal transmit-covariance matrices, which also determines the power allocation among the multiple streams of each link. In the special case where the channel matrices have circulant structure, iterative water-filling-type approaches have been widely studied [31], [25]. An approach related to the present formulation is [23],

where a noncooperative game is studied for the MIMO IC. However, since each link selfishly maximizes its own rate in [23] (that is,  $\mathbf{A}_i$  is set to  $\mathbf{0}$  in (6a)), the attained Nash equilibrium may not be socially efficient. Moreover, it is not clear how a *weighted* sum-rate maximization problem should be handled using the approach in [23] when the individual link rates are not equally weighted.

*Remark 1:* For each link  $i$  to solve (P3), it has to compute the matrix  $\mathbf{A}_i$ . As can be seen from (5), computing  $\mathbf{A}_i$  requires feedback from all links  $j \neq i$  of the Hermitian matrices  $\mathbf{B}_j \triangleq \lambda_j [\mathbf{R}_j^{-1} - (\mathbf{R}_j + \mathbf{H}_{j,j} \mathbf{Q}_j \mathbf{H}_{j,j}^H)^{-1}]$ , as well as the channels  $\mathbf{H}_{j,i}$ . The feedback may be efficiently implemented via wireless broadcast. Computing  $\mathbf{B}_j$  at link  $j$  does not require feedback from other links since  $\mathbf{R}_j$  can be estimated at link  $j$  through measurements. In a time-division duplex (TDD) system, the channels may be estimated at the transmitter of link  $i$  thanks to reciprocity. Furthermore, if  $N_j = 1$ , then  $\mathbf{B}_j$  becomes a scalar. Note that in a centralized scheme, one would normally need to collect at a central processor the MIMO channels of *all* links in the network (including the interfering links), presumably relying on multihop routing to convey channel feedback from the regions far from the central processor. On the contrary, the proposed distributed scheme requires each link  $i$  to collect only *local* channel information (i.e., the channels for the links that impinge upon the receiver of link  $i$ ), and thus possesses improved scalability and robustness properties.

*Remark 2:* There are avenues to further reduce the feedback overhead in practice. First, the feedback information may be transmitted only when it has changed significantly. Since the channels are assumed quasi-static, they do not need to be fed back in each iteration. Moreover, redundant feedback can be eliminated as the algorithm nears to convergence. Even when the channels vary slowly, warm-starting the algorithm using the last solution as an initial point can considerably reduce feedback requirements. Secondly, although the algorithm formally requires each link  $i$  to collect feedback from all links  $j \neq i$  in the network, only nearby links inflict sizable contribution of interference in practice. Thus, the feedback may be collected only in the vicinity, which constitutes an important practical advantage of the distributed algorithm in terms of scalability, compared to centralized alternatives. Note that in the centralized algorithm, all channels must be collected at a central processing unit, and the computed transmission parameters need to be fed back to the appropriate links, with the locality structure completely neglected.

$$\begin{aligned} \mathbf{A}_i &\triangleq - \left. \frac{\partial f_i}{\partial \mathbf{Q}_i} \right|_{\mathbf{Q}_i = \bar{\mathbf{Q}}_i} \\ &= \sum_{j \neq i} \lambda_j \mathbf{H}_{j,i}^H \left[ \mathbf{R}_j^{-1} - (\mathbf{R}_j + \mathbf{H}_{j,j} \mathbf{Q}_j \mathbf{H}_{j,j}^H)^{-1} \right] \mathbf{H}_{j,i} \Big|_{\mathbf{Q}_i = \bar{\mathbf{Q}}_i}. \end{aligned} \quad (5)$$

### A. Solution of (P3)

Leaving the nonnegative definiteness constraint on  $\mathbf{Q}_i$  implicit, the partial Lagrangian of (P3) is given by

$$\mathcal{L}_i(\mathbf{Q}_i, \mu_i) = \lambda_i \log \left| \mathbf{I} + \mathbf{R}_i^{-1} \mathbf{H}_{i,i} \mathbf{Q}_i \mathbf{H}_{i,i}^H \right| - \text{tr} \{ (\mathbf{A}_i + \mu_i \mathbf{I}) \mathbf{Q}_i \} + \mu_i \bar{P}_i \quad (7)$$

where  $\mu_i$  is the Lagrange multiplier corresponding to the power constraint. The dual function is

$$\mathcal{D}_i(\mu_i) = \max_{\mathbf{Q}_i \geq \mathbf{0}} \mathcal{L}_i(\mathbf{Q}_i, \mu_i). \quad (8)$$

Then the optimal solution can be found by solving the dual problem

$$\min_{\mu_i \geq 0} \mathcal{D}_i(\mu_i). \quad (9)$$

Upon factorizing  $\mathbf{Q}_i = \mathbf{G}_i \mathbf{G}_i^H$ , the maximization problem in (8) can be rewritten as

$$\max_{\mathbf{G}_i} \lambda_i \log \left| \mathbf{I} + \mathbf{G}_i^H \mathbf{H}_{i,i}^H \mathbf{R}_i^{-1} \mathbf{H}_{i,i} \mathbf{G}_i \right| - \text{tr} \{ \mathbf{G}_i^H (\mathbf{A}_i + \mu_i \mathbf{I}) \mathbf{G}_i \}. \quad (10)$$

*Proposition 1:* The optimal  $\mathbf{G}_i$  solving (10) is a generalized eigen-matrix of the pair of matrices  $\mathbf{H}_{i,i}^H \mathbf{R}_i^{-1} \mathbf{H}_{i,i}$  and  $(\mathbf{A}_i + \mu_i \mathbf{I})$ .

*Proof:* The proof relies on simplifying (10) until Hadamard's inequality applies as in e.g., [10]. The Cholesky decomposition of the matrix  $(\mathbf{A}_i + \mu_i \mathbf{I})$  is written as  $\mathbf{L}_i \mathbf{L}_i^H$ , where  $\mathbf{L}_i$  is a lower-triangular Cholesky factor. Upon defining  $\tilde{\mathbf{G}}_i \triangleq \mathbf{L}_i^H \mathbf{G}_i$ , (10) reduces to

$$\max_{\tilde{\mathbf{G}}_i} \lambda_i \log \left| \mathbf{I} + \tilde{\mathbf{G}}_i^H \mathbf{L}_i^{-1} \mathbf{H}_{i,i}^H \mathbf{R}_i^{-1} \mathbf{H}_{i,i} \mathbf{L}_i^{-H} \tilde{\mathbf{G}}_i \right| - \text{tr} \{ \tilde{\mathbf{G}}_i \tilde{\mathbf{G}}_i^H \}. \quad (11)$$

The eigen-decomposition of  $\mathbf{L}_i^{-1} \mathbf{H}_{i,i}^H \mathbf{R}_i^{-1} \mathbf{H}_{i,i} \mathbf{L}_i^{-H}$  is written as  $\mathbf{U}_i \boldsymbol{\Sigma}_i \mathbf{U}_i^H$ , where  $\mathbf{U}_i$  is unitary and  $\boldsymbol{\Sigma}_i$  is diagonal. If  $\tilde{\mathbf{Q}}_i \triangleq \mathbf{U}_i^H \tilde{\mathbf{G}}_i \tilde{\mathbf{G}}_i^H \mathbf{U}_i$ , then (11) can be expressed as

$$\max_{\tilde{\mathbf{Q}}_i \geq \mathbf{0}} \lambda_i \log \left| \mathbf{I} + \tilde{\mathbf{Q}}_i \boldsymbol{\Sigma}_i \right| - \text{tr} \{ \tilde{\mathbf{Q}}_i \}. \quad (12)$$

By Hadamard's inequality [6, p. 233], it can be seen that the optimal  $\tilde{\mathbf{Q}}_i$  must be diagonal; hence,  $\mathbf{G}_i = \mathbf{L}_i^{-H} \mathbf{U}_i \tilde{\mathbf{Q}}_i^{1/2}$ , and thus

$$\mathbf{H}_{i,i}^H \mathbf{R}_i^{-1} \mathbf{H}_{i,i} \mathbf{G}_i = \mathbf{L}_i \mathbf{L}_i^H \mathbf{L}_i^{-H} \mathbf{U}_i \tilde{\mathbf{Q}}_i^{1/2} \boldsymbol{\Sigma}_i = (\mathbf{A}_i + \mu_i \mathbf{I}) \mathbf{G}_i \boldsymbol{\Sigma}_i \quad (13)$$

which proves the result. ■

Proposition 1 specifies the directions of the optimal precoders (beamformers)  $\mathbf{G}_i$ , but not the optimal power allocated in each direction, since right-multiplying the eigen-matrix  $\tilde{\mathbf{G}}_i$  by a diagonal matrix  $\boldsymbol{\Delta}_i$  again yields an eigen-matrix. (Indeed (13) is satisfied after swapping  $\boldsymbol{\Sigma}_i$  and  $\boldsymbol{\Delta}_i$ .) To obtain the optimal power allocation, consider the normalized precoding matrix  $\bar{\mathbf{G}}_i$  such that  $\mathbf{G}_i = \bar{\mathbf{G}}_i \mathbf{P}_i^{1/2}$ , where each column of  $\bar{\mathbf{G}}_i$  has norm equal

TABLE I  
OVERALL ALGORITHM TO SOLVE (P2)

Initialize $\mathbf{Q}_i = \mathbf{0}$ for all $i$ . Repeat until convergence For $i = 1, 2, \dots, I$ Compute $\mathbf{A}_i$ with $\bar{\mathbf{Q}}_i = \mathbf{Q}_i$ . Set $\underline{\mu}_i = 0, \bar{\mu}_i = \mu'_{i,\max}$ . Repeat until convergence $\mu_i = (\underline{\mu}_i + \bar{\mu}_i)/2$ . Compute the generalized eigenmatrix $\mathbf{G}_i$ and eigenvalues $\boldsymbol{\Sigma}_i$ of $\mathbf{H}_{i,i}^H \mathbf{R}_i^{-1} \mathbf{H}_{i,i}$ and $(\mathbf{A}_i + \mu_i \mathbf{I})$ . Compute the normalized eigenmatrix $\bar{\mathbf{G}}_i$ . Compute $\boldsymbol{\Sigma}_i^{(1)} = \bar{\mathbf{G}}_i^H \mathbf{H}_{i,i}^H \mathbf{R}_i^{-1} \mathbf{H}_{i,i} \bar{\mathbf{G}}_i$ . Compute $\mathbf{P}_i = (\lambda_i \boldsymbol{\Sigma}_i - \mathbf{I})^+ \boldsymbol{\Sigma}_i^{(1)-1}$ . If $\text{tr} \{ \mathbf{P}_i \} > \bar{P}_i$ , set $\underline{\mu}_i = \mu_i$ ; otherwise, set $\bar{\mu}_i = \mu_i$ . Set $\mathbf{Q}_i = \bar{\mathbf{G}}_i \mathbf{P}_i \bar{\mathbf{G}}_i^H$ . Next $i$
---

to unity, and  $\mathbf{P}_i^{1/2}$  is diagonal and nonnegative. Since  $\bar{\mathbf{G}}_i$  is the generalized eigen-matrix of  $\mathbf{H}_{i,i}^H \mathbf{R}_i^{-1} \mathbf{H}_{i,i}$  and  $(\mathbf{A}_i + \mu_i \mathbf{I})$ , so is  $\bar{\mathbf{G}}_i$ . It then follows readily from [8, p. 462] that

$$\begin{aligned} \bar{\mathbf{G}}_i^H \mathbf{H}_{i,i}^H \mathbf{R}_i^{-1} \mathbf{H}_{i,i} \bar{\mathbf{G}}_i &= \boldsymbol{\Sigma}_i^{(1)} \text{ and} \\ \bar{\mathbf{G}}_i^H (\mathbf{A}_i + \mu_i \mathbf{I}) \bar{\mathbf{G}}_i &= \boldsymbol{\Sigma}_i^{(2)} \end{aligned} \quad (14)$$

where  $\boldsymbol{\Sigma}_i^{(1)}$  and  $\boldsymbol{\Sigma}_i^{(2)}$  are diagonal. From (13), it also holds that  $\boldsymbol{\Sigma}_i^{(1)} = \boldsymbol{\Sigma}_i \boldsymbol{\Sigma}_i^{(2)}$ . Then, (10) can be written as

$$\max_{\mathbf{P}_i \geq \mathbf{0}} \lambda_i \log \left| \mathbf{I} + \mathbf{P}_i \boldsymbol{\Sigma}_i^{(1)} \right| - \text{tr} \left\{ \mathbf{P}_i \boldsymbol{\Sigma}_i^{(2)} \right\}. \quad (15)$$

The solution of (15) is given by (see also [6, p. 252])

$$\begin{aligned} \mathbf{P}_i(k, k) &= \frac{1}{\boldsymbol{\Sigma}_i^{(1)}(k, k)} \left( \frac{\lambda_i \boldsymbol{\Sigma}_i^{(1)}(k, k)}{\boldsymbol{\Sigma}_i^{(2)}(k, k)} - 1 \right)^+ \\ &= \left( \frac{\lambda_i \boldsymbol{\Sigma}_i(k, k) - 1}{\boldsymbol{\Sigma}_i^{(1)}(k, k)} \right)^+, \\ &\text{for all } k \text{ s. t. } \boldsymbol{\Sigma}_i^{(1)}(k, k) > 0 \end{aligned} \quad (16)$$

where  $(\cdot)^+ \triangleq \max\{\cdot, 0\}$ , and the remaining elements of  $\mathbf{P}_i$  are zero.

To solve (P3), one needs to find the optimal Lagrange multiplier  $\mu_i$  by solving (9). This can be accomplished using a simple bisection search.

### B. Convergence

The overall algorithm essentially allows each link to perform the optimization in (P3) autonomously until convergence. The Gauss-Seidel iteration, the Jacobi iteration, or an entirely asynchronous iteration may be adopted to this end [18]. Table I lists a Gauss-Seidel version of the iterative algorithm. The ensuing proposition links the solution of the original problem (P2) with that of its first-order approximation (P3).

*Proposition 2:* A fixed point of the algorithm in Table I exists, and it is a KKT point of (P2). Thus, if the algorithm converges, it converges to a KKT point of (P2).

*Proof:* The result is immediate upon checking that the set of KKT conditions for (P3) for  $i \in \{1, 2, \dots, I\}$  constitute precisely the KKT conditions for (P2). ■

Convergence of the algorithm when the Gauss-Seidel iteration is applied is proved in the following proposition. The proof is inspired by [26], which established convergence of a similar algorithm for single-input single-output and multi-input single-output ad hoc network configurations. In the MIMO case, the following lemma is necessary.

*Lemma 1:* For each  $i \in \{1, 2, \dots, I\}$ , the function  $f_i(\mathbf{Q}_i, \mathbf{Q}_{-i})$  is convex w.r.t.  $\mathbf{Q}_i \in \mathcal{Q}_i \triangleq \{\mathbf{Q}_i | \mathbf{Q}_i \geq \mathbf{0}, \text{tr}\{\mathbf{Q}_i\} \leq \bar{P}_i\}$ .

*Proof:* Adopting the technique used in [30, Th. 2], convexity of  $f_i(\cdot)$  w.r.t.  $\mathbf{Q}_i$  can be proved by showing that  $\tilde{f}_i(t) \triangleq f_i(\mathbf{X}_i + t\mathbf{Y}_i, \mathbf{Q}_{-i})$  is convex w.r.t.  $t \in [0, 1]$ , where  $\mathbf{X}_i, \mathbf{Y}_i \in \mathcal{Q}_i$ . The second derivative of  $\tilde{f}_i(t)$  is (see also [30])

$$\frac{d^2 \tilde{f}_i(t)}{dt^2} = \sum_{j \neq i} \lambda_j \text{tr} \{ (\mathbf{R}_j + \mathbf{J}_j)^{-1} \mathbf{K}_j \mathbf{M}_j \mathbf{K}_j + \mathbf{M}_j \mathbf{K}_j \mathbf{R}_j^{-1} \mathbf{K}_j \} \quad (17)$$

where  $\mathbf{J}_j \triangleq \mathbf{H}_{j,j} \mathbf{Q}_j \mathbf{H}_{j,j}^H$ ,  $\mathbf{K}_j \triangleq \mathbf{H}_{j,i} \mathbf{Y}_i \mathbf{H}_{j,i}^H$ , and  $\mathbf{M}_j \triangleq (\mathbf{R}_j + \mathbf{J}_j)^{-1} \mathbf{J}_j \mathbf{R}_j^{-1}$ . Note that  $(\mathbf{R}_j + \mathbf{J}_j)^{-1} \geq \mathbf{0}$  and  $\mathbf{M}_j = \mathbf{R}_j^{-1} - (\mathbf{R}_j + \mathbf{J}_j)^{-1} \geq \mathbf{0}$ . Thus, using matrix factors  $\mathbf{C}_j$  and  $\mathbf{D}_j$ , it is possible to write  $(\mathbf{R}_j + \mathbf{J}_j)^{-1} = \mathbf{C}_j \mathbf{C}_j^H$ , and  $\mathbf{M}_j = \mathbf{D}_j \mathbf{D}_j^H$ . Then, it can be seen that (cf. (17))

$$\frac{d^2 \tilde{f}_i(t)}{dt^2} = \sum_{j \neq i} \lambda_j [\text{tr} \{ \mathbf{C}_j^H \mathbf{K}_j \mathbf{M}_j \mathbf{K}_j \mathbf{C}_j \} + \text{tr} \{ \mathbf{D}_j^H \mathbf{K}_j \mathbf{R}_j^{-1} \mathbf{K}_j \mathbf{D}_j \}] \geq 0. \quad (18)$$

■

*Proposition 3:* The Gauss-Seidel iteration in Table I converges.

*Proof:* The claim will be established by showing that after a link  $i$  solves (P3), the network-wide objective function of (P2) is nondecreasing. Suppose that  $\mathbf{Q}_j = \bar{\mathbf{Q}}_j$  for all  $j \in \{1, 2, \dots, I\}$ , from the previous iteration. Let  $\mathbf{Q}_i^*$  denote the optimal solution of (P3) for link  $i$ . Define  $r_j(\mathbf{Q}_i, \mathbf{Q}_{-i}) \triangleq \lambda_j \log |\mathbf{I} + \mathbf{R}_j^{-1} \mathbf{H}_{j,j} \mathbf{Q}_j \mathbf{H}_{j,j}^H|$ , for all  $j$ . Then,

$$\begin{aligned} & \sum_{j=1}^I r_j(\mathbf{Q}_i^*, \bar{\mathbf{Q}}_{-i}) \\ &= r_i(\mathbf{Q}_i^*, \bar{\mathbf{Q}}_{-i}) + f_i(\mathbf{Q}_i^*, \bar{\mathbf{Q}}_{-i}) \end{aligned} \quad (19a)$$

$$\geq r_i(\mathbf{Q}_i^*, \bar{\mathbf{Q}}_{-i}) + f_i(\bar{\mathbf{Q}}_i, \bar{\mathbf{Q}}_{-i}) - \text{tr} \{ \mathbf{A}_i^H (\mathbf{Q}_i^* - \bar{\mathbf{Q}}_i) \} \quad (19b)$$

$$\geq r_i(\bar{\mathbf{Q}}_i, \bar{\mathbf{Q}}_{-i}) + f_i(\bar{\mathbf{Q}}_i, \bar{\mathbf{Q}}_{-i}) - \text{tr} \{ \mathbf{A}_i^H (\bar{\mathbf{Q}}_i - \bar{\mathbf{Q}}_i) \} \quad (19c)$$

$$= \sum_{j=1}^I r_j(\bar{\mathbf{Q}}_i, \bar{\mathbf{Q}}_{-i}) \quad (19d)$$

where (19b) is a consequence of Lemma 1, and the definition of  $\mathbf{A}_i$  in (5); and (19c) holds since  $\mathbf{Q}_i^*$  is the optimal solution of (P3). Thus, the network-wide objective  $\sum_{j=1}^I r_j(\mathbf{Q}_i, \mathbf{Q}_{-i})$  is

nondecreasing after each link  $i$  updates  $\mathbf{Q}_i$ . Since the objective is bounded from above, the algorithm must converge. ■

#### IV. APPLICATION TO COGNITIVE RADIO NETWORKS

Let us return to the original formulation (P1) with the CR constraint (2c). The CR constraint couples all SU transmitters in the network. To decouple them so as to develop a distributed solution, one may consider the dual decomposition technique to relax the CR constraint. However, we aim at an on-line algorithm that iteratively attains the optimal solution as the network actually transmits and receives signals. In other words, the system will employ the values of intermediate iterates for network operation, while the algorithm may be yet converging. In contrast, an off-line algorithm will first obtain the final solution by running the iterative algorithm until convergence, and subsequently deploy the solution for network operation. In addition to distributing the computational burden over time, the on-line algorithm is often capable of coping with slow variations in the environment (e.g., mobile channels) by tracking optimal solutions over time. However, to run an algorithm in an on-line fashion, one must ensure that the intermediate iterates are feasible for network operation.

In this respect, the dual decomposition may not be appropriate because the primal variables become feasible only at convergence. Thus, while iterations proceed but have not yet converged, the CR constraint (2c) may well be violated if the dual decomposition approach is taken. In fact, the algorithm may not converge to a feasible solution of the underlying nonconvex problem (P1).

##### A. Primal Decomposition Approach

An alternative is to use the primal decomposition, which has been also adopted in various contexts; see e.g., [18]. In this approach, a master problem determines allocation of the shared resource, thus preventing violation of the constraint. In our setup, this amounts to introducing a set of auxiliary variables  $\{\bar{P}_{0,i}\}$  such that  $\bar{P}_{0,i} \geq 0$  and  $\sum_{i=1}^I \bar{P}_{0,i} = \bar{P}_0$ . Then, (P1) can be written as (20), at the bottom of the page. The master problem is then formulated as

$$(P5) \quad \max_{\{\bar{P}_{0,i}\}} g(\{\bar{P}_{0,i}\}) \quad (21a)$$

$$\text{subject to } \sum_{i=1}^I \bar{P}_{0,i} \leq \bar{P}_0 \quad (21b)$$

$$\bar{P}_{0,i} \geq 0, \quad i \in \{1, 2, \dots, I\}. \quad (21c)$$

##### B. Semidistributed Solution

Now, taking the same approach as in Section III, one can find at least a local optimum of (P4) in a distributed fashion by solving the per-link problems for each  $i = 1, 2, \dots, I$ ; that is,

$$(P6) \quad \max_{\mathbf{Q}_i \geq \mathbf{0}} \lambda_i \log |\mathbf{I} + \mathbf{R}_i^{-1} \mathbf{H}_{i,i} \mathbf{Q}_i \mathbf{H}_{i,i}^H| - \text{tr} \{ \mathbf{A}_i^H \mathbf{Q}_i \} \quad (22a)$$

$$\text{subject to } \text{tr} \{ \mathbf{Q}_i \} \leq \bar{P}_i \quad (22b)$$

TABLE II  
OVERALL ALGORITHM FOR SOLVING (P1)

1:	Initialize $\bar{P}_{0,i} = \bar{P}_0/I$ for all $i$
2:	Repeat until convergence:
2:	Feedback $\{\bar{P}_{0,i}\}$ to each link $i$
3:	Solve (P6) in a distributed manner; see Sec. IV-B
4:	Collect optimal $\{\mu_{0,i}\}_{i=1}^I$ at a central processor
5:	Update $\{\bar{P}_{0,i}\}$ via (24)

$$\text{tr}\{\mathbf{H}_{0,i}\mathbf{Q}_i\mathbf{H}_{0,i}^H\} \leq \bar{P}_{0,i}. \quad (22c)$$

The partial Lagrangian of (P6) is given by [cf. (7)]

$$\begin{aligned} \mathcal{L}'_i(\mathbf{Q}_i, \mu_i, \mu_{0,i}) &= \lambda_i \log |\mathbf{I} + \mathbf{R}_i^{-1}\mathbf{H}_{i,i}\mathbf{Q}_i\mathbf{H}_{i,i}^H| \\ &- \text{tr}\{(\mathbf{A}_i + \mu_i\mathbf{I} + \mu_{0,i}\mathbf{H}_{0,i}^H\mathbf{H}_{0,i})\mathbf{Q}_i\} + \mu_i\bar{P}_i + \mu_{0,i}\bar{P}_{0,i} \end{aligned} \quad (23)$$

where  $\mu_{0,i}$  is the Lagrange multiplier associated with the CR constraint in (22c). Note that (23) has essentially the same form as (7), except that  $(\mathbf{A}_i + \mu_i\mathbf{I})$  is replaced by  $(\mathbf{A}_i + \mu_i\mathbf{I} + \mu_{0,i}\mathbf{H}_{0,i}^H\mathbf{H}_{0,i})$ , and that there are two Lagrange multipliers rather than one. Thus, an algorithm analogous to that in Table I can be developed to solve (P4). The bisection search in Table I can be replaced by the ellipsoid algorithm [1].

To solve (P5), we rely on the subgradient algorithm. The subgradient of  $g(\cdot)$  with respect to  $\bar{P}_{0,i}$  is given by the optimal  $\mu_{0,i}$  obtained from the solution of (P6). Thus, the update for  $\{\bar{P}_{0,i}\}$  is

$$\begin{aligned} \bar{P}_{0,i}(n+1) &= \mathcal{P}\{\bar{P}_{0,i}(n) + \alpha_n\mu_{0,i}(n)\}, \\ &\quad \forall i = 1, 2, \dots, I \end{aligned} \quad (24)$$

where  $n$  denotes the iteration index,  $\alpha_n$  the step size, and  $\mathcal{P}\{\cdot\}$  the projection onto the region defined by (21b)–(21c). An efficient algorithm to compute such a projection is detailed in [16].

The overall procedure for solving (P1) is summarized in Table II. Although the convergence of the algorithm could not be established due to nonconvexity of (P4), the algorithm always converged in the extensive numerical tests that we performed. It is easy to see that once the algorithm converges, it converges to a local optimum of (P4) and (P5).

*Remark 3:* Clearly, the projection operation needs to be performed in a centralized fashion. This can be done either by a “head” node elected among the SU transceivers, or, by a specialized controller. The controller needs to collect the optimal dual prices  $\{\mu_{0,i}\}$  from the SU transmitters, and feed the allowed interference contributions  $\{\bar{P}_{0,i}\}$  back to them. The numerical tests in Section VI demonstrate that optimizing over  $\{\bar{P}_{0,i}\}$  often results in negligible increase in the objective, especially when the number of antennas is large. The rationale is that as the spatial degrees of freedom increase, the CRs (even the

ones with significantly strong links to the PUs) can simply steer sharp nulls toward the PU directions without incurring much cost in terms of the power resources. The weighted sum-rate objective will be dictated essentially by the per-node power constraints in this power-limited regime. Therefore, one may just set  $P_{0,i} = \bar{P}_0/I$  for all  $i$  to obtain a fully (as opposed to semi-) distributed algorithm without much degradation of the reward when the number of antennas is large.

*Remark 4:* A remark is due on the feedback requirement of this algorithm. Solving (P6) at link  $i$  requires collecting channel  $\mathbf{H}_{0,i}$  locally and the interference power cap  $\bar{P}_{0,i}$  from a central processor, in addition to the feedback requirement already described in Remark 1. For solving (P5), the optimal Lagrange multipliers  $\{\mu_{0,i}\}$  must be collected at a central unit. Since the master problem iterations in Table II run in a time scale larger than that of the per-link iterations, the message passing to/from the central processor occurs in a much slower pace than the feedback related to the per-link problems. Moreover, the central message passing involves merely two scalar quantities  $\mu_{0,i}$  and  $\bar{P}_{0,i}$  per link. This requires far less overhead when compared to collecting all MIMO channels at a central unit. The performance of the latter case is studied in the next section.

Next, an alternative centralized algorithm is developed for weighted sum-rate maximization of MIMO ad hoc CR networks. It also achieves local optimality, but may yield higher reward than its distributed counterpart in practice, as will be confirmed by the numerical tests in Section VI.

## V. GEOMETRIC PROGRAMMING AND NETWORK DUALITY

In this section, an alternative algorithm is developed to solve (P1) in a centralized manner using GP and network duality. In spite of the inherent nonconvexity of the underlying optimization problem, GP-based power control algorithms for single-antenna transceivers have been reported to achieve excellent performance in practice [4], [21]. On the other hand, network duality proved extremely useful in optimizing MIMO networks [2], [12], [32]. Our approach here is to leverage both of these techniques by dividing the overall nonconvex problem into smaller subproblems, each of which can be efficiently solved.

Being a centralized algorithm, the proposed approach requires the channel states of all links in the network to be first collected at a central processor. This is feasible for networks of small size. However, even when such an arrangement is impractical, performance of the centralized algorithm provides a benchmark for the (semi-)distributed algorithm developed in the preceding sections. Moreover, a subproblem to be solved as part of the algorithm, namely (P10), is an interesting problem

$$(P4) \quad g(\{\bar{P}_{0,i}\}) \triangleq \max_{\{\mathbf{Q}_i \geq \mathbf{0}\}} \sum_{i=1}^I \lambda_i \log |\mathbf{I} + \mathbf{R}_i^{-1}\mathbf{H}_{i,i}\mathbf{Q}_i\mathbf{H}_{i,i}^H| \quad (20a)$$

$$\text{subject to } \text{tr}\{\mathbf{Q}_i\} \leq \bar{P}_i, \quad i \in \{1, 2, \dots, I\} \quad (20b)$$

$$\text{tr}\{\mathbf{H}_{0,i}\mathbf{Q}_i\mathbf{H}_{0,i}^H\} \leq \bar{P}_{0,i}, \quad i \in \{1, 2, \dots, I\}. \quad (20c)$$

in its own right that has not been previously addressed in the literature.

### A. Iterative Algorithm

In order to exploit network duality, one first needs to reformulate the problem in terms of linear transmit- and receive-precoders. By factorizing  $\mathbf{Q}_i = \mathbf{G}_i \mathbf{G}_i^H$  for all  $i$ , (P1) can be equivalently written as

$$(P7) \quad \max_{\{\mathbf{G}_i\}} \sum_{i=1}^I \lambda_i \log |\mathbf{I} + \mathbf{G}_i^H \mathbf{H}_{i,i}^H \mathbf{R}_i^{-1} \mathbf{H}_{i,i} \mathbf{G}_i| \quad (25a)$$

$$\text{subject to } \text{tr} \{ \mathbf{G}_i \mathbf{G}_i^H \} \leq \bar{P}_i, \quad \forall i \in \{1, 2, \dots, I\} \quad (25b)$$

$$\sum_{i=1}^I \text{tr} \{ \mathbf{H}_{0,i} \mathbf{G}_i \mathbf{G}_i^H \mathbf{H}_{0,i}^H \} \leq \bar{P}_0. \quad (25c)$$

The following result adapted from [2, Th. 1] allows us to express (P7) in terms of linear precoders.

*Proposition 4:* For any given  $\{\mathbf{G}_i\}_{i=1}^I$ , there exist matrices  $\{\tilde{\mathbf{G}}_i\}_{i=1}^I$  that yield the same objective as in (25a) while maintaining  $\tilde{\mathbf{G}}_i \tilde{\mathbf{G}}_i^H = \mathbf{G}_i \mathbf{G}_i^H = \mathbf{Q}_i$ ,  $\forall i$ . Moreover, the same objective can be attained by linear transmit- and receive-beamforming, where the transmit-beamformers for link  $i$  are the columns of  $\tilde{\mathbf{G}}_i$  denoted by  $\tilde{\mathbf{g}}_{i,m_i} \in \mathbb{C}^{M_i}$ , and the receive-beamformers are computed as the MVDR beamformers given by  $\tilde{\mathbf{w}}_{i,m_i} = \mathbf{R}_{i,m_i}^{-1} \mathbf{H}_{i,i} \tilde{\mathbf{g}}_{i,m_i}$ , with  $\mathbf{R}_{i,m_i} \triangleq \mathbf{I} + \sum_{(j,m_j) \neq (i,m_i)} \mathbf{H}_{i,j} \tilde{\mathbf{g}}_{j,m_j} \tilde{\mathbf{g}}_{j,m_j}^H \mathbf{H}_{i,j}^H$ ,  $m_i = 1, 2, \dots, M_i$ .

Note that  $\mathbf{R}_{i,m_i}$  is related to  $\mathbf{R}_i$  as  $\mathbf{R}_{i,m_i} = \mathbf{R}_i + \sum_{l_i \neq m_i} \mathbf{H}_{i,i} \tilde{\mathbf{g}}_{i,l_i} \tilde{\mathbf{g}}_{i,l_i}^H \mathbf{H}_{i,i}^H$ , i.e.,  $\mathbf{R}_{i,m_i}$  contains the terms from the interstream interference. Proposition 4 essentially asserts that the interstream interference can be completely eliminated at every link in the network by linear transmit- and receive-beamforming.

Based on Proposition 4, it can be seen that the following problem achieves the same optimal objective as (P7)—see (26) at the bottom of the page, where  $\mathbf{w}_{i,m_i} = \tilde{\mathbf{w}}_{i,m_i} / \|\tilde{\mathbf{w}}_{i,m_i}\|$ ,  $\forall i, m_i$ .

Let  $P_{i,m_i} \triangleq \|\tilde{\mathbf{g}}_{i,m_i}\|^2$  denote the transmit-power of the  $m_i$ -th stream in link  $i$ , and define  $\mathbf{g}_{i,m_i} \triangleq \tilde{\mathbf{g}}_{i,m_i} / \|\tilde{\mathbf{g}}_{i,m_i}\|$ . The algorithm to be presented consists mainly of two components: i) power control via GP given the transmit-precoders  $\{\mathbf{g}_{i,m_i}\}$ ; and ii) network duality-based precoder updates; see also [32]. Related algorithms were studied in [5] and [27] for a MIMO cellular downlink setup without formal claims of optimality. Specifically, the following iterative algorithm is proposed:

**S1)** Given the transmit-beamformers  $\{\tilde{\mathbf{g}}_{i,m_i}\}$ , determine the receive-beamformers  $\{\mathbf{w}_{i,m_i}\}$  as the normalized MVDR beamformers

$$\mathbf{w}_{i,m_i} = \frac{\mathbf{R}_{i,m_i}^{-1} \mathbf{H}_{i,i} \tilde{\mathbf{g}}_{i,m_i}}{\|\mathbf{R}_{i,m_i}^{-1} \mathbf{H}_{i,i} \tilde{\mathbf{g}}_{i,m_i}\|}, \quad \forall i, m_i. \quad (27)$$

**S2)** Given  $\{\mathbf{w}_{i,m_i}\}$  and  $\{\mathbf{g}_{i,m_i}\}$ , determine the (locally) optimal transmit-powers  $\{P_{i,m_i}\}$  and the per-stream SINRs  $\{\gamma_{i,m_i}\}$  for (P9), at the bottom of the next page.

**S3)** Given  $\{\mathbf{w}_{i,m_i}\}$  and  $\{\gamma_{i,m_i}\}$ , update the transmit-beamformers  $\{\tilde{\mathbf{g}}_{i,m_i}\}$  by solving (for an auxiliary variable  $\alpha$ )

$$(P10) \quad \min_{\alpha, \{\tilde{\mathbf{g}}_{i,m_i}\}} \alpha \sum_{i=1}^I \bar{P}_i \quad (29a)$$

$$\text{subject to } \frac{|\mathbf{w}_{i,m_i}^H \mathbf{H}_{i,i} \tilde{\mathbf{g}}_{i,m_i}|^2}{1 + \sum_{(j,m_j) \neq (i,m_i)} |\mathbf{w}_{i,m_i}^H \mathbf{H}_{i,j} \tilde{\mathbf{g}}_{j,m_j}|^2} \geq \gamma_{i,m_i}, \quad \forall i, m_i \quad (29b)$$

$$\sum_{m_i=1}^{M_i} \|\tilde{\mathbf{g}}_{i,m_i}\|^2 \leq \alpha \bar{P}_i, \quad \forall i \quad (29c)$$

$$\sum_{i=1}^I \sum_{m_i=1}^{M_i} \|\mathbf{H}_{0,i} \tilde{\mathbf{g}}_{i,m_i}\|^2 \leq \alpha \bar{P}_0. \quad (29d)$$

**S4)** Repeat S1)–S3) until convergence.

Step S1) updates the receive-beamformers given the transmit-beamformers and powers, which does not affect the interference. Step S2) optimizes the per-stream powers and SINRs while adhering to the power and CR interference constraints. Problem (P10) in step S3) is in the form of a feasibility optimization problem for the power and CR interference constraints, through the auxiliary variable  $\alpha$ . Since

$$(P8) \quad \max_{\{\mathbf{w}_{i,m_i}\}, \{\tilde{\mathbf{g}}_{i,m_i}\}} \sum_{i=1}^I \lambda_i \sum_{m_i=1}^{M_i} \log \left( 1 + \frac{|\mathbf{w}_{i,m_i}^H \mathbf{H}_{i,i} \tilde{\mathbf{g}}_{i,m_i}|^2}{1 + \sum_{(j,m_j) \neq (i,m_i)} |\mathbf{w}_{i,m_i}^H \mathbf{H}_{i,j} \tilde{\mathbf{g}}_{j,m_j}|^2} \right) \quad (26a)$$

$$\text{subject to } \sum_{m_i=1}^{M_i} \|\tilde{\mathbf{g}}_{i,m_i}\|^2 \leq \bar{P}_i, \quad \forall i \quad (26b)$$

$$\sum_{i=1}^I \sum_{m_i=1}^{M_i} \|\mathbf{H}_{0,i} \tilde{\mathbf{g}}_{i,m_i}\|^2 \leq \bar{P}_0 \quad (26c)$$

$$\|\mathbf{w}_{i,m_i}\| = 1, \quad \forall i, m_i \quad (26d)$$

the constraints were already enforced in the previous step, the optimal  $\alpha$  cannot be larger than 1. By minimizing a constant  $(\sum_{i=0}^I \bar{P}_i)$  times  $\alpha$ , step S3) essentially minimizes the maximum of the per-node transmit-powers normalized by  $\bar{P}_i$ 's, and the interference power normalized by  $\bar{P}_0$ .

Since the MVDR beamformers maximize the SINRs of the individual streams given the transmit-beamformers and transmit-powers, step S1) increases (or maintains) the objective value of (P8) from the previous iteration. Step S2) optimizes the transmit-powers to maximize the objective with the transmit- and the receive-beamformers fixed. Thus, this step also increases (or maintains) the objective value of (P8). Finally, step S3) uniformly maximizes the margins of the power and the CR interference constraints through the auxiliary variable  $\alpha \in [0, 1]$ , while maintaining the same per-stream SINRs  $\{\gamma_{i,m_i}\}$  for all  $i$ 's and  $m_i$ 's, and hence achieving the same objective value for (P8) from step S2). Therefore, this step provides "room" for the next iteration to increase the objective value. Overall, it is clear that the iterative algorithm S1)–S4) converges. In fact, it can be shown that the algorithm achieves at least local optimality.<sup>1</sup>

*Proposition 5:* Iterative application of S1)–S4) converges to a KKT point of (P8).

*Proof:* See Appendix A. ■

Problem (P9) in S2) falls in the class of truly nonconvex problems, called signomial programs. It cannot be transformed to a convex program unless  $\gamma_{i,m_i} \gg 1, \forall i, m_i$ , which is not the case here. However, a local optimum of (P9) can be obtained by solving a series of GPs as detailed in [4]. The resultant solution very often corresponds to the global optimum [4].

Although (P10) in S3) is nonconvex in its present form, it can be transformed to a convex problem, and thus be solved efficiently [28]. To this end, the alternative approach developed next relies on the network duality concept. Capitalizing on Lagrangian duality, the optimal solution of the MIMO cellular downlink problem with per-antenna transmit-power constraints can be obtained by solving an uplink problem with the worst-case noise covariance matrix [32]. We will considerably broaden the scope of [32] by showing that a similar

<sup>1</sup>In the conference precursor of this paper [14], it was erroneously reported that the algorithm does not attain local optimality.

derivation is possible as in Section V-B for: i) the MISO IC with ii) the CR constraints added. The key observation is that the per-antenna power constraints in [32] become analogous to the per-link power constraints (29c), or the CR interference constraints (29d).

### B. Solution of (P10)

The nonconvex problem (P10) can be solved optimally. The key enabler is the strong duality between (P10) and its dual. The proof follows closely the ones in [28] and [32] developed for an analogous claim in the multiantenna downlink setup. Here, it is shown that strong duality claims carry over to the multiantenna ad hoc network topology as well.

*Lemma 2:* Strong duality holds for (P10) and its dual.

*Proof:* It is first noted that an arbitrary phase rotation can be applied to the transmit-beamforming vectors  $\{\tilde{\mathbf{g}}_{i,m_i}\}$  without altering the optimal solution of (P10). In particular, the phase can be chosen so as to make the imaginary part of  $\mathbf{w}_{i,m_i}^H \mathbf{H}_{i,i} \tilde{\mathbf{g}}_{i,m_i}$  vanish,  $\forall i, m_i$ . Then, one can rewrite the SINR constraints (29b) as second-order cone constraints expressed as

$$\sqrt{1 + \frac{1}{\gamma_{i,m_i}} \mathbf{w}_{i,m_i}^H \mathbf{H}_{i,i} \tilde{\mathbf{g}}_{i,m_i}} \geq \left\| \begin{bmatrix} 1 \\ \tilde{\mathbf{G}}_1^H \mathbf{H}_{i,1}^H \mathbf{w}_{i,m_i} \\ \tilde{\mathbf{G}}_2^H \mathbf{H}_{i,2}^H \mathbf{w}_{i,m_i} \\ \vdots \\ \tilde{\mathbf{G}}_I^H \mathbf{H}_{i,I}^H \mathbf{w}_{i,m_i} \end{bmatrix} \right\|, \quad \forall i, m_i \quad (30)$$

where  $\|\cdot\|$  denotes the Euclidean norm. Replacing (29b) in (P10) by the convex constraints (30) renders the problem convex, and thus strong duality holds for this reformulated problem and its dual. The remaining task is to show that the Lagrangian dual of (P10) coincides with the Lagrangian dual of this reformulated problem. This can be done in a completely analogous fashion to the proof for the multiantenna downlink in [32, Appendix A]; for this reason, details are omitted. ■

Although a general-purpose convex optimization software can solve the convexified problem, the following proposition provides a means to develop an efficient solver for (P10), based on the preceding lemma.

$$(P9) \quad \max_{\{P_{i,m_i}\}, \{\gamma_{i,m_i}\}} \sum_{i=1}^I \lambda_i \sum_{m_i=1}^{M_i} \log(1 + \gamma_{i,m_i}) \quad (28a)$$

$$\text{subject to } \gamma_{i,m_i} \leq \frac{|\mathbf{w}_{i,m_i}^H \mathbf{H}_{i,i} \mathbf{g}_{i,m_i}|^2 P_{i,m_i}}{1 + \sum_{(j,m_j) \neq (i,m_i)} |\mathbf{w}_{i,m_i}^H \mathbf{H}_{i,j} \mathbf{g}_{j,m_j}|^2 P_{j,m_j}}, \quad \forall i, m_i \quad (28b)$$

$$\sum_{m_i=1}^{M_i} P_{i,m_i} \leq \bar{P}_i, \quad \forall i \quad (28c)$$

$$\sum_{i=1}^I \sum_{m_i=1}^{M_i} \|\mathbf{H}_{0,i} \mathbf{g}_{i,m_i}\|^2 P_{i,m_i} \leq \bar{P}_0 \quad (28d)$$

$$P_{i,m_i} \geq 0 \quad \forall i, m_i. \quad (28e)$$

TABLE III  
 THE ALGORITHM SOLVING (P10)

Input: $\{\mathbf{w}_{i,m_i}\}$ and $\{\gamma_{i,m_i}\}$
1: Initialize $\mathbf{E}^{-1} = \text{diag} \left\{ (I+1) \left( \frac{\mu_{i,\max}}{2} \right)^2 \right\}_{i=0}^I$ , and $\boldsymbol{\mu} = \left[ \frac{\mu_{i,\max}}{2} \right]_{i=0}^I$ .
2: Repeat until convergence lines 3–11:
3: While $\boldsymbol{\mu} \notin \left\{ \boldsymbol{\mu} \mid \boldsymbol{\mu} \geq \mathbf{0}, \sum_{i=0}^I \mu_i \bar{P}_i \leq \sum_{i=0}^I \bar{P}_i \right\}$
4: If $\mu_i < 0$ , set $\mathbf{d} = \mathbf{e}_{i+1}$ ( $(i+1)$ -st canonical basis); or, if $\sum_{i=0}^I \mu_i \bar{P}_i > \sum_{i=0}^I \bar{P}_i$ , set $\mathbf{d} = -[\bar{P}_0 \ \cdots \ \bar{P}_I]^T$ .
5: Perform the ellipsoid update: $\mathbf{d} \leftarrow \mathbf{d} / \sqrt{\mathbf{d}^T \mathbf{E}^{-1} \mathbf{d}}$ $\boldsymbol{\mu} \leftarrow \boldsymbol{\mu} + \mathbf{E}^{-1} \mathbf{d} / (I+2)$ $\mathbf{E}^{-1} \leftarrow \frac{(I+1)^2}{(I+1)^2 - 1} \left( \mathbf{E}^{-1} - \frac{2}{I+2} \mathbf{E}^{-1} \mathbf{d} \mathbf{d}^T \mathbf{E}^{-1} \right)$
6: Repeat until convergence: $\zeta_{i,m_i} \leftarrow \begin{cases} \frac{1}{\left(1 + \frac{1}{\gamma_{i,m_i}}\right) \mathbf{w}_{i,m_i}^H \mathbf{H}_{i,m_i} \hat{\mathbf{R}}_{i,m_i}^\dagger \mathbf{H}_{i,m_i}^H \mathbf{w}_{i,m_i}}, & \text{if } \gamma_{i,m_i} \neq 0 \\ 0, & \text{if } \gamma_{i,m_i} = 0 \end{cases}, \forall i, m_i.$
7: Set $\hat{\mathbf{g}}_{i,m_i} = \hat{\mathbf{R}}_{i,m_i}^\dagger \mathbf{H}_{i,m_i}^H \mathbf{w}_{i,m_i}, \forall i, m_i.$
8: Set $\mathbf{g}_{i,m_i} = \hat{\mathbf{g}}_{i,m_i} / \ \hat{\mathbf{g}}_{i,m_i}\ , \forall i, m_i.$
9: Compute $\{P_{i,m_i}\}$ by solving (32).
10: Set $\mathbf{d} = \left[ \sum_{i,m_i} \ \mathbf{H}_{0,i} \mathbf{g}_{i,m_i}\ ^2 P_{i,m_i} \sum_{m_1} P_{1,m_1} \cdots \sum_{m_I} P_{I,m_I} \right]^T$ .
11: Perform the ellipsoid update (same as line 5 above).
Output: $\{\mathbf{g}_{i,m_i}\}$ and $\{P_{i,m_i}\}$ .

*Proposition 6:* Problem (P10) can be solved via the following joint receive-beamforming and transmit-power control problem, where  $\zeta_{j,m_j}$  plays the role of transmit-power,  $\mathbf{w}_{j,m_j}$  the transmit-beamformer, and  $\hat{\mathbf{g}}_{i,m_i}$  the receive-beamformer, over the “flipped” channel  $\mathbf{H}_{j,i}^H$ .

$$(P11) \quad \max_{\substack{\mu_0 \geq 0, \dots, \mu_I \geq 0 \\ \sum_{i=0}^I \mu_i \bar{P}_i \leq \sum_{i=0}^I \bar{P}_i}} \min_{\{\zeta_{i,m_i} \geq 0\}, \{\mathbf{g}_{i,m_i}\}} \sum_{i,m_i} \zeta_{i,m_i} \quad (31a)$$

$$\text{subject to } \frac{|\hat{\mathbf{g}}_{i,m_i}^H \mathbf{H}_{i,i}^H \mathbf{w}_{i,m_i}|^2 \zeta_{i,m_i}}{\hat{\mathbf{g}}_{i,m_i}^H \hat{\mathbf{R}}_{i,m_i} \hat{\mathbf{g}}_{i,m_i}} \geq \gamma_{i,m_i}, \quad \forall i, m_i \quad (31b)$$

where  $\hat{\mathbf{R}}_{i,m_i} \triangleq \sum_{(j,m_j) \neq (i,m_i)} \zeta_{j,m_j} \mathbf{H}_{j,i}^H \mathbf{w}_{j,m_j} \mathbf{w}_{j,m_j}^H \mathbf{H}_{j,i} + \mu_i \mathbf{I} + \mu_0 \mathbf{H}_{0,i}^H \mathbf{H}_{0,i}$ . The optimal  $\mathbf{g}_{i,m_i}$  is given by  $\hat{\mathbf{g}}_{i,m_i} / \|\hat{\mathbf{g}}_{i,m_i}\|$  for each  $i$  and  $m_i$ , and the powers  $\{P_{i,m_i}\}$  are obtained by solving the system of equations

$$\frac{|\mathbf{w}_{i,m_i}^H \mathbf{H}_{i,i} \mathbf{g}_{i,m_i}|^2 P_{i,m_i}}{\sum_{(j,m_j) \neq (i,m_i)} |\mathbf{w}_{i,m_i}^H \mathbf{H}_{i,j} \mathbf{g}_{j,m_j}|^2 P_{j,m_j} + 1} = \gamma_{i,m_i}, \quad \forall i, m_i. \quad (32)$$

*Proof:* See Appendix B.  $\blacksquare$

The algorithm to solve (P10) is summarized in Table III, where  $(\cdot)^\dagger$  denotes matrix pseudoinverse. To determine the initial ellipsoid, it is noted that  $0 \leq \mu_i \leq \mu_{i,\max} \triangleq (\sum_{j=0}^I \bar{P}_j) / \bar{P}_i, \forall i$ , since  $\sum_{i=0}^I \mu_i \bar{P}_i \leq \sum_{i=0}^I \bar{P}_i$  must hold. Lines 3–5 ensure that the feasibility conditions for  $\boldsymbol{\mu} \triangleq [\mu_0 \ \mu_1 \ \dots \ \mu_I]^T$  are met. For given  $\{\mu_i\}$ , the minimization in (P11) can be performed by an iterative algorithm, given in line 6, whose detailed derivation is omitted as it is analogous

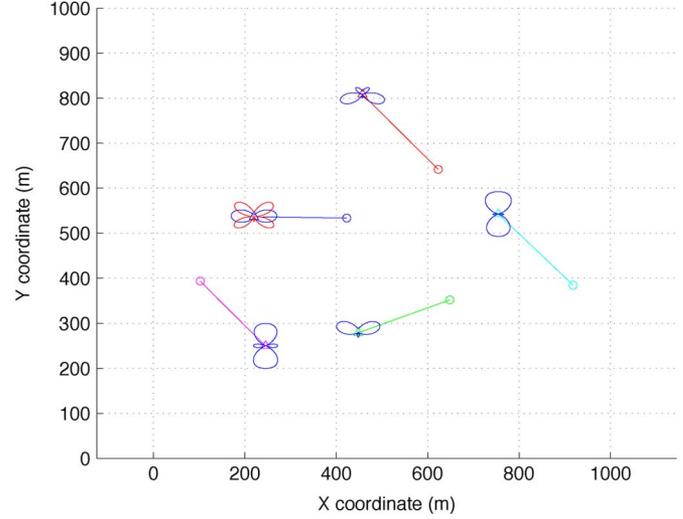


Fig. 1. Example MIMO CR network topology.

to the ones presented in [28] and [32]. To update  $\{\mu_i\}$  for the maximization in (P11), the ellipsoid method is again used in lines 10–11, where the subgradient of the optimal value of the minimization w.r.t.  $\boldsymbol{\mu}$  is given by [32]

$$\left[ \sum_{i,m_i} \|\mathbf{H}_{0,i} \hat{\mathbf{g}}_{i,m_i}\|^2, \sum_{m_1} \|\hat{\mathbf{g}}_{1,m_1}\|^2, \dots, \sum_{m_I} \|\hat{\mathbf{g}}_{I,m_I}\|^2 \right]^T. \quad (33)$$

## VI. NUMERICAL RESULTS

The configuration depicted in Fig. 1 is a representative scenario. In this setup, there are five links that operate over the same frequency band, while transmitters and receivers are equipped with  $N_i = M_i = 3$  antennas each. The pathloss exponent is assumed to be 2, and the thermal noise power is set to  $2 \times 10^{-6}$  W. The maximum transmit-power  $\bar{P}_i$  is equal to 10 W for all  $i = 1, 2, \dots, 5$ . The flat fading channel per node pair was generated to have a line-of-sight (LOS) path as well as two reflected paths with angles  $\pm 30^\circ$  w.r.t. the LOS. The reflected paths have power gain half that of the LOS path. The subsequent numerical results are based on this particular realization of channels, except for Fig. 5, for which 50 different realizations were tested.

Fig. 2 shows the evolution of the weighted sum-rate when (P2) is solved via the algorithm in Table I using Gauss-Seidel. Two sets of weights  $\boldsymbol{\lambda} = \frac{1}{5}[1 \ 1 \ 1 \ 1 \ 1]$  and  $\boldsymbol{\lambda} = \frac{1}{15}[5 \ 4 \ 3 \ 2 \ 1]$  were tested. The weighted sum-rates are seen to converge to the near-optimal values quite rapidly. Also, the monotonic increase of the weighted sum-rates can be verified. Fig. 1 shows the optimized beam patterns for the equal-weight case. It is seen that one link uses two streams while the others use only a single stream. The beams are steered so as to balance between transmitting more power toward the desired receivers while avoiding interference imposed to others.

Fig. 3 pertains to the case when the rightmost link in Fig. 1 becomes a PU, and the rest four links are constrained to generate total interference at the PU of no more than  $\bar{P}_0 = 1$  W. Equal weights were used. The maximum transmit-power at each link is again  $\bar{P}_i = 10$  for all  $i$ . The top panel of Fig. 3 depicts the

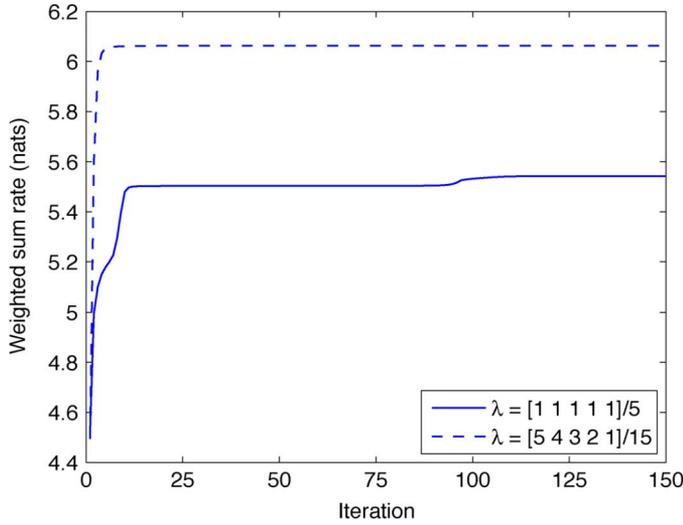


Fig. 2. Weighted sum-rate without the CR constraint.

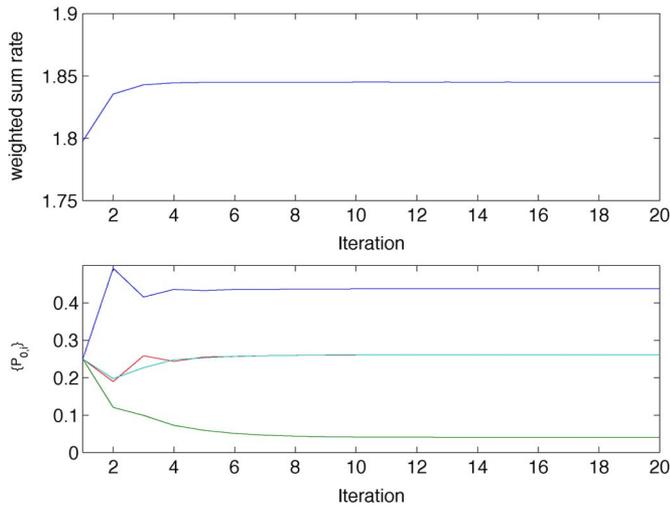


Fig. 3. Weighted sum-rate with the CR constraint (top) and the evolution of  $\{\bar{P}_{0,i}\}$  (bottom).

evolution of the sum-rate each time the maximum interference powers  $\{\bar{P}_{0,i}\}$  are updated following (24) with  $\alpha_n = 1$  for all  $n$ . The bottom panel shows the evolution of  $\{\bar{P}_{0,i}\}$  starting from the equal assignment of  $\bar{P}_{0,i} = \bar{P}_0/I$  for all  $i$ . It can be seen that although convergence is observed within 10 iterations, the sum-rate is increased no more than 3% by updating  $\{\bar{P}_{0,i}\}$ . This is due to the fact that enough degrees of freedom become available in the spatial domain to avoid interference to the PU such that the additional degrees of freedom to play with the powers add little value. Thus, the centralized operation of adjusting  $\{\bar{P}_{0,i}\}$  can be removed without much sacrificing performance when the system has sufficient capability to steer away the beams from the PUs. In this case, the central controller is no longer necessary.

Performance of the centralized algorithm in Table III is illustrated in Fig. 4 for the scenario in Fig. 3. In this particular case, the centralized algorithm yields a sum-rate higher than the distributed algorithm, as can be verified from the top panel of the figure. However, its computational complexity is considerably higher than that of its distributed counterpart. The bottom panel

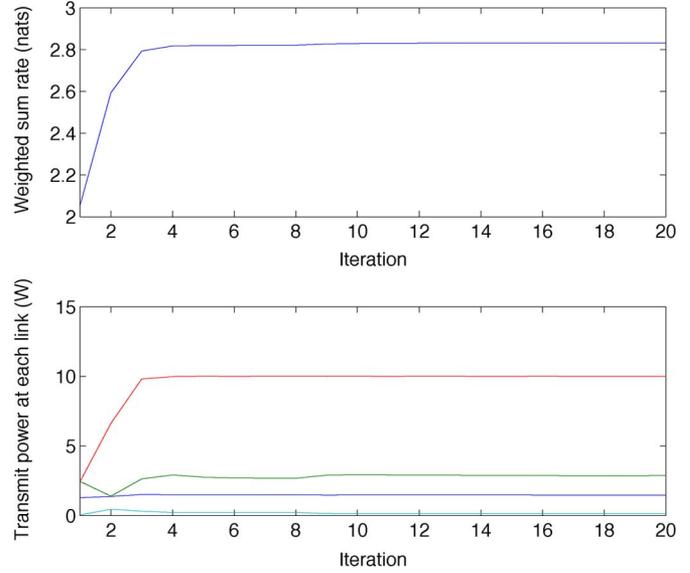


Fig. 4. Sum-rate (top) and transmit-powers  $\{\sum_{m_i} P_{i,m_i}\}$  (bottom) corresponding to the centralized algorithm.

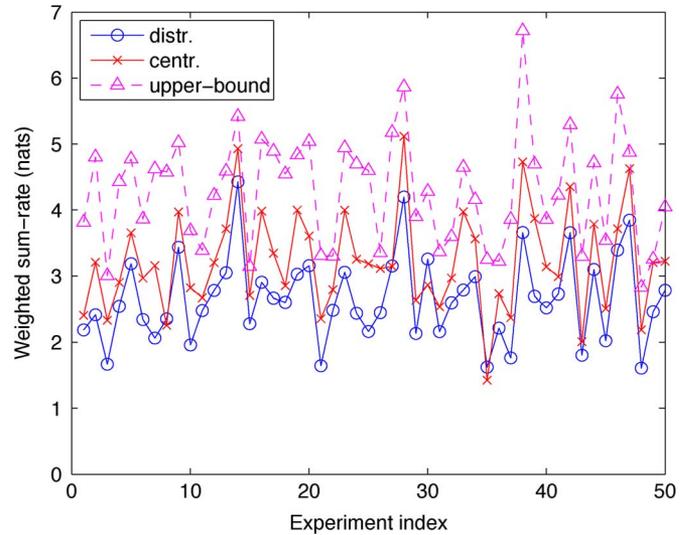


Fig. 5. Sum-rates achieved by the distributed and the centralized algorithms.

depicts the transmit-power at each link, which is constrained to be less than or equal to 10 W.

To compare the actual performances of the proposed distributed and centralized algorithms, the two algorithms were applied to 50 random topologies with the node positions uniformly distributed in a 1000 m  $\times$  1000 m area. Four links were present in each experiment with  $M_i = N_i = 3$  antennas, and the CR constraints were not imposed. Fig. 5 depicts the achieved sum-rate in the 50 different experiments. The upper-bounds, which are also shown in Fig. 5, were obtained by artificially setting all interference channels to zero matrices. Since both algorithms aim at locally optimal solutions, there is no guarantee that one will uniformly outperform the other. However, one can see that the centralized algorithm usually performs better than the distributed one, which partially justifies the heavy feedback overhead associated with the centralized

algorithm. It is noted that a similar trend has also been observed in [21] for the single-antenna case.

## VII. CONCLUSION

Distributed resource allocation algorithms for MIMO ad hoc CR networks have been developed to maximize the weighted sum-rate of any-to-any links under per-node transmit-power and CR interference constraints. When the CR interference constraints are absent, the problem amounts to weighed sum-rate maximization for MIMO ad hoc networks with per-node power constraints when other-user interferences are treated as noise, which is, in itself, an important problem. A provably convergent distributed algorithm was developed that yields a locally optimal solution.

When the CR interference constraints are added, the primal decomposition technique was adopted to ensure that the constraints are enforced even during the iterative procedure, and a semidistributed algorithm was proposed.

An alternative centralized algorithm was also developed based on network duality and GP-based power control. Existing network duality results for the single-cell cellular network topology were extended to the ad hoc network configuration under the per-node power and CR interference constraints. The novel algorithm views MIMO optimization as a multi-beam beamforming problem without loss of optimality, and iteratively updates receive-beamformers, transmit-powers, and transmit-beamformers in cycle to attain a KKT-optimal solution. The numerical tests verified the efficacy of the proposed algorithms.

Overall, although optimization of MIMO interference network is a difficult problem even when simple single-user decoding is assumed per link, the study finds that “good” solutions can be obtained in a distributed and on-line fashion exploiting

only local feedback, and CR interference constraints do not necessarily alter the fundamental solution structure. Given a more lenient budget on computational complexity and feedback overhead, a centralized alternative often provides an improved performance, tapping into the power of GP and network duality.

Recently, the interference alignment idea advanced the understanding of ICs in a major way by characterizing the inherent degrees of freedom in the high-SNR regime [3]. However, from a practical standpoint, it is still important to have weighted sum-rate-optimal solutions for the low- and mid-range SNRs. In fact, algorithms that specifically target interference alignment seem to be often outperformed by algorithms that directly maximize the sum-rates in finite SNRs [22]. The bottom line is that a weighted sum-rate maximization does not preclude interference alignment, and often leads to desirable solutions in the SNR ranges of practical interest.

## APPENDIX A

### PROOF OF PROPOSITION 5

Convergence is clear as each of S1)–S3) does not decrease the objective of (P8). It remains to show that the convergence point corresponds to a KKT point of (P8).

For ease of exposition, let us ignore the CR interference constraints, which does not change the essential steps of the proof. Since the inequalities (28b) in (P9) are tight at the optimum, (P9) can be equivalently rewritten as (34), at the bottom of the page. The Lagrangian of (34) is given by (35), at the bottom of the page, where  $\{\beta_i \geq 0\}$  and  $\{\nu_{i,m_i} \geq 0\}$  are the Lagrange multipliers associated with the constraints (34b) and (34c), respectively. Then, the conditions shown in (36)–(38), at the bottom of the next page, are necessary for optimality, and hence are enforced at the fixed point of the iterative procedure S1)–S4). It is

$$\max_{\{P_{i,m_i}\}} \sum_{i=1}^I \lambda_i \sum_{m_i=1}^{M_i} \log \left( 1 + \frac{|\mathbf{w}_{i,m_i}^H \mathbf{H}_{i,i} \mathbf{g}_{i,m_i}|^2 P_{i,m_i}}{1 + \sum_{(j,m_j) \neq (i,m_i)} |\mathbf{w}_{i,m_i}^H \mathbf{H}_{i,j} \mathbf{g}_{j,m_j}|^2 P_{j,m_j}} \right) \quad (34a)$$

$$\text{subject to} \quad \sum_{m_i=1}^{M_i} P_{i,m_i} \leq \bar{P}_i, \quad \forall i \quad (34b)$$

$$P_{i,m_i} \geq 0 \quad \forall i, m_i. \quad (34c)$$

$$\begin{aligned} & L_1(\{P_{i,m_i}\}, \{\beta_i\}, \{\nu_{i,m_i}\}) \\ &= \sum_{i=1}^I \lambda_i \sum_{m_i=1}^{M_i} \log \left( 1 + \frac{|\mathbf{w}_{i,m_i}^H \mathbf{H}_{i,i} \mathbf{g}_{i,m_i}|^2 P_{i,m_i}}{1 + \sum_{(j,m_j) \neq (i,m_i)} |\mathbf{w}_{i,m_i}^H \mathbf{H}_{i,j} \mathbf{g}_{j,m_j}|^2 P_{j,m_j}} \right) \\ & \quad - \sum_{i=1}^I \beta_i \left( \sum_{m_i=1}^{M_i} P_{i,m_i} - \bar{P}_i \right) + \sum_{i=1}^I \sum_{m_i=1}^{M_i} \nu_{i,m_i} P_{i,m_i} \end{aligned} \quad (35)$$

useful to note that the following relations hold for all  $i, m_i$  at the fixed point due to (27):

$$1 + \sum_{(j,m_j) \neq (i,m_i)} |\mathbf{w}_{i,m_i}^H \mathbf{H}_{i,j} \mathbf{g}_{j,m_j}|^2 P_{j,m_j} = \frac{\mathbf{g}_{i,m_i}^H \mathbf{H}_{i,i} \mathbf{R}_{i,m_i}^{-1} \mathbf{H}_{i,i} \mathbf{g}_{i,m_i}}{\|\mathbf{R}_{i,m_i}^{-1} \mathbf{H}_{i,i} \mathbf{g}_{i,m_i}\|^2} \quad (39)$$

$$|\mathbf{w}_{i,m_i}^H \mathbf{H}_{i,i} \mathbf{g}_{i,m_i}|^2 = \frac{(\mathbf{g}_{i,m_i}^H \mathbf{H}_{i,i} \mathbf{R}_{i,m_i}^{-1} \mathbf{H}_{i,i} \mathbf{g}_{i,m_i})^2}{\|\mathbf{R}_{i,m_i}^{-1} \mathbf{H}_{i,i} \mathbf{g}_{i,m_i}\|^2} \quad (40)$$

$$\gamma_{i,m_i} = \tilde{\mathbf{g}}_{i,m_i}^H \mathbf{H}_{i,i} \mathbf{R}_{i,m_i}^{-1} \mathbf{H}_{i,i} \tilde{\mathbf{g}}_{i,m_i}. \quad (41)$$

Substituting (39)–(41) into (36), multiplying the both sides by  $P_{i,m_i}$ , and using (37) yields

$$0 = \beta_i P_{i,m_i} + \frac{\lambda_i \gamma_{i,m_i}}{1 + \gamma_{i,m_i}} - \sum_{(j,m_j) \neq (i,m_i)} \frac{\lambda_j |\tilde{\mathbf{g}}_{j,m_j}^H \mathbf{H}_{j,j}^H \mathbf{R}_{j,m_j}^{-1} \mathbf{H}_{j,i} \tilde{\mathbf{g}}_{i,m_i}|^2}{1 + \gamma_{j,m_j}}, \quad \forall i, m_i. \quad (42)$$

Consider now (P10) without the CR interference constraints:

$$\min_{\alpha, \{\tilde{\mathbf{g}}_{i,m_i}\}} \alpha \sum_{i=1}^I \bar{P}_i \quad (43a)$$

$$\text{subject to } \frac{|\mathbf{w}_{i,m_i}^H \mathbf{H}_{i,i} \tilde{\mathbf{g}}_{i,m_i}|^2}{1 + \sum_{(j,m_j) \neq (i,m_i)} |\mathbf{w}_{i,m_i}^H \mathbf{H}_{i,j} \tilde{\mathbf{g}}_{j,m_j}|^2} \geq \gamma_{i,m_i}, \quad \forall i, m_i \quad (43b)$$

$$\sum_{m_i=1}^{M_i} \|\tilde{\mathbf{g}}_{i,m_i}\|^2 \leq \alpha \bar{P}_i, \quad \forall i. \quad (43c)$$

Assuming w.l.o.g. that  $\gamma_{i,m_i} > 0, \forall i, m_i$ , the Lagrangian of (43) is given by

$$L_2(\alpha, \{\tilde{\mathbf{g}}_{i,m_i}\}, \{\kappa_i\}, \{\eta_{i,m_i}\}) = \alpha \sum_{i=1}^I \bar{P}_i + \sum_{i=1}^I \kappa_i \left( \sum_{m_i=1}^{M_i} \|\tilde{\mathbf{g}}_{i,m_i}\|^2 - \alpha \bar{P}_i \right)$$

$$- \sum_{i=1}^I \sum_{m_i=1}^{M_i} \eta_{i,m_i} \left( \frac{1}{\gamma_{i,m_i}} |\mathbf{w}_{i,m_i}^H \mathbf{H}_{i,i} \tilde{\mathbf{g}}_{i,m_i}|^2 - \sum_{(j,m_j) \neq (i,m_i)} |\mathbf{w}_{i,m_i}^H \mathbf{H}_{i,j} \tilde{\mathbf{g}}_{j,m_j}|^2 - 1 \right) \quad (44)$$

where  $\{\kappa_i \geq 0\}$  and  $\{\eta_{i,m_i} \geq 0\}$  are the Lagrange multipliers. The necessary conditions for optimality for (P10) include

$$0 = \frac{\partial L_2}{\partial \alpha} = \sum_{i=1}^I \bar{P}_i - \sum_{i=1}^I \kappa_i \bar{P}_i \quad (45)$$

$$0 = \frac{\partial L_2}{\partial \tilde{\mathbf{g}}_{i,m_i}} = \kappa_i \tilde{\mathbf{g}}_{i,m_i} - \frac{\eta_{i,m_i} \mathbf{H}_{i,i}^H \mathbf{w}_{i,m_i} \mathbf{w}_{i,m_i}^H \mathbf{H}_{i,i} \tilde{\mathbf{g}}_{i,m_i}}{\gamma_{i,m_i}} + \sum_{(j,m_j) \neq (i,m_i)} \eta_{j,m_j} \mathbf{H}_{j,i}^H \mathbf{w}_{j,m_j} \mathbf{w}_{j,m_j}^H \mathbf{H}_{j,i} \tilde{\mathbf{g}}_{i,m_i}, \quad \forall i, m_i \quad (46)$$

$$0 = \kappa_i \left( \sum_{m_i=1}^{M_i} \|\tilde{\mathbf{g}}_{i,m_i}\|^2 - \bar{P}_i \right), \quad \forall i \quad (47)$$

$$0 = \eta_{i,m_i} \left( \frac{1}{\gamma_{i,m_i}} |\mathbf{w}_{i,m_i}^H \mathbf{H}_{i,i} \tilde{\mathbf{g}}_{i,m_i}|^2 - \sum_{(j,m_j) \neq (i,m_i)} |\mathbf{w}_{i,m_i}^H \mathbf{H}_{i,j} \tilde{\mathbf{g}}_{j,m_j}|^2 - 1 \right), \quad \forall i, m_i \quad (48)$$

where the fact that the optimal  $\alpha = 1$  at the fixed point has been used in (47), and it is noted that (48) holds automatically as the inequalities in (28b) are tight. Substituting (27) into (46) and left-multiplying the both sides by  $\tilde{\mathbf{g}}_{i,m_i}^H$  yields

$$0 = \kappa_i P_{i,m_i} - \frac{\eta_{i,m_i} \gamma_{i,m_i}}{\|\mathbf{R}_{i,m_i}^{-1} \mathbf{H}_{i,i} \tilde{\mathbf{g}}_{i,m_i}\|^2} + \sum_{(j,m_j) \neq (i,m_i)} \frac{\eta_{j,m_j} |\tilde{\mathbf{g}}_{j,m_j}^H \mathbf{H}_{j,j}^H \mathbf{R}_{j,m_j}^{-1} \mathbf{H}_{j,i} \tilde{\mathbf{g}}_{i,m_i}|^2}{\|\mathbf{R}_{j,m_j}^{-1} \mathbf{H}_{j,j} \tilde{\mathbf{g}}_{j,m_j}\|^2}, \quad \forall i, m_i. \quad (49)$$

$$0 = \frac{\partial L_1}{\partial P_{i,m_i}} = -\beta_i + \nu_{i,m_i} + \frac{\lambda_i |\mathbf{w}_{i,m_i}^H \mathbf{H}_{i,i} \mathbf{g}_{i,m_i}|^2}{1 + \sum_{(j,m_j)} |\mathbf{w}_{i,m_i}^H \mathbf{H}_{i,j} \mathbf{g}_{j,m_j}|^2} P_{j,m_j} - \sum_{(j,m_j) \neq (i,m_i)} \frac{\lambda_j |\mathbf{w}_{j,m_j}^H \mathbf{H}_{j,j} \mathbf{g}_{j,m_j}|^2 P_{j,m_j} \cdot |\mathbf{w}_{j,m_j}^H \mathbf{H}_{j,i} \mathbf{g}_{i,m_i}|^2}{\left(1 + \sum_{(k,m_k) \neq (j,m_j)} |\mathbf{w}_{j,m_j}^H \mathbf{H}_{j,k} \mathbf{g}_{k,m_k}|^2 P_{k,m_k}\right) \left(1 + \sum_{(k,m_k)} |\mathbf{w}_{j,m_j}^H \mathbf{H}_{j,k} \mathbf{g}_{k,m_k}|^2 P_{k,m_k}\right)}, \quad \forall i, m_i \quad (36)$$

$$0 = \nu_{i,m_i} P_{i,m_i}, \quad \forall i, m_i \quad (37)$$

$$0 = \beta_i \left( \sum_{m_i=1}^{M_i} P_{i,m_i} - \bar{P}_i \right), \quad \forall i. \quad (38)$$

By comparing (42) with (49), one can deduce that  $\kappa_i$  and  $\eta_{i,m_i}$  must be proportional to  $\beta_i$  and  $\frac{\lambda_i}{1+\gamma_{i,m_i}} \|\mathbf{R}_{i,m_i}^{-1} \mathbf{H}_{i,i} \tilde{\mathbf{g}}_{i,m_i}\|^2$ , respectively, with the proportionality constant determined to satisfy (45). Substituting these relations into (46), and using (41) yields (50), at the bottom of the page.

The goal is to show that the conditions explored so far imply the KKT conditions for (P8). Replacing the receive-beamformers by the optimal MVDR beamformers, (P8) is equivalent to

$$\max_{\{\tilde{\mathbf{g}}_{i,m_i}\}} \sum_{i=1}^I \lambda_i \sum_{m_i=1}^{M_i} \log(1 + \tilde{\mathbf{g}}_{i,m_i}^{\mathcal{H}} \mathbf{H}_{i,i}^{\mathcal{H}} \mathbf{R}_{i,m_i}^{-1} \mathbf{H}_{i,i} \tilde{\mathbf{g}}_{i,m_i}) \quad (51a)$$

$$\text{subject to } \sum_{m_i=1}^{M_i} \|\tilde{\mathbf{g}}_{i,m_i}\|^2 \leq \bar{P}_i, \quad \forall i. \quad (51b)$$

Upon introducing the Lagrange multipliers  $\{\theta_i\}$  to relax (51b), the Lagrangian is given by

$$\begin{aligned} L_3(\{\tilde{\mathbf{g}}_{i,m_i}\}, \{\theta_i\}) &= \sum_{i=1}^I \lambda_i \sum_{m_i=1}^{M_i} \log(1 + \tilde{\mathbf{g}}_{i,m_i}^{\mathcal{H}} \mathbf{H}_{i,i}^{\mathcal{H}} \mathbf{R}_{i,m_i}^{-1} \mathbf{H}_{i,i} \tilde{\mathbf{g}}_{i,m_i}) \\ &\quad - \sum_{i=1}^I \theta_i \left( \sum_{m_i=1}^{M_i} \|\tilde{\mathbf{g}}_{i,m_i}\|^2 - \bar{P}_i \right) \end{aligned} \quad (52)$$

and the KKT conditions are

$$\begin{aligned} \mathbf{0} &= \frac{\partial L_3}{\partial \tilde{\mathbf{g}}_{i,m_i}} = \frac{\lambda_i \mathbf{H}_{i,i}^{\mathcal{H}} \mathbf{R}_{i,m_i}^{-1} \mathbf{H}_{i,i} \tilde{\mathbf{g}}_{i,m_i}}{1 + \tilde{\mathbf{g}}_{i,m_i}^{\mathcal{H}} \mathbf{H}_{i,i}^{\mathcal{H}} \mathbf{R}_{i,m_i}^{-1} \mathbf{H}_{i,i} \tilde{\mathbf{g}}_{i,m_i}} \\ &\quad - \sum_{(j,m_j) \neq (i,m_i)} \lambda_j \\ &\quad \times \frac{\mathbf{H}_{j,i}^{\mathcal{H}} \mathbf{R}_{j,m_j}^{-1} \mathbf{H}_{j,j} \tilde{\mathbf{g}}_{j,m_j} \tilde{\mathbf{g}}_{j,m_j}^{\mathcal{H}} \mathbf{H}_{j,j}^{\mathcal{H}} \mathbf{R}_{j,m_j}^{-1} \mathbf{H}_{j,i} \tilde{\mathbf{g}}_{i,m_i}}{1 + \tilde{\mathbf{g}}_{j,m_j}^{\mathcal{H}} \mathbf{H}_{j,j}^{\mathcal{H}} \mathbf{R}_{j,m_j}^{-1} \mathbf{H}_{j,j} \tilde{\mathbf{g}}_{j,m_j}} \\ &\quad - \theta_i \tilde{\mathbf{g}}_{i,m_i}, \quad \forall i, m_i \end{aligned} \quad (53)$$

$$0 = \theta_i \left( \sum_{m_i=1}^{M_i} \|\tilde{\mathbf{g}}_{i,m_i}\|^2 - \bar{P}_i \right), \quad \forall i \quad (54)$$

$$\sum_{m_i=1}^{M_i} \|\tilde{\mathbf{g}}_{i,m_i}\|^2 \leq \bar{P}_i, \quad \forall i \quad (55)$$

$$\theta_i \geq 0, \quad \forall i. \quad (56)$$

By choosing  $\theta_i = \beta_i, \forall i$ , one can see from (38) and (50) that the fixed point  $\{\tilde{\mathbf{g}}_{i,m_i}\}$  of the iterative procedure S1)–S4) satisfies the KKT conditions (53)–(56).

## APPENDIX B

### PROOF OF PROPOSITION 6

The proof is adapted from [32]. Letting  $\{\zeta_{i,m_i}\}, \{\mu_i\}_{i=1}^I$ , and  $\mu_0$  denote the Lagrange multipliers associated with constraints (29b), (29c), and (29d), respectively, the Lagrangian corresponding to (P10) can be written as

$$\begin{aligned} L(\alpha, \{\tilde{\mathbf{g}}_{i,m_i}\}, \{\mu_i\}_{i=0}^I, \{\zeta_{i,m_i}\}) &= \alpha \sum_{i=0}^I \bar{P}_i + \mu_0 \left( \sum_{i=1}^I \sum_{m_i=1}^{M_i} \|\mathbf{H}_{0,i} \tilde{\mathbf{g}}_{i,m_i}\|^2 - \alpha \bar{P}_0 \right) \\ &\quad + \sum_{i=1}^I \mu_i \left( \sum_{m_i=1}^{M_i} \|\tilde{\mathbf{g}}_{i,m_i}\|^2 - \alpha \bar{P}_i \right) \\ &\quad - \sum_{i=1}^I \sum_{m_i=1}^{M_i} \zeta_{i,m_i} \left( \frac{1}{\gamma_{i,m_i}} |\mathbf{w}_{i,m_i}^{\mathcal{H}} \mathbf{H}_{i,i} \tilde{\mathbf{g}}_{i,m_i}|^2 \right. \\ &\quad \left. - \sum_{(j,m_j) \neq (i,m_i)} |\mathbf{w}_{i,m_i}^{\mathcal{H}} \mathbf{H}_{i,j} \tilde{\mathbf{g}}_{j,m_j}|^2 - 1 \right) \quad (57) \\ &= \sum_{i=1}^I \sum_{m_i=1}^{M_i} \zeta_{i,m_i} - \alpha \left( \sum_{i=0}^I \mu_i \bar{P}_i - \sum_{i=0}^I \bar{P}_i \right) \\ &\quad + \sum_{i=1}^I \sum_{m_i=1}^{M_i} \tilde{\mathbf{g}}_{i,m_i}^{\mathcal{H}} \left[ \mu_0 \mathbf{H}_{0,i}^{\mathcal{H}} \mathbf{H}_{0,i} + \mu_i \mathbf{I} \right. \\ &\quad \left. - \frac{\zeta_{i,m_i}}{\gamma_{i,m_i}} \mathbf{H}_{i,i}^{\mathcal{H}} \mathbf{w}_{i,m_i} \mathbf{w}_{i,m_i}^{\mathcal{H}} \mathbf{H}_{i,i} \right. \\ &\quad \left. + \sum_{(j,m_j) \neq (i,m_i)} \zeta_{j,m_j} \mathbf{H}_{j,i}^{\mathcal{H}} \mathbf{w}_{j,m_j} \mathbf{w}_{j,m_j}^{\mathcal{H}} \mathbf{H}_{j,i} \right] \tilde{\mathbf{g}}_{i,m_i}. \end{aligned} \quad (58)$$

Thus, the dual function is given by

$$\begin{aligned} D(\{\mu_i\}, \{\zeta_{i,m_i}\}) &= \min_{\{\tilde{\mathbf{g}}_{i,m_i}\}} \min_{\alpha} L(\alpha, \{\tilde{\mathbf{g}}_{i,m_i}\}, \{\mu_i\}, \{\zeta_{i,m_i}\}) \quad (59) \end{aligned}$$

and the dual problem by

$$\max_{\{\mu_i \geq 0\}} \max_{\{\zeta_{i,m_i} \geq 0\}} \sum_{i=1}^I \sum_{m_i=1}^{M_i} \zeta_{i,m_i} \quad (60a)$$

$$\text{subject to } \sum_{i=0}^I \mu_i \bar{P}_i \leq \sum_{i=0}^I \bar{P}_i \quad (60b)$$

$$\begin{aligned} \mathbf{0} &= \beta_i \tilde{\mathbf{g}}_{i,m_i} - \frac{\lambda_i \mathbf{H}_{i,i}^{\mathcal{H}} \mathbf{R}_{i,m_i}^{-1} \mathbf{H}_{i,i} \tilde{\mathbf{g}}_{i,m_i}}{1 + \tilde{\mathbf{g}}_{i,m_i}^{\mathcal{H}} \mathbf{H}_{i,i}^{\mathcal{H}} \mathbf{R}_{i,m_i}^{-1} \mathbf{H}_{i,i} \tilde{\mathbf{g}}_{i,m_i}} \\ &\quad + \sum_{(j,m_j) \neq (i,m_i)} \frac{\lambda_j \mathbf{H}_{j,i}^{\mathcal{H}} \mathbf{R}_{j,m_j}^{-1} \mathbf{H}_{j,j} \tilde{\mathbf{g}}_{j,m_j} \tilde{\mathbf{g}}_{j,m_j}^{\mathcal{H}} \mathbf{H}_{j,j}^{\mathcal{H}} \mathbf{R}_{j,m_j}^{-1} \mathbf{H}_{j,i} \tilde{\mathbf{g}}_{i,m_i}}{1 + \tilde{\mathbf{g}}_{j,m_j}^{\mathcal{H}} \mathbf{H}_{j,j}^{\mathcal{H}} \mathbf{R}_{j,m_j}^{-1} \mathbf{H}_{j,j} \tilde{\mathbf{g}}_{j,m_j}}, \quad \forall i, m_i. \end{aligned} \quad (50)$$

$$\hat{\mathbf{R}}_{i,m_i} \succeq \frac{\zeta_{i,m_i}}{\gamma_{i,m_i}} \mathbf{H}_{i,i}^{\mathcal{H}} \mathbf{w}_{i,m_i} \mathbf{w}_{i,m_i}^{\mathcal{H}} \mathbf{H}_{i,i}, \quad \forall i, m_i \quad (60c)$$

where (60b) and (60c) are necessary to ensure that  $D(\cdot) > -\infty$ .

To establish that the dual problem (60) is equivalent to (P11), it is first noted that the optimal receive-beamformers  $\{\hat{\mathbf{g}}_{i,m_i}\}$  for (P11) can be obtained in closed form as

$$\hat{\mathbf{g}}_{i,m_i} = \hat{\mathbf{R}}_{i,m_i}^{\dagger} \mathbf{H}_{i,i}^{\mathcal{H}} \mathbf{w}_{i,m_i}, \quad \forall i, m_i \quad (61)$$

where  $\dagger$  denotes Moore-Penrose pseudoinverse, which is necessary here as  $\hat{\mathbf{R}}_{i,m_i}$  may be rank-deficient. Plugging (61) into (31b) yields

$$\frac{\zeta_{i,m_i}}{\gamma_{i,m_i}} \mathbf{w}_{i,m_i}^{\mathcal{H}} \mathbf{H}_{i,i} \hat{\mathbf{R}}_{i,m_i}^{\dagger} \mathbf{H}_{i,i}^{\mathcal{H}} \mathbf{w}_{i,m_i} \geq 1, \quad \forall i, m_i. \quad (62)$$

Consider next the following optimization problem

$$\max_{\substack{\mu_0 \geq 0, \dots, \mu_J \geq 0 \\ \sum_{i=0}^I \mu_i \bar{P}_i \leq \sum_{i=1}^I \bar{P}_i}} \max_{\{\zeta_{i,m_i} \geq 0\}, \{\gamma_{i,m_i}\}} \sum_{i,m_i} \zeta_{i,m_i} \quad (63a)$$

$$\text{subject to } \frac{\zeta_{i,m_i}}{\gamma_{i,m_i}} \mathbf{w}_{i,m_i}^{\mathcal{H}} \mathbf{H}_{i,i} \hat{\mathbf{R}}_{i,m_i}^{\dagger} \mathbf{H}_{i,i}^{\mathcal{H}} \mathbf{w}_{i,m_i} \leq 1, \quad \forall i, m_i \quad (63b)$$

in which the inequality in (62) is reversed, and the minimization in (31a) is replaced by maximization. Problem (63) admits identically optimal reward as (P11) because both achieve their optimum when the SINR constraints are met with equality; i.e.,

$$\frac{\zeta_{i,m_i}}{\gamma_{i,m_i}} \mathbf{w}_{i,m_i}^{\mathcal{H}} \mathbf{H}_{i,i} \hat{\mathbf{R}}_{i,m_i}^{\dagger} \mathbf{H}_{i,i}^{\mathcal{H}} \mathbf{w}_{i,m_i} = 1, \quad \forall i, m_i. \quad (64)$$

The remaining task is to prove that (63b) is equivalent to (60c). For this, [32, Lemma 1] states that for a positive semidefinite matrix  $\mathbf{R}$ ,  $\mathbf{R} \succeq \mathbf{b}\mathbf{b}^{\mathcal{H}}$  if and only if  $\mathbf{b}^{\mathcal{H}}\mathbf{R}^{\dagger}\mathbf{b} \leq 1$ . The desired result now follows immediately after substituting  $\mathbf{R} := \hat{\mathbf{R}}_{i,m_i}$  and  $\mathbf{b} := \sqrt{\frac{\zeta_{i,m_i}}{\gamma_{i,m_i}}} \mathbf{H}_{i,i}^{\mathcal{H}} \mathbf{w}_{i,m_i}$  in the said Lemma.

To determine the desired transmit-beamformers  $\{\tilde{\mathbf{g}}_{i,m_i}\}$ , equate to zero the gradient of  $L(\cdot)$  w.r.t.  $\tilde{\mathbf{g}}_{i,m_i}$  for  $\forall i, m_i$ ; that is

$$\frac{\partial L}{\partial \tilde{\mathbf{g}}_{i,m_i}} = \left[ \hat{\mathbf{R}}_{i,m_i} - \frac{\zeta_{i,m_i}}{\gamma_{i,m_i}} \mathbf{H}_{i,i}^{\mathcal{H}} \mathbf{w}_{i,m_i} \mathbf{w}_{i,m_i}^{\mathcal{H}} \mathbf{H}_{i,i} \right] \tilde{\mathbf{g}}_{i,m_i} = \mathbf{0}, \quad \forall i, m_i. \quad (65)$$

Thus, the optimal transmit beamformer  $\tilde{\mathbf{g}}_{i,m_i}$  can be represented as

$$\begin{aligned} \tilde{\mathbf{g}}_{i,m_i} &= \hat{\mathbf{R}}_{i,m_i}^{\dagger} \frac{\zeta_{i,m_i}}{\gamma_{i,m_i}} \mathbf{H}_{i,i}^{\mathcal{H}} \mathbf{w}_{i,m_i} \mathbf{w}_{i,m_i}^{\mathcal{H}} \mathbf{H}_{i,i} \tilde{\mathbf{g}}_{i,m_i} \\ &= \left( \frac{\zeta_{i,m_i}}{\gamma_{i,m_i}} \mathbf{w}_{i,m_i}^{\mathcal{H}} \mathbf{H}_{i,i} \mathbf{H}_{i,i} \tilde{\mathbf{g}}_{i,m_i} \right) \tilde{\mathbf{g}}_{i,m_i} \end{aligned} \quad (66)$$

from which it is clear that  $\tilde{\mathbf{g}}_{i,m_i}$  is collinear to  $\hat{\mathbf{g}}_{i,m_i}$ . Since phase rotation is irrelevant, only the transmit-powers  $\{P_{i,m_i}\}$  need to be determined. (The optimal transmit-beamformers can

then be computed as  $\tilde{\mathbf{g}}_{i,m_i} = \sqrt{P_{i,m_i}} \hat{\mathbf{g}}_{i,m_i} / \|\hat{\mathbf{g}}_{i,m_i}\|$ ,  $\forall i, m_i$ .) The transmit-powers can be obtained by solving the system of (32), based on the fact that the SINR constraints are met with equality at the optimum.

## REFERENCES

- [1] S. Boyd, Ellipsoid method, 2003 [Online]. Available: <http://www.stanford.edu/class/ee392o/elp.pdf>
- [2] M. C. Bromberg, "Optimizing MIMO multipoint wireless networks assuming Gaussian other-user interference," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2352–2362, Oct. 2003.
- [3] V. R. Cadambe and S. A. Jafar, "Interference alignment and degrees of freedom of the  $k$ -user interference channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3425–3441, Aug. 2008.
- [4] M. Chiang, C. W. Tan, D. P. Palomar, D. O'Neill, and D. Julian, "Power control by geometric programming," *IEEE Trans. Wireless Commun.*, vol. 6, no. 7, pp. 2640–2651, Jul. 2007.
- [5] M. Codreanu, A. Tölli, M. Juntti, and M. Latva-aho, "Joint design of Tx-Rx beamformers in MIMO downlink channel," *IEEE Trans. Signal Process.*, vol. 55, no. 9, pp. 4639–4655, Sep. 2007.
- [6] T. Cover and J. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.
- [7] A. Gjendemsjø, D. Gesbert, G. E. Øien, and S. G. Kiani, "Binary power control for sum rate maximization over multiple interfering links," *IEEE Trans. Wireless Commun.*, vol. 7, no. 8, pp. 3164–3173, Aug. 2008.
- [8] G. H. Golub and C. F. van Loan, *Matrix Computations*, 3rd ed. Baltimore, MD: Johns Hopkins Univ. Press, 1996.
- [9] K. Gomadam, V. R. Cadambe, and S. A. Jafar, "Approaching the capacity of wireless networks through distributed interference alignment," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, New Orleans, LA, Nov.–Dec. 2008, pp. 1–6.
- [10] D. Hoang and R. A. Iltis, "Noncooperative eigencoding for MIMO ad hoc networks," *IEEE Trans. Signal Process.*, vol. 56, no. 2, pp. 865–869, Feb. 2008.
- [11] J. Huang, R. A. Berry, and M. L. Honig, "Distributed interference compensation for wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 5, pp. 1074–1084, May 2006.
- [12] R. A. Iltis, S.-J. Kim, and D. Hoang, "Noncooperative iterative MMSE beamforming algorithms for ad hoc networks," *IEEE Trans. Commun.*, vol. 54, no. 4, pp. 748–759, Apr. 2006.
- [13] S. A. Jafar and S. Srinivasa, "Capacity limits of cognitive radio with distributed and dynamic spectral activity," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 3, pp. 529–537, Apr. 2007.
- [14] S.-J. Kim and G. B. Giannakis, "Optimal resource allocation for MIMO ad hoc cognitive radio networks," in *Proc. 46th Annu. Allerton Conf. Commun., Control, Comput.*, Monticello, IL, Sep. 2008, pp. 39–45.
- [15] Z.-Q. Luo and S. Zhang, "Dynamic spectrum management: Complexity and duality," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 57–73, Feb. 2008.
- [16] C. Michelot, "A finite algorithm for finding the projection of a point onto the canonical simplex of  $\mathbb{R}^n$ ," *J. Optim. Theory Appl.*, vol. 50, no. 1, pp. 195–200, Jul. 1986.
- [17] J. Mitola, III and G. Q. Maguire, Jr., "Cognitive radio: Making software radios more personal," *IEEE Pers. Commun.*, vol. 6, no. 4, pp. 13–18, Aug. 1999.
- [18] D. P. Palomar and M. Chiang, "A tutorial on decomposition methods for network utility maximization," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 8, pp. 1439–1451, Aug. 2006.
- [19] J. M. Peha, "Approaches to spectrum sharing," *IEEE Commun. Mag.*, vol. 43, no. 2, pp. 10–12, Feb. 2005.
- [20] S. W. Peters and R. W. Heath, Jr., "Interference alignment via alternating minimization," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Taipei, Taiwan, May 2009, pp. 2445–2448.
- [21] L. Qian, Y. J. Zhang, and J. Huang, "MAPEL: Achieving global optimality for a non-convex wireless power control problem," *IEEE Trans. Wireless Commun.*, vol. 8, no. 3, pp. 1553–1563, Mar. 2009.
- [22] M. Razaviyayn, M. Sanjabi, and Z.-Q. Luo, "Linear transceiver design for interference alignment: Complexity and computation," Sep. 2010 [Online]. Available: <http://arxiv.org/abs/1009.3481>
- [23] G. Scutari, D. P. Palomar, and S. Barbarossa, "Competitive design of multiuser MIMO systems based on game theory: A unified view," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 7, pp. 1089–1103, Sep. 2008.

- [24] G. Scutari, D. P. Palomar, and S. Barbarossa, "MIMO cognitive radio: A game theoretical approach," in *Proc. Workshop Signal Process. Adv. Wireless Commun.*, Recife, Brazil, Jul. 2008, pp. 426–430.
- [25] G. Scutari, D. P. Palomar, and S. Barbarossa, "Optimal linear precoding strategies for wideband noncooperative systems based on game theory—Parts I and II," *IEEE Trans. Signal Process.*, vol. 56, no. 3, pp. 1230–1267, Mar. 2008.
- [26] C. Shi, R. A. Berry, and M. L. Honig, "Monotonic convergence of distributed interference pricing in wireless networks," presented at the ISIT Conf., Seoul, Korea, Jun.–Jul. 2009.
- [27] A. Tölli, M. Codreanu, and M. Juntti, "Linear multiuser MIMO transceiver design with quality of service and per-antenna power constraints," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 3049–3055, Jul. 2008.
- [28] A. Wiesel, Y. C. Eldar, and S. Shamai, "Linear precoding via conic optimization for fixed MIMO receivers," *IEEE Trans. Signal Process.*, vol. 54, no. 1, pp. 161–176, Jan. 2006.
- [29] L. Xiao, M. Johansson, and S. P. Boyd, "Simultaneous routing and resource allocation via dual decomposition," *IEEE Trans. Commun.*, vol. 52, no. 7, pp. 1136–1144, Jul. 2004.
- [30] S. Ye and R. S. Blum, "Optimized signaling for MIMO interference systems with feedback," *IEEE Trans. Signal Process.*, vol. 51, no. 11, pp. 2839–2848, Nov. 2003.
- [31] W. Yu, G. Ginis, and J. M. Cioffi, "Distributed multiuser power control for digital subscriber lines," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 5, pp. 1105–1115, Jun. 2002.
- [32] W. Yu and T. Lan, "Transmitter optimization for the multi-antenna downlink with per-antenna power constraints," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2646–2660, Jun. 2007.
- [33] J. Yuan and W. Yu, "Distributed cross-layer optimization of wireless sensor networks: A game theoretic approach," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, San Francisco, CA, Nov. 2006, pp. 1–5.
- [34] 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, no. 1, pp. 38–51, Jan. 2008.
- [35] R. Zhang and Y.-C. Liang, "Exploiting multi-antennas for opportunistic spectrum sharing in cognitive radio networks," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 88–102, Feb. 2008.
- [36] Q. Zhao and B. M. Sadler, "A survey of dynamic spectrum access," *Signal Process. Mag.*, vol. 24, no. 3, pp. 78–89, May 2007.



**Seung-Jun Kim** (M'07) received the B.S. and M.S. degrees from Seoul National University, Seoul, Korea in 1996 and 1998, respectively, and the Ph.D. degree from the University of California, Santa Barbara, in 2005, all in electrical engineering.

From 2005 to 2008, he worked for NEC Laboratories America, Princeton, NJ, initially as a Visiting Researcher, and then as a Research Staff Member. Since 2008, he has been with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, where he is currently a Research Assistant Professor. His research interests lie in wireless communications and networking, statistical signal processing, and cross-layer optimization.



**Georgios B. Giannakis** (F'97) received the Diploma degree in electrical engineering from the National Technical University of Athens, Greece, in 1981 and the M.Sc. degree in electrical engineering, the M.Sc. degree in mathematics, and the Ph.D. degree in electrical engineering from the University of Southern California (USC) in 1983, 1986, and 1986, respectively.

Since 1999, he has been a Professor with the University of Minnesota, where he now holds an ADC Chair in Wireless Telecommunications in the Electric and Computer Engineering Department and serves as Director of the Digital Technology Center. His general interests span the areas of communications, networking and statistical signal processing subjects on which he has published more than 300 journal papers, 500 conference papers, two edited books, and two research monographs. Current research focuses on compressive sensing, cognitive radios, network coding, cross-layer designs, mobile ad hoc networks, wireless sensor, and social networks.

Dr. Giannakis is the (co)inventor of 20 patents issued and the (co)recipient of seven paper awards from the IEEE Signal Processing (SP) and Communications Societies, including the G. Marconi Prize Paper Award in Wireless Communications. He also received Technical Achievement Awards from the SP Society (2000), from EURASIP (2005), a Young Faculty Teaching Award, and the G. W. Taylor Award for Distinguished Research from the University of Minnesota. He is a Fellow of EURASIP, and has served the IEEE in a number of posts, including that of Distinguished Lecturer for the IEEE SP Society.