Power Control for Cooperative Dynamic Spectrum Access Networks with Diverse QoS Constraints

Nikolaos Gatsis, Antonio G. Marques, *Member, IEEE,*
and Georgios B. Giannakis, *Fellow, IEEE*

Abstract

Dynamic spectrum access (DSA) is an integral part of cognitive radio technology aiming at efficient management of the available power and bandwidth resources. The present paper deals with cooperative DSA networks, where collaborating terminals adhere to diverse (maximum and minimum) quality-of-service (QoS) constraints in order to not only effect hierarchies between primary and secondary users but also prevent abusive utilization of the available spectrum. Peer-to-peer networks with co-channel interference are considered in both single- and multi-channel settings. Utilities that are functions of the signal-to-interference-plus-noise-ratio (SINR) are employed as QoS metrics. By adjusting their transmit power, users can mitigate the generated interference and also meet the QoS requirements. A novel formulation accounting for heterogeneous QoS requirements is obtained after introducing a suitable relaxation and recasting a constrained sum-utility maximization as a convex optimization problem. The optimality of the relaxation is established under general conditions. Based on this relaxation, an algorithm for optimal power control that is amenable to distributed implementation is developed, and its convergence is established. Numerical tests verify the analytical claims and demonstrate performance gains relative to existing schemes.

I. Introduction

The Federal Communications Commission (FCC) has recognized that the perceived spectrum scarcity is caused by the currently inflexible bandwidth assignments [1]. In response to this problem,

Manuscript received September 21, 2008; revised June 22, 2009. This work was supported by the USDoD ARO Grant No. W911NF-05-1-0283; through collaborative participation in the Communications and Networks Consortium sponsored by the U. S. ARL under the CTA Program, Cooperative Agreement DAAD19-01-2-0011; the NSF grants CCF-0830480 and ECCS-0824007; and by the Gov. of C.A. Madrid Grant No. P-TIC-000223-0505. The U. S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation thereon. Part of this paper was presented at the *IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Las Vegas, NV, Mar.-Apr. 2008.

N. Gatsis and G. B. Giannakis are with the Dept. of ECE, Univ. of Minnesota, USA; A. G. Marques is with the Dept. of TSC, Rey Juan Carlos Univ., Spain (e-mails: {gatsi001, antonio, georgios}@umn.edu).

a spectrum policy reform has been proposed under the term dynamic spectrum access (DSA) [2]. The premise is allocation of the spectrum in a more flexible and market-driven manner, potentially by allowing services beyond those licensed, or, by accommodating more users, who may or may not be licensed. DSA is in fact an integral part of the emerging cognitive radio (CR) technology, which aims at enhancing spectrum utilization through ‘smart’ transceivers able to sense the operating environment and adapt to it; see e.g., [2] and references therein.

DSA schemes can be classified depending on whether users cooperate to share the available spectrum or not [2], [3]. In the non-cooperative setup, secondary (unlicensed) users either transmit over frequency slots not occupied by primary (licensed) users (spectrum overlay) or retain their transmission power below the primaries’ noise floor (spectrum underlay). On the other hand, more efficient sharing of the spectrum is expected in *cooperative* alternatives, for which two different models are typically considered. One is the *open sharing* model (also known as commons model), where all users are treated as ‘peers’ or primaries [2], [4], [1]. Such a network is envisioned to e.g., be deployed over an unlicensed band along with a set of rules to ensure efficient resource management. The second one is a *flexible primary* model, where primary users negotiate access with secondary users [3], if e.g., the latter pay a fee for using a pre-specified level of the resources.

The present work deals with resource allocation in cooperative DSA networks for both open sharing and flexible primary models. Design challenges addressed include the accommodation of diverse application-specific constraints, mechanisms for encouraging efficient spectrum utilization, and decentralizing the management schemes, as advocated by the FCC. This paper’s main contribution is the incorporation of *diverse* (heterogeneous) individual QoS requirements. In a flexible primary model, access is regulated by bounding the maximum level of a commodity a secondary user receives, which may be communication rate, bit error rate, or any other QoS figure; while ensuring a minimum level for primary users. In an open sharing model, users voluntarily adapt usage of network resources to their application requirements. This way, minimum and maximum bounds on the received QoS become constraints that the resource allocation task must account for [5], [6].

Focus here is placed on peer-to-peer networks where users transmit over the same bandwidth both in single- and multi-channel settings. The co-channel interference present in such networks intimately couples individual power control decisions. Each user’s satisfaction with the received QoS level is captured by utility functions that depend on the received signal-to-interference-plus-noise ratio (SINR). Adjusting the individual transmit power offers the potential to satisfy the individual QoS requirements and is a critical network task. The required power control scheme is obtained by solving a sum-utility maximization problem subject to maximum and minimum utility (or SINR) constraints. Two features of this novel approach are: (i) incorporation of heterogeneous QoS requirements and (ii) a provably convergent algorithm for optimal power control amenable to distributed implementations.

In recent years, the design of resource allocation schemes for CR and DSA networks has received considerable attention. Maximization of network utility with diverse QoS constraints in cooperative CR has been pursued in [5], but orthogonal access and a central controller were assumed. Different decentralized power control algorithms maximizing the total utility in networks with non-orthogonal access (e.g., CDMA) but without accounting for individual users' QoS constraints were presented in [7], [8]. Minimum SINR constraints were also accommodated in [9, Chapter 4], [10, Sec. 3.3], but maximum ones were not included. More recently, two suboptimal algorithms for distributed power control in multi-channel DSA networks with diverse QoS constraints have been reported in [6].

The rest of the paper is organized as follows. In Section II, the optimal power control in single-channel networks is formulated and a convex relaxation to enable its efficient solution is introduced. An algorithm for optimal power control amenable to distributed implementation is developed in Section III. Results for multi-channel networks are presented in Section IV, while simulations in Section V and conclusions in Section VI wrap up this paper.

II. Optimal Power Control

Consider the power control problem for a single-channel (i.e., single-carrier) DSA network in which users share the same frequency band, e.g., as in CDMA. Assuming a peer-to-peer operating setup, there is a set of $\mathcal{M} := \{1, \ldots, M\}$ links, where each link $i \in \mathcal{M}$ entails a user with a dedicated transmitter (Tx$_i$) wishing to communicate with a corresponding receiver (Rx$_i$), as in [7]. The terms pair, user and link will be used interchangeably. Let h_{ij} denote the (power) path gain from Tx$_i$ to Rx$_j$, assumed static. The path gain h_{ij} models the relationship between the transmitted and received power and captures any signal processing technique taking place at the transmitter or the receiver, such as (de-)spreading in CDMA. Also, let n_i denote the noise power at Rx$_i$; p_i the transmission power[1] of Tx$_i$; and $p_i^{\max}$ the maximum power budget Tx$_i$ can afford, i.e., $0 \leq p_i \leq p_i^{\max}$. The received

[1] Although the power values here are considered continuous, adaptive modulation schemes may welcome a discrete set of power levels. The optimal design then also requires the continuous solution pursued in this paper as a first step, is highly non-trivial, and goes beyond the scope and space limits of this paper; see, e.g., [11] and references thereof.

SINR γ_i at Rx$_i$ is a function of the powers $\boldsymbol{p} := [p_1, \ldots, p_M]^T$ given by $\gamma_i := h_{ii}p_i / \left(n_i + \sum_{k \neq i} h_{ki}p_k\right)$. Let us define vectors $\boldsymbol{p}^{\max} := [p_1^{\max}, \ldots, p_M^{\max}]^T$, $\boldsymbol{\gamma} := [\gamma_1, \ldots, \gamma_M]^T$, $\boldsymbol{\eta} := [n_1/h_{11}, \ldots, n_M/h_{MM}]^T$; and the matrix $\mathbf{A} = [a_{ij}]$ with $a_{ij} := h_{ji}/h_{ii}$ if $i \neq j$ and $a_{ij} := 0$ if $i = j$. Also let $\mathbf{D}(\boldsymbol{x})$ denote an $M \times M$ diagonal matrix with diagonal elements $[x_1, \ldots, x_M]^T := \boldsymbol{x}$.

The utility associated with each link $i \in \mathcal{M}$ will be described by a generic function $u_i(\gamma_i)$. The goal is to maximize the sum of all link utilities subject to QoS constraints. The QoS per link i will also be generically described by a function $v_i(\gamma_i)$, which can e.g., represent rate when $v_i(\gamma_i) = \ln(1 + \gamma_i)$. If $v_i(\gamma_i)$ is chosen monotonic, then constraints on v_i map one-to-one to SINR bounds; i.e., $v_i(\gamma_i) \in [v_i(\gamma_i^{\min}), v_i(\gamma_i^{\max})] \Leftrightarrow \gamma_i \in [\gamma_i^{\min}, \gamma_i^{\max}]$. The lower bounds ensure a minimum QoS level while the upper bounds prevent abuse of the available resources. Recall that these are design objectives in both flexible-primary as well as in open-sharing DSA models. For both models, the associated power control problem for DSA/CR networks amounts to solving the following:

$$\max_{\mathbf{0} \leq \boldsymbol{p} \leq \boldsymbol{p}^{\max}} \quad \sum_{i=1}^{M} u_i(\gamma_i) \tag{1a}$$

$$\text{subj. to} \qquad \gamma_i^{\min} \leq \gamma_i \leq \gamma_i^{\max}, \quad \forall\, i \in \mathcal{M}. \tag{1b}$$

In most DSA setups, not all constraints in (1b) will be present. Indeed, $\gamma_i^{\max}$ may not be enforced if i is a primary user; while if i is a secondary user both $\gamma_i^{\max}$ and $\gamma_i^{\min}$ may (or may not) be present.

The maximum QoS requirements is the key difference between problem (1) and related ones in power control for non-orthogonal access networks. These requirements capture the design objectives for certain DSA networks, which would be difficult with existing formulations. For example, while properly selected spectral masks regulating transmit power can limit the interference received by other users, they cannot guarantee that the received SINR will not exceed a prescribed level. Similarly, judicious choices of utilities, e.g., proportionally fair, cannot ensure that the received SINR (and hence QoS) is within an allowable range if (1b) is absent.

Problem (1) is generally non-convex and hence challenging to solve, especially in a distributed fashion suitable for the peer-to-peer setup at hand. Upon selecting $\{u_i(\cdot)\}$ properly, a convex reformulation of (1) is possible using the methods in [12]. Such reformulation could only be solved in a centralized manner, while the methods in [12] do not readily lead to algorithms. Moreover, the special case of (1) with minimum SINR constraints *only* is addressed in [9], [10] for certain utilities, but the solutions developed in these works cannot handle two-sided SINR constraints.

A novel approach to solving (1) is described in the ensuing subsection. It entails a suitable relaxation, which allows the use of convex optimization and will also form the basis for the design of the distributed power allocation algorithm presented in Section III.

A. *Efficient optimization via convex relaxation*

To solve (1) efficiently, we adopt the following assumptions:

AS1. *The individual utilities are chosen so that: (a) $u_i(\gamma_i)$ are strictly increasing and twice continuously differentiable; and (b) $-\gamma_i u_i''(\gamma_i)/u_i'(\gamma_i) \geq 1$ for $\gamma_i > 0$ ($'$ denotes differentiation).*

AS2. *The noise power is non-zero for all i, i.e., $n_i > 0$; and the gain matrix $\mathbf{A}$ is irreducible.*

AS3. *If every user has a maximum SINR constraint, there is no power vector $\tilde{\boldsymbol{p}}$ with $\mathbf{0} < \tilde{\boldsymbol{p}} \leq \boldsymbol{p}^{\max}$ such that the resulting SINRs $\tilde{\gamma}_i$ satisfy $\tilde{\gamma}_i = \gamma_i^{\max}$ for all $i \in \mathcal{M}$.*

AS1 is standard in the power control literature [13, Chapter 5]. Specifically, it implies that $u_i(\gamma_i)$ is strictly concave in γ_i and effects the fairness condition $\lim_{\gamma_i \to 0^+} u_i(\gamma_i) = -\infty$ [9, p. 15], which guarantees that non-zero power is allocated to all users. Examples of utilities satisfying AS1 are $u_i(\gamma_i) = \ln \gamma_i$, and $u_i(\gamma_i) = \gamma_i^{\alpha}/\alpha$ with $\alpha < 0$ [13, Sec. 5.2.5]. Although AS1 refers only to the utilities u_i in (1a), the v_i functions used to obtain the SINR constraints (1b) are not restricted by any condition other than being monotonic. Furthermore, the irreducibility of $\mathbf{A}$ in AS2 is also a standard assumption in power control problems [12].

AS3 pertains to the case where all users have maximum SINR constraints. In this case, the equations $\gamma_i = \gamma_i^{\max}$, $i = 1, \ldots, M$, can be easily written as a system of linear equations in $\boldsymbol{p}$ (cf. (13a)). AS3 then means that this linear system has no solution satisfying $\mathbf{0} < \boldsymbol{p} \leq \boldsymbol{p}^{\max}$. Satisfaction of AS3 can be checked as explained in Section III. But even when it is not satisfied, $\tilde{\boldsymbol{p}}$ in AS3 is the optimal solution of (1) and no further optimization is needed, because the $u_i(\gamma_i)$ are strictly increasing and all users can achieve their $\gamma_i^{\max}$. Last but not least, AS3 is automatically satisfied when primary users do not upper-bound their QoS, i.e., when $\gamma_i^{\max} = \infty$ for some i.

Having clarified the operating conditions, we will relax (1) to facilitate its solution through convex optimization. To this end, let q_i denote an auxiliary variable associated with link i, upper-bounding the interference-plus-noise (IpN) term $n_i + \sum_{k \neq i} h_{ki} p_k$. Collecting all variables q_i in $\boldsymbol{q} := [q_1, \ldots, q_M]^T$, consider the following relaxed version of (1) ($\mathbb{R}_{++}$ denotes the positive reals):

$$\max_{\mathbf{0} \leq \boldsymbol{p} \leq \boldsymbol{p}^{\max};\, \boldsymbol{q} \in \mathbb{R}_{++}^M} \quad \sum_{i=1}^{M} u_i(h_{ii} p_i q_i^{-1}) \tag{2a}$$

$$\text{subj. to} \quad \gamma_i^{\min} \leq h_{ii} p_i q_i^{-1} \leq \gamma_i^{\max}, \forall\, i \in \mathcal{M} \tag{2b}$$

$$q_i \geq n_i + \sum_{k\neq i} h_{ki} p_k, \quad \forall\, i \in \mathcal{M}. \tag{2c}$$

Clearly, if (2c) were equality constraints, then (1) and (2) would be equivalent. In order for the relaxation to be useful, two issues need to be addressed: (i) optimality of the relaxation needs to be established, i.e., that the solution of (2) is also a solution of (1); and (ii) problem (2) must be efficiently solvable. Using the change of variables $p_i = e^{y_i}$ and $q_i = e^{z_i}$, we have shown in [14] that AS1 ensures convexity of problem (2) in $\boldsymbol{y} := [y_1, \ldots, y_M]^T$ and $\boldsymbol{z} := [z_1, \ldots, z_M]^T$; hence, (ii) is settled. To address (i), we prove in Appendix A the following.

Proposition 1. *Assume that* (1) *is feasible, and let AS1a, AS2 and AS3 hold. If* $\boldsymbol{p}^*, \boldsymbol{q}^*$ *solve* (2), *then* (2c) *holds as equality at* $\boldsymbol{p}^*, \boldsymbol{q}^*$*; i.e.,*

$$q_i^* = n_i + \sum_{k\neq i} h_{ki} p_k^* \quad \forall i \in \mathcal{M}. \tag{3}$$

Proposition 1 asserts that the optimal powers for problems (1) and (2) are identical and the optimal $\boldsymbol{q}^*$ of problem (2) is given by (3). It also follows from Proposition 1 that the values of the optimal sum-utility in (1) and (2) are identical. Hence, the relaxation incurs no loss of optimality.

Interestingly, Proposition 1 holds for *any* strictly increasing utility, e.g., $\ln(1+\gamma_i)$; that is, convexity is not required. Nonetheless, it is the convexity guaranteed by AS1 together with Proposition 1 that facilitate efficient optimization of the power allocation in (2), as explained in Section III.

It is remarked that introduction of local IpN variables and a related relaxation appear in [15], and also as a method to accommodate general interference functions in [9, Chapter 4]. Nevertheless, the optimality of the relaxation in (2) cannot follow from any of these works.

The convex relaxation of (1) has been carried out in two steps: first by introducing q_i, and then by transforming (p_i, q_i) into (y_i, z_i). The next remark elaborates on why the form of the relaxed problem is potentially solvable in a distributed fashion.

Remark 1. *The relaxed problem* (2) *has two features which facilitate a distributed solution:*
(a) The objective in (2a) *is a sum of* M *utility functions, one for each user. Moreover, each utility* $u_i(.)$, $i = 1, \ldots, M$, *depends only on the variables* p_i *and* q_i, *pertaining to user* i*; and*
(b) For each user i, *the constraints* (2b) *and* (2c) *depend only on* p_i, q_i, *as well the IpN* $n_i + \sum_{k\neq i} h_{ki} p_k$. *This quantity seemingly 'couples' all optimization variables. The key element though is that* $n_i + \sum_{k\neq i} h_{ki} p_k$ *in* (2c) *can be measured at receiver* i.
These features (a) and (b) are also present in problem (1). *Unlike* (2), *problem* (1) *is non-convex and cannot be rendered convex while retaining (a) and (b).*

III. Power Allocation Algorithm for Single-Channel Networks

In this section, an algorithm based on Lagrangian techniques is developed to solve (1) via (2).[2] This algorithm will have provable convergence, exhibit tracking capability, entail low complexity and be suitable for distributed implementation, features certainly desirable in DSA/CR networks.

Before solving (2), the validity of AS3 must be ensured by checking whether there are powers solving $\gamma_i = \gamma_i^{\max}$ for all $i \in \mathcal{M}$ with feasible $\boldsymbol{p} \leq \boldsymbol{p}^{\max}$. This can be checked using the standard power control algorithm of [16, eq. (21)], which has guaranteed convergence and can be implemented in a distributed fashion without information exchange among users. If *all* maximum SINR constraints are exactly met, then the powers returned by this algorithm are the optimal solution of (1), due to AS1a. If not, these powers may be used as initialization for the solver of (2), developed next.

With the objective of solving (2), set $y_i^{\max} := \ln p_i^{\max}$, $\mathcal{Y} := \prod_{i=1}^{M}(-\infty, y_i^{\max}]$ and observe that in addition to (2b) and (2c), problem (2) has an additional convex set constraint $(\boldsymbol{y}, \boldsymbol{z}) \in \mathcal{Y} \times \mathbb{R}^M$. Let ν_i, λ_i, μ_i denote Lagrange multipliers corresponding to minimum and maximum SINR constraints (2b) and (2c), respectively. The Lagrangian function of the convex equivalent of (2) is then

$$L(\boldsymbol{y}, \boldsymbol{z}, \boldsymbol{\nu}, \boldsymbol{\lambda}, \boldsymbol{\mu}) := -\sum_i u_i\left(\frac{h_{ii}e^{y_i}}{e^{z_i}}\right) + \sum_i \mu_i\left[e^{-z_i}\left(n_i + \sum_{k\neq i} h_{ki}e^{y_k}\right) - 1\right] + \sum_i \nu_i\left(\gamma_i^{\min}\frac{e^{z_i}}{h_{ii}e^{y_i}} - 1\right) + \sum_i \lambda_i\left(\frac{1}{\gamma_i^{\max}}\frac{h_{ii}e^{y_i}}{e^{z_i}} - 1\right). \quad (4)$$

For brevity, let $\boldsymbol{\omega} := \{\boldsymbol{y}, \boldsymbol{z}, \boldsymbol{\nu}, \boldsymbol{\lambda}, \boldsymbol{\mu}\}$ denote all optimization variables and Lagrange multipliers. Problem (2) is solved via the following first-order algorithm that utilizes the gradient of $L(\boldsymbol{\omega})$ to simultaneously update primal and dual variables with constant stepsize β and $[x]^+ := \max\{0, x\}$:

$$y_i(t+1) = \min\left\{y_i(t) - \beta\frac{\partial L(\boldsymbol{\omega})}{\partial y_i}\bigg|_{\boldsymbol{\omega}(t)}, y_i^{\max}\right\} \quad (5a)$$

$$z_i(t+1) = z_i(t) - \beta\frac{\partial L(\boldsymbol{\omega})}{\partial z_i}\bigg|_{\boldsymbol{\omega}(t)} \quad (5b)$$

$$\nu_i(t+1) = \left[\nu_i(t) + \beta\left(\gamma_i^{\min}e^{z_i(t)-y_i(t)}/h_{ii} - 1\right)\right]^+ \quad (5c)$$

$$\lambda_i(t+1) = \left[\lambda_i(t) + \beta\left(h_{ii}e^{y_i(t)-z_i(t)}/\gamma_i^{\max} - 1\right)\right]^+ \quad (5d)$$

$$\mu_i(t+1) = \left[\mu_i(t) + \beta\left(e^{-z_i(t)}\left(n_i + \sum_{k\neq i} h_{ki}e^{y_k(t)}\right) - 1\right)\right]^+. \quad (5e)$$

[2]Throughout this section references to (2) will in fact refer to its convex equivalent after the transformation $p_i = e^{y_i}$ and $q_i = e^{z_i}$.

The gradient $\nabla_{\boldsymbol{\omega}} L(\boldsymbol{\omega})$ is used in (5) to minimize $L(\boldsymbol{\omega})$ with respect to $\boldsymbol{y}$, $\boldsymbol{z}$, and maximize it with respect to $\boldsymbol{\nu}$, $\boldsymbol{\lambda}$, $\boldsymbol{\mu}$; i.e., a saddle point is sought. Convergence is analyzed in the next subsection.

From an implementation perspective, it is worth stressing that in compliance with FCC, the power constraints are respected *throughout the iterations* due to the projection operation in (5a). In addition, updates in (5) use a constant β, which enables tracking and is thus attractive for mobile CR networks. Means of distributing the iterations (5) are explored in Subsection III-B.

A. Convergence and sensitivity analysis

In order to analyze the convergence of (5), an additional assumption is due:

AS4. *Problem* (2) *is strictly feasible, i.e., there exist* $\bar{\boldsymbol{p}}$, $\bar{\boldsymbol{q}}$ *with* $\mathbf{0} < \bar{\boldsymbol{p}} \leq \boldsymbol{p}^{\max}$ *such that* (2b) *and* (2c) *hold as strict inequalities.*

This last assumption corresponds to Slater's constraint qualification, which guarantees the existence of optimal Lagrange multipliers [17, Sec. 3.3.5]. Capitalizing on AS4, the following lemma characterizes the optimal Lagrange multipliers of (2); its proof is in Appendix A.

Lemma 1. *If* (1) *is feasible and AS1-AS4 hold, then: (i) the optimal Lagrange multipliers for constraints* (2c) *are positive, i.e.,* $\boldsymbol{\mu}^* > \mathbf{0}$*; and (ii) the Lagrangian function at the optimal Lagrange multipliers,* $L(\boldsymbol{y}, \boldsymbol{z}, \boldsymbol{\nu}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$*, is strictly convex in* $\boldsymbol{y}$ *and* $\boldsymbol{z}$ *over* $\mathbb{R}^{2M}$*.*

The first part of Lemma 1 is a strict complementary slackness result, which in general does not follow from the Karush-Kuhn-Tucker (KKT) necessary conditions for optimality; for details on these notions, see e.g., [17, Sec. 3.3]. Moreover, notice that part (ii) of Lemma 1 holds even for utilities that are not strictly convex in $\boldsymbol{y}$ and $\boldsymbol{z}$, e.g., $u_i(h_{ii}e^{y_i}/e^{z_i}) = \ln(h_{ii}e^{y_i}/e^{z_i})$.

Now let $\mathrm{dist}(\boldsymbol{x}, \mathcal{X}) := \min_{\boldsymbol{\xi} \in \mathcal{X}} ||\boldsymbol{x} - \boldsymbol{\xi}||_2$ denote the distance of a point $\boldsymbol{x}$ from a set $\mathcal{X}$; and Ω^* the set of optimal $\boldsymbol{\omega}$ vectors. Using Lemma 1, the following proposition establishes the global convergence of iterations (5) to a neighborhood of Ω^*.

Proposition 2. *Suppose* (1) *is feasible, and AS1-AS4 hold. For any* ϵ *and* δ *with* $0 < \epsilon < \delta$*, there exist positive* $\beta_0(\epsilon, \delta)$ *and* $t_0(\epsilon, \delta)$ *such that for any stepsize* $0 < \beta \leq \beta_0(\epsilon, \delta)$ *and any initial point* $\boldsymbol{\omega}(0) \in \mathcal{Y} \times \mathbb{R}^M \times \mathbb{R}_+^{3M}$ *with* $\mathrm{dist}(\boldsymbol{\omega}(0), \Omega^*) \leq \delta$*, the iterates* $\boldsymbol{\omega}(t)$ *in* (5) *satisfy* $\mathrm{dist}(\boldsymbol{\omega}(t), \Omega^*) \leq \epsilon$ *for all* $t \geq t_0(\epsilon, \delta)/\beta$*.*

Proposition 2 asserts that the iterates $\boldsymbol{\omega}(t)$ reach (and remain within) an arbitrarily small neighborhood of Ω^* from *any* initial point. The stepsize and the number of iterations depend on the initialization and the desired neighborhood size. The proof provided in Appendix A relies on

Lemma 1. The numerical examples presented in Section V will demonstrate that the iterations not only remain arbitrarily close to the optimal solution, but actually converge.

It is well-known that the activation of a constraint in an optimization problem entails a penalty in the achieved optimal value. Sensitivity analysis can be used to study the effect of changes in the constraints on the optimal utility value. Such analysis is pertinent when the constraints are fixed beforehand (e.g., if they are QoS levels dictated by a specific application), but also when they have to be settled by the system designer. A brief sensitivity analysis for problem (1) (via (2)) is presented next. Since incorporating maximum SINR constraints is the main feature of (1), the focus here is on the effect of varying $\gamma_i^{\max}$. The analysis for the minimum SINR constraint is similar.

To specify the problem, let λ_i^*, $i = 1, \ldots, M$, be the optimal Lagrange multipliers returned by (5) and u_{tot}^* the optimal value of problem (2); and hence of (1) in view of Proposition 1. Suppose that $\gamma_i^{\max}$ is changed to $\gamma_i^{\max} + \delta_i \gamma_i^{\max}$, $\delta_i \in \mathbb{R}$. The objective is to quantify the effect of $\delta_i \gamma_i^{\max}$ on u_{tot}^*. Both smaller as well as larger changes of $\boldsymbol{\delta} := [\delta_1, \ldots, \delta_M]^T$ are of interest.

Let $u_{\text{tot}}(\boldsymbol{\delta})$ be the optimal value of (1) and (2) under the aforementioned perturbation, and suppose that AS3 holds also with $\gamma_i^{\max} + \delta_i \gamma_i^{\max}$ instead of $\gamma_i^{\max}$. With this notation, $u_{\text{tot}}^* = u_{\text{tot}}(\mathbf{0})$. The effects of small values of $\boldsymbol{\delta}$ are studied first. To this end, the value of the derivative of $u_{\text{tot}}(\boldsymbol{\delta})$ can be used, and it is computed next based on known quantities.

Let $\{\mathbf{e}_i\}_{i=1}^M$ denote the vectors in $\mathbb{R}^M$ with 1 on the i-th component and 0 elsewhere. Also let $\Theta^* \subset \mathbb{R}^{3M}$ denote the set of optimal Lagrange multiplier vectors $[\boldsymbol{\nu}^T, \boldsymbol{\lambda}^T, \boldsymbol{\mu}^T]^T$ of (2). Under AS1–AS4, [18, Theorem 2.3.2] asserts that $u_{\text{tot}}(\boldsymbol{\delta})$ has directional derivative in any direction in $\mathbb{R}^M$; its values in the directions $\mathbf{e}_i$ and $-\mathbf{e}_i$ along with bounds for the derivative values are listed in Table I. These bounds depend on $\gamma_i^{\max}$ and the optimal λ_i^* returned by (5); hence, they are easily computable. The first bound is immediate; the second is derived by setting $\partial L/\partial z_i = 0$ (cf. (18)), using (3) and assuming that the $\gamma_i^{\max}$ constraint is active, so that $\gamma_i^* = \gamma_i^{\max}$ and $\nu_i^* = 0$.

The derivatives are used to evaluate the increase or decrease of the sum-utility value when the SINR constraints $\gamma_i^{\max}$ change. In particular, if $\gamma_i^{\max}$ is changed to $\gamma_i^{\max} + \delta_i \gamma_i^{\max}$ with $\delta_i > 0$ small, then u_{tot}^* is increased by $D_{\mathbf{e}_i} u_{\text{tot}}(\mathbf{0}) \cdot \delta_i$ approximately. On the other hand, if $\gamma_i^{\max}$ is decreased to $\gamma_i^{\max} - \delta_i \gamma_i^{\max}$ with $\delta_i > 0$ again small, then u_{tot}^* is approximately decreased by $D_{-\mathbf{e}_i} u_{\text{tot}}(\mathbf{0}) \cdot \delta_i$.

The optimal multipliers λ_i^* can also be used to assess the effect of larger changes in the perturbation $\boldsymbol{\delta}$. The following inequality holds for all $\boldsymbol{\delta} \in \mathbb{R}^M$ (cf. [19, eq. (6.23)])

$$u_{\text{tot}}(\boldsymbol{\delta}) \leq u_{\text{tot}}^* + \sum_{i=1}^{M} \lambda_i^* \delta_i. \qquad (6)$$

Inequality (6) offers an upper bound on the optimal sum-utility with the following qualitative implications. If λ_i^* is large and $\delta_i < 0$, then the sum-utility decreases considerably. If λ_i^* is small and $\delta_i > 0$, then the sum-utility increases, but not much. Note though that from inequality (6) one cannot draw conclusions for other combinations of signs of δ_i and values of λ_i^*.

B. Distributed implementation

To develop a distributed counterpart of (5), consider the derivatives in (5a) and (5b)

$$\frac{\partial L}{\partial y_i} = -u_i'\left(\frac{h_{ii}e^{y_i}}{e^{z_i}}\right)\frac{h_{ii}e^{y_i}}{e^{z_i}} + e^{y_i}\sum\nolimits_{j\neq i} h_{ij}\mu_j e^{-z_j} + \frac{\lambda_i}{\gamma_i^{\max}}\frac{h_{ii}e^{y_i}}{e^{z_i}} - \nu_i\gamma_i^{\min}\frac{e^{z_i}}{h_{ii}e^{y_i}} \tag{7a}$$

$$\frac{\partial L}{\partial z_i} = u_i'\left(\frac{h_{ii}e^{y_i}}{e^{z_i}}\right)\frac{h_{ii}e^{y_i}}{e^{z_i}} - \mu_i e^{-z_i}\left(n_i + \sum\nolimits_{k\neq i} h_{ki}e^{y_k}\right) - \frac{\lambda_i}{\gamma_i^{\max}}\frac{h_{ii}e^{y_i}}{e^{z_i}} + \nu_i\gamma_i^{\min}\frac{e^{z_i}}{h_{ii}e^{y_i}}. \tag{7b}$$

The updates (5) take place at Tx_i. It is assumed that Rx_i is able to estimate the gain h_{ii} and the SINR $h_{ii}e^{y_i(t)}/(n_i+\sum_{k\neq i} h_{ki}e^{y_k(t)})$, and feed the latter back to its peer Tx_i per time slot t. Tx_i needs also to obtain h_{ii} via feedback but this may happen only during the start-up phase provided that h_{ii} changes at a scale much slower than the algorithm's convergence time. Then, all terms needed for the updates (5) are known locally at Tx_i, with the exception of the sum $\sum_{j\neq i} h_{ij}\mu_j(t)e^{-z_j(t)}$, which is associated with the IpN constraints in (2c).

In order to make the aforementioned sum available at Tx_i, two schemes that have been proposed for power control problems different from (2) can be adapted to the problem at hand: message passing [10, Sec. 3.4], [7], [6], and "the reversed network" [13, Sec. 6.5], [8], [9, Chapter 4]. The latter has the attractive feature of not requiring exchange of information among links.

1) Message passing: Users in this scheme exchange information over a control channel to facilitate power management decisions, as in e.g., [4, Sec. 3.2.3]. To be specific, each Tx_j *broadcasts* its variable $\mu_j(t)e^{-z_j(t)}$, which can be readily interpreted as the current estimate of the cost paid due to local interference. Moreover, each Tx_i needs to know the path gains h_{ij} of the links causing interference to the non-peer receivers Rx_j. This is possible if reciprocity holds and the Rx_j transmits a training signal; alternatively, Tx_i can transmit a training signal so that Rx_j estimates h_{ij} and feeds it back. The quantities involved in the message passing are illustrated in Fig. 1.

2) Reversed network: All links here are assumed reciprocal. Every receiver becomes a transmitter and vice-versa. In order to use the reversed network, the term $e^{y_i}\sum_{j\neq i} h_{ij}\mu_j e^{-z_j}$ of $\partial L/\partial y_i$ in (7a) is re-written as $e^{y_i}\sum_{j=1}^{M} h_{ij}\mu_j e^{-z_j} - e^{y_i}h_{ii}\mu_i e^{-z_i}$. The main idea is that the sum $\sum_{j=1}^{M} h_{ij}\mu_j e^{-z_j} \geq 0$ represents received power at each Tx_i when all transmitters of the reversed network (i.e., all Rx_j)

transmit simultaneously symbols with power $\mu_j e^{-z_j}$. These symbols do not need to be known at the Tx_i; only the *total received power* needs to be estimated.

Notice that each $\mu_i e^{-z_i}$ term is unknown at Rx_i, but known at Tx_i. The feature that the power for the reversed network transmission is unknown at the corresponding transmitters is not present in previous works. In order to address this, variables $z_i(t)$, $\mu_i(t)$, $\lambda_i(t)$, $\nu_i(t)$ are also updated at Rx_i. The key is that each receiver already measures all quantities needed for these updates, namely the received power $h_{ii}e^{y_i(t)}$ and the IpN term $n_i + \sum_{k \neq i} h_{ki}e^{y_k(t)}$ in order to have an estimate of the current SINR. Clearly, for the peers Tx_i and Rx_i to have identical copies of $z_i(t)$, $\mu_i(t)$, $\lambda_i(t)$ and $\nu_i(t)$, the initializations must be identical, requiring only coordination between peers.

IV. Multi-Channel Networks

The approach pursued so far will be generalized in this section to devise globally convergent algorithms for optimal power control in multi-channel networks. Due to space limitation, emphasis will be placed on stressing the differences with respect to the single-channel case.

A. Optimal power control

Users here may transmit over an orthogonal set of frequency bands $\mathcal{F} := \{1, \ldots, F\}$, also referred to as channels, subcarriers or tones. The power of Tx_i on channel f is $p_{i,f}$, the noise power at Rx_i on channel f is $n_{i,f}$, and the (power) path gain from Tx_i to Rx_j on channel f is $h_{ij,f}$. Moreover, each user adheres to a *spectral mask* $p_{i,f} \leq p_{i,f}^{\max}$, and maximum power budget $\sum_f p_{i,f} \leq p_i^{\max}$. Hence, each user's power must lie in $\mathcal{P}_i := \{p_{i,f} | 0 \leq p_{i,f} \leq p_{i,f}^{\max} \; \forall \, f \in \mathcal{F}; \sum_f p_{i,f} \leq p_i^{\max}\}$. The received SINR at Rx_i on channel f is $\gamma_{i,f} := h_{ii,f}p_{i,f}/(n_{i,f} + \sum_{k \neq i} h_{ki,f}p_{k,f})$; vector $\boldsymbol{p}_i := [p_{1,f}, \ldots, p_{M,f}]^T$ contains the power loadings for user i; and similar to the single-channel case, $\mathbf{A}_f$ is the gain matrix for channel f.

The aim is to formulate the power control problem for a multi-channel network incorporating diverse QoS constraints. Two ways of generalizing the QoS bounds in (1) are possible: (i) individual bounds per user; and (ii) individual bounds per user and channel.[3] The optimal solution of (ii) can be readily obtained by implementing the single-channel algorithm of Section III per channel, and projecting $\boldsymbol{p}_i$ onto $\mathcal{P}_i$ per iteration. For this reason, emphasis here is placed on generalization (i).

[3]As a way of illustration, suppose QoS is measured in terms of rate. Clearly (i) corresponds to bounding the aggregate rate of each user (sum-rate across channels), while (ii) corresponds to bounding each user's rate on every channel.

The QoS that each user receives is an aggregate measure of the performance attained when all channels are utilized. Utility functions $u_{i,f}$, $U_{i,f}$ and $V_{i,f}$ model the contribution of the performance over individual channels $f \in \mathcal{F}$ to the total QoS. These functions may represent different performance measures; one example is communication rate. The performance over an individual channel is a function of the SINR γ_i^f; this is made explicit by writing $u_{i,f}(\gamma_i^f)$, $U_{i,f}(\gamma_i^f)$ and $V_{i,f}(\gamma_i^f)$. Furthermore, the contribution of the per-channel utility to the total QoS is linear. Therefore the sums $\sum_{f\in\mathcal{F}} u_{i,f}(\gamma_i^f)$, $\sum_{f\in\mathcal{F}} U_{i,f}(\gamma_i^f)$ and $\sum_{f\in\mathcal{F}} V_{i,f}(\gamma_i^f)$ are measures of the total QoS per user. The first amounts to the objective to be maximized, the second is used to ensure minimum QoS $U_i^{\min}$, and the third to set an upper bound on the received QoS $V_i^{\max}$. Thus, the optimization problem generalizing (1) to multi-channel networks is

$$\max_{\{\boldsymbol{p}_i\in\mathcal{P}_i\,\forall\, i\in\mathcal{M}\}} \sum\nolimits_{i=1}^{M}\sum\nolimits_{f=1}^{F} u_{i,f}(\gamma_{i,f}) \tag{8a}$$

$$\text{subj. to} \quad \sum\nolimits_{f=1}^{F} U_{i,f}(\gamma_{i,f}) \geq U_i^{\min} \text{ and } \sum\nolimits_{f=1}^{F} V_{i,f}(\gamma_{i,f}) \leq V_i^{\max} \quad \forall\, i \in \mathcal{M}. \tag{8b}$$

Recall that in the single-channel case QoS constraints are mapped one-to-one to SINR constraints when link-specific utilities are selected to be monotonic (cf. (1b)). For this reason, there was no need to introduce $U_{i,f}(\gamma_{i,f})$ and $V_{i,f}(\gamma_{i,f})$ in the optimization problem (1). But this is impossible for the multi-channel generalization in (8) because the sum-utilities are involved in (8b).

Similar to the single-channel case, a solution to (8) will be pursued through a suitable relaxation. With $\boldsymbol{q}_i := [q_{1,f},\ldots,q_{M,f}]^T$ representing the local IpN vector, we will solve:

$$\max_{\{\boldsymbol{p}_i\in\mathcal{P}_i,\,\boldsymbol{q}_i\in\mathbb{R}_{++}^{M}\,\forall\, i\in\mathcal{M}\}} \sum\nolimits_{i=1}^{M}\sum\nolimits_{f=1}^{F} u_{i,f}(h_{ii,f}p_{i,f}/q_{i,f}) \tag{9a}$$

$$\text{subj. to} \quad \sum\nolimits_{f=1}^{F} U_{i,f}(h_{ii,f}p_{i,f}/q_{i,f}) \geq U_i^{\min}, \forall\, i \in \mathcal{M} \tag{9b}$$

$$\sum\nolimits_{f=1}^{F} V_{i,f}(h_{ii,f}p_{i,f}/q_{i,f}) \leq V_i^{\max}, \forall\, i \in \mathcal{M} \tag{9c}$$

$$q_{i,f} \geq n_{i,f} + \sum\nolimits_{j\neq i} h_{ji,f}p_{j,f}, \qquad \forall\, i \in \mathcal{M},\, \forall\, f \in \mathcal{F}. \tag{9d}$$

The assumptions that will ensure optimality and convexity of the relaxed problem are:

AS5. *Utilities $u_{i,f}(\gamma_{i,f})$ are chosen so that: (a) $u_{i,f}(\gamma_{i,f})$ are strictly increasing, twice continuously differentiable, with $\lim_{\gamma_{i,f}\to 0^+} u_{i,f}(\gamma_{i,f}) = -\infty$; and (b) $-\gamma_i u_i''(\gamma_i)/u_i'(\gamma_i) \geq 1$ for $\gamma_i > 0$.*

AS6. *Utilities $U_{i,f}(\gamma_{i,f})$ satisfy AS1.*

AS7. *Utilities $V_{i,f}(\gamma_{i,f})$ are chosen so that: (a) are strictly increasing, concave, and twice continuously differentiable; and (b) satisfy $-\gamma_{i,f}V_{i,f}''(\gamma_{i,f})/V_{i,f}'(\gamma_{i,f}) \leq 1$ for $\gamma_{i,f} > 0$.*

AS8. *It holds that* $n_{i,f} > 0$ *for all* i *and* f, *and gain matrix* $\mathbf{A}_f$ *is irreducible for all* f.

AS9. *If every user has a maximum utility constraint (cf.* (9c)*), there are no* $\tilde{\boldsymbol{p}}_i$, $\tilde{\boldsymbol{q}}_i$ *with* $\tilde{\boldsymbol{p}}_i \in \mathcal{P}_i$, $\tilde{\boldsymbol{q}}_i \in \mathbb{R}^M_{++}$ *such that* (9c) *holds with equality for all* i.

As in the single-channel case, AS5-AS7 guarantee the convexity of (9) under the transformation $p_{i,f} = e^{y_{i,f}}$, $q_{i,f} = e^{z_{i,f}}$. Examples of utilities satisfying AS7 are $V_{i,f}(\gamma_{i,f}) = \ln \gamma_{i,f}$, $V_{i,f}(\gamma_{i,f}) = \gamma_{i,f}$, and $V_{i,f}(\gamma_{i,f}) = \ln(1 + \gamma_{i,f})$. Utilities satisfying AS5 and AS6 are those satisfying AS1. Similar to [7], the fairness condition in AS5a precludes assignment of zero power to any channel, which may be restrictive for some multi-channel systems. Note also that if just one terminal does not upper-bound its QoS (e.g., when primary users are present), AS9 is satisfied. However, different from the single-channel case, there is no standard algorithm available to validate AS9 for the hypothetical case of all users meeting their maximum QoS constraints with equality.

The optimality of the relaxation is established in the following result, proved in Appendix B.

Proposition 3. *Assume that problem* (8) *is feasible, and AS5a, AS6a, AS7a, AS8, and AS9 hold. Then at the optimal solution* $p^*_{i,f}$, $q^*_{i,f}$ *of* (9)*, constraint* (9d) *holds as equality, i.e.,*

$$q^*_{i,f} = n_{i,f} + \sum\nolimits_{j \neq i} h_{ji,f} p^*_{j,f} \quad \forall\, i \in \mathcal{M},\, \forall\, f \in \mathcal{F}. \tag{10}$$

Proposition 3 states that the optimal power allocations as well as the optimal objective values of (8) and (9) coincide. As with Proposition 1, no assumption on convexity is needed. Furthermore, Proposition 3 implies that an efficient solution of (8) can be found via (9); this is pursued next.

B. Power allocation algorithm

Let ν_i, λ_i be Lagrange multipliers for the minimum and maximum QoS constraints, (9b) and (9c), and $\mu_{i,f}$ for (9d). Also let $\boldsymbol{y}$, $\boldsymbol{z}$, $\boldsymbol{\nu}$, $\boldsymbol{\lambda}$, $\boldsymbol{\mu}$ denote vectors collecting variables $y_{i,f}$, $z_{i,f}$, ν_i, λ_i, $\mu_{i,f}$, respectively, for all i and f. The notation $\boldsymbol{\omega}$ is used for $\boldsymbol{y}$, $\boldsymbol{z}$, $\boldsymbol{\nu}$, $\boldsymbol{\lambda}$, $\boldsymbol{\mu}$ collectively. Further, define $\mathcal{Y}_i := \{y_{i,f} | y_{i,f} \leq \ln p^{\max}_{i,f}\ \forall\, f \in \mathcal{F}; \sum_f e^{y_{i,f}} \leq p^{\max}_i\}$ and $\mathcal{Y} := \prod_{i=1}^M \mathcal{Y}_i$. The Lagrangian of (9) is

$$\begin{aligned} L(\boldsymbol{\omega}) := &- \sum\nolimits_{i,f} u_{i,f}(h_{ii,f} e^{y_{i,f}} / e^{z_{i,f}}) + \sum\nolimits_{i,f} \mu_{i,f} \Big[e^{-z_{i,f}} \Big(n_{i,f} + \sum\nolimits_{k \neq i} h_{ki,f} e^{y_{k,f}} \Big) - 1 \Big] \\ &- \sum\nolimits_i \nu_i \Big(\sum\nolimits_f U_{i,f}(h_{ii,f} e^{y_{i,f}} / e^{z_{i,f}}) - U^{\min}_i \Big) + \sum\nolimits_i \lambda_i \Big(\sum\nolimits_f V_{i,f}(h_{ii,f} e^{y_{i,f}} / e^{z_{i,f}}) - V^{\max}_i \Big). \end{aligned} \tag{11}$$

As in Section III, a first-order (gradient) algorithm is employed to solve (9) iteratively using

$$\boldsymbol{y}_i(t+1) = [\boldsymbol{y}_i(t) - \beta \nabla_{\boldsymbol{y}_i} L(\boldsymbol{\omega}(t))]_{\mathcal{Y}_i} \tag{12a}$$

$$\boldsymbol{z}_i(t+1) = \boldsymbol{z}_i(t) - \beta \nabla_{\boldsymbol{z}_i} L(\boldsymbol{\omega}(t)) \tag{12b}$$

$$\nu_i(t+1) = \Big[\nu_i(t) + \beta\Big(-\sum\nolimits_f U_{i,f}(h_{ii,f}e^{y_{i,f}(t)}/e^{z_{i,f}(t)}) + U_i^{\min}\Big)\Big]^+ \tag{12c}$$

$$\lambda_i(t+1) = \Big[\lambda_i(t) + \beta\Big(\sum\nolimits_f V_{i,f}(h_{ii,f}e^{y_{i,f}(t)}/e^{z_{i,f}(t)}) - V_i^{\max}\Big)\Big]^+ \tag{12d}$$

$$\mu_{i,f}(t+1) = \Big[\mu_{i,f}(t) + \beta\Big(e^{-z_{i,f}(t)}\Big(n_{i,f} + \sum\nolimits_{k\neq i} h_{ki,f}e^{y_{k,f}(t)}\Big) - 1\Big)\Big]^+ \tag{12e}$$

where β is a constant stepsize, and $[\boldsymbol{x}]_{\mathcal{Y}_i}$ is the projection of $\boldsymbol{x}$ onto the set $\mathcal{Y}_i$. Since $\mathcal{Y}_i$ is a closed convex set, the projection in (12a) can be implemented efficiently. Note that spectral mask and sum-power constraints are respected throughout the algorithm, thanks to the projection in (12a).

The convergence analysis parallels the single-channel case; AS10, Lemma 2 and Proposition 4 are the counterparts of AS4, Lemma 1 and Proposition 2, respectively. Proofs are in Appendix B.

AS10. *Problem* (9) *is strictly feasible, i.e., there exist* $\bar{\boldsymbol{p}}$, $\bar{\boldsymbol{q}}$ *with* $\bar{\boldsymbol{p}}_i \in \mathcal{P}_i$, $\bar{\boldsymbol{q}}_i \in \mathbb{R}^M_{++}$ *for all* i *such that* (9b), (9c), *and* (9d) *hold with strict inequality.*

Lemma 2. *If* (1) *is feasible and AS5-AS10 hold, then: (i) the optimal Lagrange multipliers for constraints* (9d) *are positive, i.e.,* $\boldsymbol{\mu}^* > \mathbf{0}$*; and (ii) the Lagrangian function at the optimal Lagrange multipliers,* $L(\boldsymbol{y}, \boldsymbol{z}, \boldsymbol{\nu}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$*, is strictly convex in* $\boldsymbol{y}$ *and* $\boldsymbol{z}$ *over* $\mathbb{R}^{2MF}$.

Proposition 4. *Assume that* (1) *is feasible and AS5-AS10 hold. For any* ϵ *and* δ *with* $0 < \epsilon < \delta$*, there exist positive* $\beta_0(\epsilon,\delta)$ *and* $t_0(\epsilon,\delta)$ *such that for any stepsize* $0 < \beta \leq \beta_0(\epsilon,\delta)$ *and any initial point* $\boldsymbol{\omega}(0) \in \mathcal{Y} \times \mathbb{R}^{MF} \times \mathbb{R}_+^{M(F+2)}$ *with* $\mathrm{dist}(\boldsymbol{\omega}(0), \Omega^*) \leq \delta$*, the iterates* $\boldsymbol{\omega}(t)$ *in* (12) *satisfy* $\mathrm{dist}(\boldsymbol{\omega}(t), \Omega^*) \leq \epsilon$ *for all* $t \geq t_0(\epsilon,\delta)/\beta$*, where* Ω^* *is the set of optimal* $\boldsymbol{\omega}$ *vectors.*

Distributed implementation: It can be easily verified that if path gains $h_{ii,f}$ and SINR for all channels are fed back from Rx_i, then all terms in (12) are known at Tx_i, except the sum $\sum_{j\neq i} h_{ij,f}\mu_{j,f}(t) e^{-z_{j,f}(t)}$ for all f. For the latter to become available, message passing or the reversed network approach can be utilized. The operations are the same as in the single-channel case, with the additional feature that they are performed for every channel f.

V. Numerical results

Numerical tests are presented in this section to corroborate the analytical claims and also to compare the performance of the developed algorithm with that of various existing algorithms.

Test case 1: Single-channel networks. Consider a peer-to-peer network using CDMA. With d_{ij} denoting the distance between Tx_i and Rx_j and B the spreading gain, it is assumed that gains h_{ij} follow a (deterministic) path loss model with $h_{ii} = d_{ii}^{-4}$ and $h_{ij} = B^{-1}d_{ij}^{-4}$ for $i \neq j$. In this case, matrix $\mathbf{A}$ is irreducible (cf. AS2). The parameters describing the setup tested are listed in Table II, while the Tx_i-Rx_i positions are shown in Table III. The selected utility satisfies AS1. First, algorithm (5) is applied to power control without constraints, and it is seen to obtain the same power allocation as other algorithms in the literature used for this problem. Then, focus is turned to a problem with minimum and maximum QoS constraints. In this case, the QoS requirements adopted are similar to those in [6, Sec. 7], mapped to SINR values, and listed in Table II as well.

The developed algorithm is applied first to power control without constraints, namely for the solution of (1a). This is done by setting very small minimum SINR constraints and very large maximum SINR constraints, so that they are all inactive. In this case, AS3 is automatically satisfied. The values selected are $\gamma_i^{\min} = 10^{-5}$ and $\gamma_i^{\max} = 10^5$ for all i. There are several algorithms in the literature which solve (1a) optimally under AS1, namely ADP [7], gradient projection for minimization [8], and variable splitting [9, Sec. 4.3]; results from all these will be the same. The optimal sum-utility and SINR per user obtained with the developed algorithm (labeled as "Lagrangian") and the ones in [7], [8] are listed in Table IV. The results are identical, as expected.

Consider next a problem having diverse QoS constraints with values listed in Table II. Algorithms QoS-ps-DSA and QoSe-DSA in [6] rely on game theory to solve (1). Each of these is developed in general for multichannel networks and each has two versions: in one version power is allocated over all channels (MC-QoS-ps-DSA, MC-QoSe-DSA), while in the other only one channel is selected for transmission (SC-QoS-ps-DSA, SC-QoSe-DSA). In order to solve (1), the algorithms are restricted to the case where there is a single available channel; then the two versions (MC- and SC-) reduce to the same algorithm. The sum-utility and SINR per user achieved by the Lagrangian algorithm and the two alternatives are provided in Table V, where the SINRs violating the constraints are shown in boldface. For completeness, the SINRs obtained from the standard power control algorithm are listed in the last column of Table V. Observe that $\gamma_i < \gamma_i^{\max}$ for $i \in \{1, 5, 6, 8\}$, confirming that AS3 indeed holds. These values were used to initialize (5). It is observed that QoS-ps-DSA and QoSe-DSA cannot always meet all users' SINR requirements (although these requirements are feasible, see, e.g., user 1). Note also that the sum-utility is not maximized (compare 32.4 with 23.6). On the other hand, it is expected that the optimal sum-utility of the unconstrained problem (1a) will be higher than that of (1) because the constraints (1b) are imposed on the SINRs. This is quantified in this test by comparing the corresponding entries of Tables IV and V.

Time trajectories of powers and Lagrange multipliers are depicted in Fig. 2. The plots corroborate that the proposed iterations converge (cf. Proposition 2), and the fact that all the IpN constraints are active ($\mu_i^* > 0$), as asserted by Lemma 1. However, although the convergence is relatively fast (100-

300 iterations), this number is one order of magnitude higher than its suboptimal game-theoretic counterparts QoS-ps-DSA and QoSe-DSA. This happens because convergence of the Lagrange multipliers slows down to satisfy the diverse (two-sided) QoS requirements.

Test Case 2: Multi-Channel Networks. Each Tx_i-Rx_i pair is placed on the same position as in the previous test case, but now a frequency selective model is tested. Specifically, there are $F = 16$ channels available and each path gain $h_{ij,f}$ is obtained from a realization of a 4-tap channel. The taps follow Rayleigh fading, are equally spaced, and have power delay profile (1,1/2,1/8,1/10). The realizations across links are independent. The path loss over each channel follows the model with $h_{ij,f} = d_{ij}^{-4}$. The remaining parameters are listed in Table VI.

First, algorithm (12) is used for the solution of the unconstrained problem (8a), using $U_i^{\min} = -150$ and $V_i^{\max} = 150$. The objective value ($\sum_{i,f} u_{i,f}(\gamma_{i,f})$) and the sum-utility per user ($\sum_f u_{i,f}(\gamma_{i,f})$ for $i = 1, \ldots, M$) are listed in Table VII. The corresponding ones obtained from DADP [7], which solves (8a) optimally, are also shown in Table VII. The results coincide, as expected.

When the QoS constraints of Table VI are imposed, results obtained by different algorithms are listed in Table VIII. Algorithms MC-QoS-ps-DSA and MC-QoSe-DSA attempt to solve (8) [6]. As in the single-channel case, the results of Table VIII illustrate that existing schemes might not always satisfy all QoS constraints, and may achieve lower objective value than the Lagrangian algorithm.

VI. CONCLUSIONS

Power control algorithms were developed for DSA networks with primary and secondary users or peer users willing to cooperate. A distinct feature of the novel design is the incorporation of diverse (maximum and/or minimum) QoS constrains per user. Peer-to-peer networks with co-channel interference were considered for both single- and multi-channel settings. The QoS level of each user was captured through utility functions that depend on the received SINR.

The novel power control algorithm has been obtained as the solution of a sum-utility maximization subject to maximum and minimum utility (or SINR) constraints. The presence of interference intimately couples the users' power control decisions and represents a challenge to develop efficient optimal solutions. However, a two-step relaxation rendering the problem convex and amenable to distributed implementation was presented for a broad class of utilities.

Using this relaxation, a first-order Lagrangian method that simultaneously updates primal and dual variables was developed and its convergence to the optimum solution established. Two distributed implementations were also introduced. Finally, numerical tests confirming the analytical claims and comparing the performance gains relative to existing schemes were presented.[4]

APPENDIX

A. Single-channel networks

To prove Proposition 1, the following lemma, which applies to the case where all users have maximum SINR constraints, is required.

Lemma 3. *If AS2 holds and there is no* $\boldsymbol{p}$ *in the feasible set of* (1) *such that* $\gamma_i = \gamma_i^{\max}$ *for all* $i \in \mathcal{M}$ *(cf. AS3), then there are no* $\boldsymbol{p}$*,* $\boldsymbol{q}$ *in the feasible set of* (2) *such that* $h_{ii}p_i/q_i = \gamma_i^{\max}$ *for all* $i \in \mathcal{M}$.

Proof of Lemma 3: The feasibility problem of the SINRs $\gamma_i^{\max}$ in (1) can be written as

$$\boldsymbol{p} = \mathbf{D}(\boldsymbol{\gamma}^{\max})\mathbf{A}\boldsymbol{p} + \mathbf{D}(\boldsymbol{\gamma}^{\max})\boldsymbol{\eta} \tag{13a}$$

$$\mathbf{0} < \boldsymbol{p} \leq \boldsymbol{p}^{\max}. \tag{13b}$$

If the spectral radius of $\mathbf{D}(\boldsymbol{\gamma}^{\max})\mathbf{A}$ (see [20, p. 35] for a definition) satisfies $\rho(\mathbf{D}(\boldsymbol{\gamma}^{\max})\mathbf{A}) < 1$, then the linear system in (13a) accepts a unique positive solution $\boldsymbol{p}(\boldsymbol{\gamma}^{\max}) := (\mathbf{I} - \mathbf{D}(\boldsymbol{\gamma}^{\max})\mathbf{A})^{-1}\,\mathbf{D}(\boldsymbol{\gamma}^{\max})\boldsymbol{\eta}$; see, e.g., [13, Theorem A.35]. Since (13) does not have a solution by assumption, then either $\rho(\mathbf{D}(\boldsymbol{\gamma}^{\max})\mathbf{A}) \geq 1$, or, $\rho(\mathbf{D}(\boldsymbol{\gamma}^{\max})\mathbf{A}) < 1$ but with $\boldsymbol{p}(\boldsymbol{\gamma}^{\max}) \nleq \boldsymbol{p}^{\max}$.

Achievability of $\boldsymbol{\gamma}^{\max}$ in (2) can now be posed as the following feasibility problem in $\boldsymbol{p}$, $\boldsymbol{q}$:

$$\gamma_i^{\max} = h_{ii}p_i/q_i, \quad q_i \geq n_i + \sum\nolimits_{k \neq i} h_{ki}p_k, \; \forall i \in \mathcal{M}; \quad \mathbf{0} < \boldsymbol{p} \leq \boldsymbol{p}^{\max}. \tag{14}$$

Clearly $\boldsymbol{q}$ can be eliminated, so (14) becomes

$$\boldsymbol{p} \geq \mathbf{D}(\boldsymbol{\gamma}^{\max})\mathbf{A}\boldsymbol{p} + \mathbf{D}(\boldsymbol{\gamma}^{\max})\boldsymbol{\eta} \tag{15a}$$

$$\mathbf{0} < \boldsymbol{p} \leq \boldsymbol{p}^{\max}. \tag{15b}$$

If $\rho(\mathbf{D}(\boldsymbol{\gamma}^{\max})\mathbf{A}) \geq 1$, then (15a) cannot have a nonnegative solution ($\boldsymbol{p} \geq \mathbf{0}$). Otherwise, the Subinvariance Theorem [13, Lemma A.37] and $\boldsymbol{\eta} > \mathbf{0}$ leads to a contradiction.

If $\rho(\mathbf{D}(\boldsymbol{\gamma}^{\max})\mathbf{A}) < 1$, the solutions of (15a) form a cone with apex $\boldsymbol{p}(\boldsymbol{\gamma}^{\max})$, and $\boldsymbol{p} \geq \boldsymbol{p}(\boldsymbol{\gamma}^{\max})$ for all $\boldsymbol{p}$ in the cone [21]. If $\boldsymbol{p}(\boldsymbol{\gamma}^{\max}) \nleq \boldsymbol{p}^{\max}$, then (15) represents an empty set [21, Lemma 3]. □

[4]The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U. S. Government.

Proof of Proposition 1: First note that the feasibility of (1) implies the feasibility of (2), and a solution to (2) exists due to Weierstrass Theorem [17, Prop. A.8].

Having shown the existence of solution to (2), the proof of (3) is by contradiction. Assume that there exists a user i with *dominant* q_i, meaning that at the optimum (2c) is inactive for user i, i.e.,

$$q_i^* > n_i + \sum\nolimits_{k \neq i} h_{ki} p_k^*. \tag{16}$$

If all users have maximum SINR constraints, then from Lemma 3 it follows that at the optimum of (2) (in fact at any feasible $\boldsymbol{p}, \boldsymbol{q}$ of (2)) at least one user m will have inactive $\gamma_m^{\max}$; i.e., $\gamma_m^* = h_{mm} p_m^* / q_m^* < \gamma_m^{\max}$. Any such user at the optimal point must have *non-dominant* q_m^*, i.e.,

$$q_m^* = n_m + \sum\nolimits_{k \neq i} h_{km} p_k^*; \tag{17}$$

otherwise, q_m^* could be reduced, yielding higher objective value. In the case of at least one not having maximum SINR constraint, (17) obviously holds (for that user). Comparing (16) with (17), it follows that $i \neq m$. Moreover, since it has been assumed that q_i^* is dominant, then $h_{ii} p_i^* / q_i^* = \gamma_i^{\max}$. Thus, the user set $\mathcal{M}$ can be divided into three disjoint groups $\mathcal{G}_1$, $\mathcal{G}_2$, $\mathcal{G}_3$ (cf. Fig. 3). In $\mathcal{G}_1$ are the users with inactive (or absent) $\gamma_m^{\max}$ (these must have non-dominant q_m^*). Groups $\mathcal{G}_2$, $\mathcal{G}_3$ contain users with active max SINR constraint and in particular $\mathcal{G}_2$ contains the ones with dominant q_i.

Now consider the user $i \in \mathcal{G}_2$ with dominant q_i^* (cf. (16)) and active $\gamma_i^{\max}$; and the user $m \in \mathcal{G}_1$ with non-dominant q_m^* (cf. (17)). Due to the irreducibility of $\mathbf{A}$ there exists a sequence of distinct indices $i = k_0, k_1, \ldots, k_{l-1}, k_l = m$ with the property $\{k_1, \ldots, k_{l-1}\} \in \mathcal{G}_2 \cup \mathcal{G}_3$ for *some* $m \in \mathcal{G}_1$ such that the corresponding channels are *positive*, i.e., $h_{k_0 k_1} > 0, \ldots, h_{k_{l-1}, k_l} > 0$ [20, Sec. 6.2].

The main argument is that one can successively decrease $p_{k_\iota}^*$ and $q_{k_\iota}^*$ for $\iota = 0, 1, \ldots, l-1$, but keep the same 'local SINR' $h_{k_\iota k_\iota} p_{k_\iota}^* / q_{k_\iota}^* = \gamma_{k_\iota}^{\max}$, until reaching user m with inactive $\gamma_m^{\max}$. Note that $p_{k_\iota}^* > 0$ for $\iota = 0, 1, \ldots, l-1$, since $\gamma_{k_\iota}^{\max} > 0$. Specifically, attempt to decrease both p_i^*, q_i^* by the same proportion, i.e., set $\check{p}_i = \alpha_{k_0} p_i^*$, $\check{q}_i = \alpha_{k_0} q_i^*$ with $\alpha_{k_0} < 1$. The resulting 'local SINR' for i is still maximum, but $q_{k_1}^*$ has become dominant since $h_{k_0 k_1} > 0$, i.e., $q_{k_1}^* > n_{k_1} + \sum_{k \neq i, k_1} h_{k,k_1} p_k^* + h_{i k_1} \check{p}_i$. Then $p_{k_1}^*$ and $q_{k_1}^*$ can be reduced, rendering $q_{k_2}^*$ dominant. Proceeding likewise across $\iota = 0, \ldots, l-1$, $p_{k_\iota}^*$ and $q_{k_\iota}^*$ can be reduced, yielding

$$q_{k_{\iota+1}}^* > n_{k_{\iota+1}} + \sum_{k \in \{k_0, \ldots, k_\iota\}} h_{k, k_{\iota+1}} \check{p}_k + \sum_{k \notin \{k_0, \ldots, k_\iota\}} h_{k, k_{\iota+1}} p_k^*.$$

When user $m \in \mathcal{G}_1$ is reached (i.e., $\iota + 1 = l$), q_m^* is decreased but without changing p_m^* (recall that $\gamma_m^* < \gamma_m^{\max}$). This yields a higher γ_m, and hence higher objective value, which is a contradiction. □

Now proofs of Lemma 1 and Proposition 2 are provided; footnote 2 also applies here.

Proof of Lemma 1: (i) Since problem (2) has an additional convex set constraint, $(\boldsymbol{y}, \boldsymbol{z}) \in \mathcal{Y} \times \mathbb{R}^M$, we use the necessary conditions of [17, Prop. 3.3.11]. These conditions are more general than the KKT, in that they also include a multiplier for the gradient of the objective function (not only the constraints). But when Slater's constraint qualification holds (cf. AS4), such a multiplier is not needed (see e.g., [17, pp. 334–335]). Due to the special structure of the constraint set ($y_i \leq y_i^{\max}$, $z_i \in \mathbb{R}$), the first of the aforementioned conditions can be written as

$$\left.\frac{\partial L}{\partial y_i}\right|_{(\boldsymbol{y}^*,\boldsymbol{z}^*,\boldsymbol{\nu}^*,\boldsymbol{\lambda}^*,\boldsymbol{\mu}^*)} \leq 0, \ \left.\frac{\partial L}{\partial z_i}\right|_{(\boldsymbol{y}^*,\boldsymbol{z}^*,\boldsymbol{\nu}^*,\boldsymbol{\lambda}^*,\boldsymbol{\mu}^*)} = 0, \ \forall i \in \mathcal{M}. \tag{18}$$

It will be shown that $\boldsymbol{\mu}^* > \mathbf{0}$. This cannot be concluded from $\partial L/\partial z_i = 0$ alone (using (7b) into (18)), due to the term arising from the maximum SINR constraint. Substituting (7b) into (18) and (7a) into $\partial L/\partial y_i = -\theta_i$ for some $\theta_i \geq 0$, summing the previous two equations, arranging them into matrix form and using (3), gives the equation for the optimal $\boldsymbol{\mu}^*$

$$[\mathbf{I} - \mathbf{D}(e^{y_i^*})\mathbf{A}^T\mathbf{D}(h_{ii}e^{-z_i^*})]\boldsymbol{\mu}^* = \boldsymbol{\theta}, \tag{19}$$

where slightly abusing notation, here $\mathbf{D}(x_i)$ denotes an $M \times M$ diagonal matrix with elements $x_1, \ldots, x_M$ on the diagonal. The matrix $\mathbf{D}(e^{y_i^*})\mathbf{A}^T\mathbf{D}(h_{ii}e^{-z_i^*})$ is irreducible, and has column sums smaller than 1 due to (3) and $n_i > 0$; hence $\rho[\mathbf{D}(e^{y_i^*})\mathbf{A}^T\mathbf{D}(h_{ii}e^{-z_i^*})] < 1$ [20, Theorem 8.1.22]. Furthermore, we have $\boldsymbol{\theta} \geq \mathbf{0}$ and $\boldsymbol{\theta} \neq \mathbf{0}$ (the reason why $\boldsymbol{\theta} \neq \mathbf{0}$ will be explained soon). Now using [13, Theorem A.36] it follows readily that the solution of system (19) is *positive*, i.e., $\boldsymbol{\mu}^* > \mathbf{0}$.

Assume that $\boldsymbol{\theta} = \mathbf{0}$. Since $\rho[\mathbf{D}(e^{y_i^*})\mathbf{A}^T\mathbf{D}(h_{ii}e^{-z_i^*})] < 1$, matrix $\mathbf{I} - \mathbf{D}(e^{y_i^*})\mathbf{A}^T\mathbf{D}(h_{ii}e^{-z_i^*})$ is invertible and the solution of (19) is $\boldsymbol{\mu}^* = \mathbf{0}$. Now from AS3, there is a user i for whom $\gamma_i^* < \gamma_i^{\max}$. From the (weak) complementary slackness condition in [17, Prop. 3.3.11], it follows that $\lambda_i^* = 0$. Setting (7b) to zero (cf. (18)) and substituting $\lambda_i^* = 0$, AS1a yields $\mu_i^* > 0$, contradicting $\boldsymbol{\mu}^* = \mathbf{0}$.

(ii) The main idea is to show that the Hessian (with respect to the primal variables $\boldsymbol{y}$, $\boldsymbol{z}$) of the Lagrangian function (4) evaluated at the optimal Lagrange multipliers is *positive definite* for all $(\boldsymbol{y}, \boldsymbol{z}) \in \mathbb{R}^{2M}$. In particular, the Hessian is positive semidefinite, since problem (2) is convex. Here it is shown that for the optimal Lagrange multipliers, the Hessian is invertible for all $(\boldsymbol{y}, \boldsymbol{z}) \in \mathbb{R}^{2M}$.

The Hessian with respect to the primal variables $\boldsymbol{y}$, $\boldsymbol{z}$ takes the partitioned form

$$\nabla^2 L(\boldsymbol{y}, \boldsymbol{z}, \boldsymbol{\nu}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*) = \begin{bmatrix} \nabla^2_{\boldsymbol{yy}} L & \nabla_{\boldsymbol{y}}\nabla_{\boldsymbol{z}} L \\ \nabla_{\boldsymbol{z}}\nabla_{\boldsymbol{y}} L & \nabla^2_{\boldsymbol{zz}} L \end{bmatrix} := \begin{bmatrix} \mathbf{H}_{11} & \mathbf{H}_{12} \\ \mathbf{H}_{21} & \mathbf{H}_{22} \end{bmatrix}. \tag{20}$$

Diagonal blocks $\mathbf{H}_{11}$, $\mathbf{H}_{22}$ (not shown for brevity) are diagonal matrices, with *positive* diagonal elements due to AS1, AS2, and $\boldsymbol{\mu}^* > \mathbf{0}$. The off-diagonal blocks satisfy $\mathbf{H}_{21} = \mathbf{H}_{12}^T$ with

$$\mathbf{H}_{12} = \mathbf{D}[u_i''(h_{ii}e^{y_i}/e^{z_i})(h_{ii}e^{y_i}/e^{z_i})^2 + u_i'(h_{ii}e^{y_i}/e^{z_i})(h_{ii}e^{y_i}/e^{z_i})] \\ - \mathbf{D}(e^{y_i})\mathbf{A}^T\mathbf{D}(\mu_i^* h_{ii}/e^{z_i}) - \mathbf{D}[(\lambda_i^* h_{ii}e^{y_i})/(e^{z_i}\gamma_i^{\max}) + (\nu_i^*\gamma_i^{\min}e^{z_i})/(h_{ii}e^{y_i})].$$

The off-diagonal blocks $\mathbf{H}_{12}$ and $\mathbf{H}_{21}$ are nonpositive matrices. To show that the Hessian is nonsingular, we apply [22, Chapter 6, Theorem 2.3, Condition (J_{30})]. The vector that satisfies the aforementioned condition for the Hessian matrix here is the vector of length $2M$ with 1 in each entry. Then, with $(\mathbf{H})_{ij}$ denoting the i, j entry of the Hessian, the condition becomes

$$\sum_{j=1}^{i}(\mathbf{H})_{ij} > 0, \quad \sum_{j=1}^{M+i}(\mathbf{H})_{M+i,j} > 0, \quad i = 1, \ldots, M. \tag{21}$$

It holds that $\sum_{j=1}^{i}(\mathbf{H})_{ij} = (\mathbf{H})_{ii}$ and $\sum_{j=1}^{M+i}(\mathbf{H})_{M+i,j} = \mu_i^* n_i/e^{z_i}$, $i = 1, \ldots, M$. Then the first condition in (21) is true because the diagonal entries of $\mathbf{H}_{11}$ are positive; while the second holds because $\boldsymbol{\mu}^* > \mathbf{0}$ and $n_i > 0$ (cf. AS2). □

Proof of Proposition 2: The iterations (5) solve for a saddle point of the Lagrangian (4) over $\mathcal{Y} \times \mathbb{R}^M \times \mathbb{R}_+^{3M}$. So the first step is to assert that the optimal $\boldsymbol{\omega}$'s in (2) are exactly these saddle points. Then the convergence claim is proved directly after invoking [23, Theorem 1], and therefore it suffices to show that the conditions required by the theorem are satisfied.

Indeed, the optimal primal solutions and geometric multipliers of (2) are exactly the saddle points of (4) over $\mathcal{Y} \times \mathbb{R}^M \times \mathbb{R}_+^{3M}$ [17, Prop. 5.1.6]. But the geometric multipliers coincide with the Lagrange multipliers associated with the optimal solution [19, Prop. 6.1.2] since the problem is convex and a solution exists (cf. the proof of Proposition 1). Finally, the set of Lagrange multipliers associated with the optimal primal solution is nonempty due to AS4 (cf. the proof of Lemma 1).

Now it is shown that the three conditions of [23, Theorem 1] hold for the problem at hand:

(i) The sets over which the saddle points are sought ($\mathcal{Y} \times \mathbb{R}^M \times \mathbb{R}_+^{3M}$) are closed and convex.

(ii) The set of saddle points of the Lagrangian is bounded. First it has to be shown that the set of optimal primal solutions is bounded; but this follows readily from Weierstrass' theorem (cf. the proof of Proposition 1). Moreover, the set of Lagrange multipliers associated with the optimal primal solution is bounded [19, Prop. 6.4.3], due to AS4.

(iii) For any $(\boldsymbol{y}, \boldsymbol{z}) \neq (\boldsymbol{y}^*, \boldsymbol{z}^*)$ it holds that $L(\boldsymbol{y}^*, \boldsymbol{z}^*, \boldsymbol{\nu}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*) < L(\boldsymbol{y}, \boldsymbol{z}, \boldsymbol{\nu}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$ (referred to in [23] as stability of the saddle points with respect to $(\boldsymbol{y}, \boldsymbol{z})$). This follows immediately from the strict convexity of $L(\boldsymbol{y}, \boldsymbol{z}, \boldsymbol{\nu}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$ in $(\boldsymbol{y}, \boldsymbol{z})$ over $\mathbb{R}^{2M}$ (cf. Lemma 1). □

B. Multi-channel networks

The proofs for this case are very similar to the single-channel case. Here only the points differentiating the arguments in the two cases are described.

Regarding Proposition 3, the proof is again by contradiction. The main argument must be made for every channel, hence the need for AS8. Moreover, note that $p^*_{i,f} > 0$ for all i and f due to AS5; hence, it is indeed possible to successively reduce the powers and arrive to a contradiction.

Now the first part of Lemma 2 can be shown again by manipulating the necessary optimality conditions $\partial L/\partial y_{i,f} \leq 0$, $\partial L/\partial z_{i,f} = 0$ and arriving to a linear system of the form (19) per channel. For the second part, note that the Hessian with respect to $\boldsymbol{y}$ and $\boldsymbol{z}$ is block diagonal, where each block corresponds to the variables organized per channel and has the form of (20). The proof then follows the proof of Lemma 1(ii); we apply again [22, Chapter 6, Theorem 2.3, Condition (J_{30})], where now the vector of all ones and length $2MF$ works.

Finally, Proposition 4 can be proved by invoking [23, Theorem 1] and using arguments similar to those in the proof of Proposition 2.

REFERENCES

[1] "Spectrum Policy Task Force," Report, ET Docket No. 02-135, Federal Communications Commission, Washington, D.C., Nov. 2002.

[2] Q. Zhao and A. Swami, "A survey of dynamic spectrum access: Signal processing and networking perspectives," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Honolulu, HI, Apr. 2007, pp. IV–1349–IV–1352.

[3] J. M. Peha, "Approaches to spectrum sharing," *IEEE Commun. Mag.*, vol. 43, no. 2, pp. 10–12, Feb. 2005.

[4] W. Lehr and J. Crowcroft, "Managing shared access to a spectrum commons," in *Proc. 1st IEEE Int. Symp. New Frontiers Dynamic Spectrum Access Networks*, Baltimore, MD, Nov. 2005, pp. 420–444.

[5] A. G. Marques, X. Wang, and G. B. Giannakis, "Optimal stochastic dual resource allocation for cognitive radios based on quantized CSI," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Las Vegas, NV, Mar.-Apr. 2008, pp. 2801–2804.

[6] C. Zou, T. Jin, C. Chigan, and Z. Tian, "QoS-aware distributed spectrum sharing for heterogeneous wireless cognitive networks," *Comput. Netw.*, vol. 52, no. 4, pp. 864–878, Mar. 2008.

[7] J. Huang, R. A. Berry, and M. L. Honig, "Distributed interference compensation for wireless networks," *IEEE J. Select. Areas Commun.*, vol. 24, no. 5, pp. 1074–1084, May 2006.

[8] S. Stańczak, M. Wiczanowski, and H. Boche, "Distributed utility-based power control: Objectives and algorithms," *IEEE Trans. Signal Process.*, vol. 55, no. 10, pp. 5058–5068, Oct. 2007.

[9] M. Wiczanowski, "Algorithmic and analytic framework for optimization of multi-user performance in wireless networks with interference," Ph.D. dissertation, Dept. Elect. Eng. Comput. Sci., Tech. Univ. Berlin, Berlin, Germany, 2007.

[10] M. Chiang, "Geometric programming for communication systems," *Foundations and Trends in Communications and Information Theory*, vol. 2, no. 1/2, 2006.

[11] X. Wang, A. G. Marques, and G. B. Giannakis, "Power-efficient resource allocation and quantization for TDMA using adaptive transmission and limited-rate feedback," *IEEE Trans. on Signal Process.*, vol. 56, no. 8, pp. 4470–4485, Sep. 2008.

[12] H. Boche and S. Stańczak, "Convexity of some feasible QoS regions and asymptotic behavior of the minimum total power in CDMA systems," *IEEE Trans. Commun.*, vol. 52, no. 12, pp. 2190–2197, Dec. 2004.

[13] S. Stańczak, M. Wiczanowski, and H. Boche, *Resource Allocation in Wireless Networks: Theory and Algorithms*. Berlin, Germany: Springer, 2006.

[14] N. Gatsis, A. G. Marques, and G. B. Giannakis, "Utility-based power control for peer-to-peer cognitive radio networks with heterogeneous QoS constraints," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Las Vegas, NV, Mar.-Apr. 2008, pp. 2805–2808.

[15] D. Palomar and M. Chiang, "A tutorial on decomposition methods for network utility maximization," *IEEE J. Select. Areas Commun.*, vol. 24, no. 8, pp. 1439–1451, Aug. 2006.

[16] R. D. Yates, "A framework for uplink power control in cellular radio systems," *IEEE J. Select. Areas Commun.*, vol. 13, no. 7, pp. 1341–1347, Sep. 1995.

[17] D. P. Bertsekas, *Nonlinear Programming*. Belmont, MA: Athena Scientific, 1999.

[18] A. V. Fiacco, *Introduction to Sensitivity and Stability Analysis in Nonlinear Programming*. New York, NY: Academic Press, 1983.

[19] D. P. Bertsekas, A. Nedić, and A. E. Ozdaglar, *Convex Analysis and Optimization*. Belmont, MA: Athena Scientific, 2003.

[20] R. A. Horn and C. R. Johnson, *Matrix Analysis*. New York, NY: Cambridge Univ. Press, 1985.

[21] S. A. Grandhi, J. Zander, and R. D. Yates, "Constrained power control," *Wireless Pers. Commun.*, vol. 1, no. 4, pp. 257–270, Dec. 1994.

[22] A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*. New York, NY: Academic Press, 1979.

[23] D. Maistrovskii, "Gradient methods for finding saddle points," *Matekon*, vol. 14, no. 1, pp. 3–22, 1977.

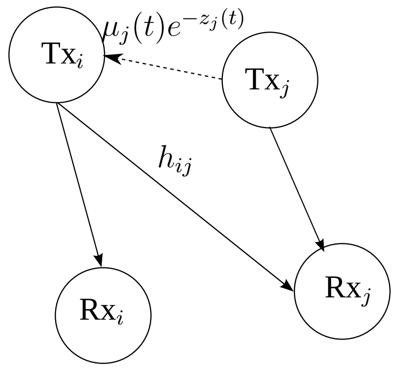

Fig. 1. Quantities involved in message passing.

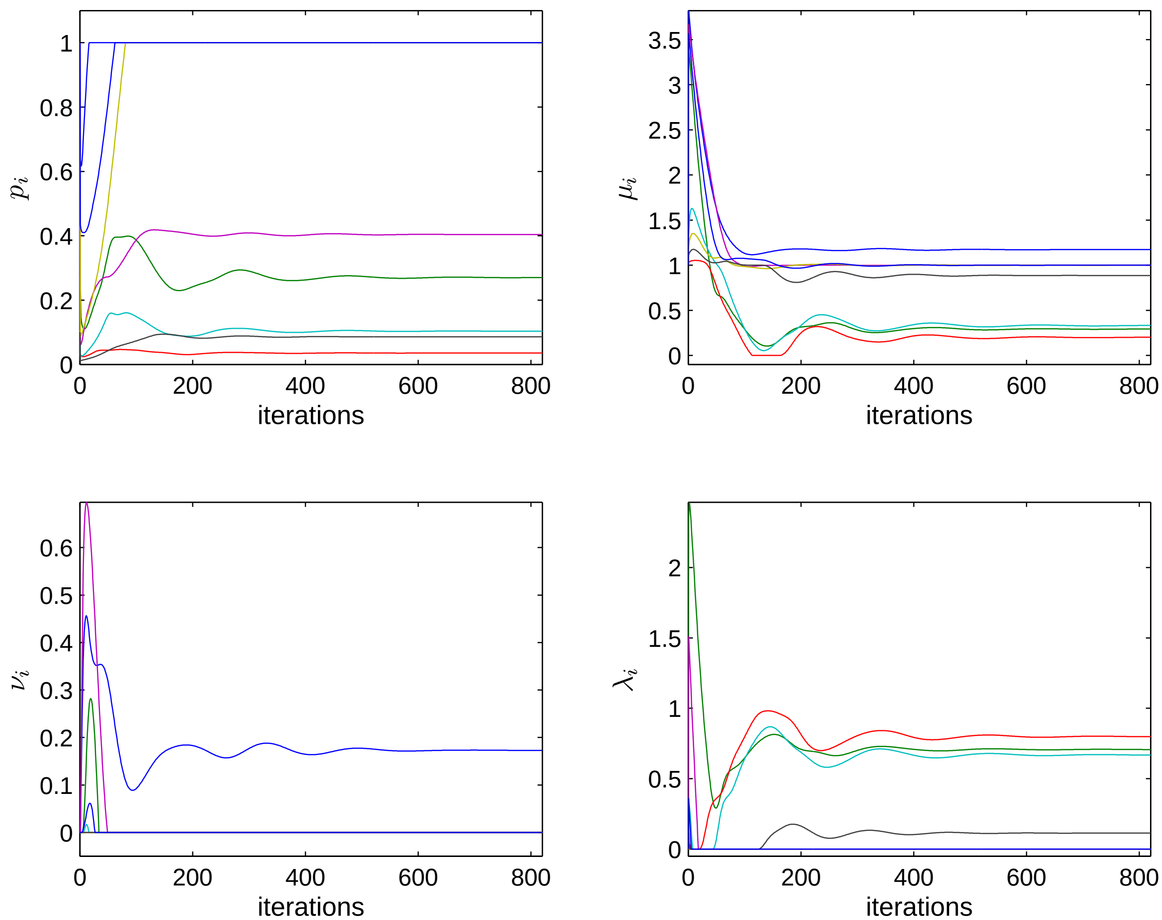

Fig. 2. Convergence of powers and Lagrange multipliers.

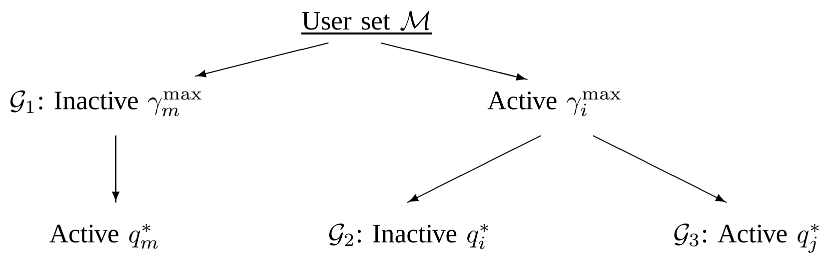

Fig. 3. Division of user set in proof of Proposition 1.

TABLE I
DIRECTIONAL DERIVATIVES OF SUM-UTILITY AS FUNCTION OF THE PERTURBATION.

$D_{\mathbf{e}_i} u_{\text{tot}}(\mathbf{0}) = \min\{\lambda_i \mid \exists\, \boldsymbol{\nu},\ \exists\, \boldsymbol{\mu},\ \exists\, \lambda_j,\ j \neq i \text{ s.t. } (\boldsymbol{\nu}, \boldsymbol{\lambda}, \boldsymbol{\mu}) \in \Theta^*\}$
$0 \leq D_{\mathbf{e}_i} u_{\text{tot}}(\mathbf{0}) \leq \lambda_i^*$
$D_{(-\mathbf{e}_i)} u_{\text{tot}}(\mathbf{0}) = -\max\{\lambda_i \mid \exists\, \boldsymbol{\nu},\ \exists\, \boldsymbol{\mu},\ \exists\, \lambda_j,\ j \neq i \text{ s.t. } (\boldsymbol{\nu}, \boldsymbol{\lambda}, \boldsymbol{\mu}) \in \Theta^*\}$
$-u_i(\gamma_i^{\max})\gamma_i^{\max} \leq D_{(-\mathbf{e}_i)} u_{\text{tot}}(\mathbf{0}) \leq -\lambda_i^*$

TABLE II
SIMULATION PARAMETERS FOR TEST CASE 1.

$M = 8,\ B = 128,\ \beta = 0.1$
$u_i = \ln(\gamma_i)\ \forall i$
$p_i^{\max} = 1\,\text{W},\ p_i^{\max}/n_i = 40\,\text{dB}\ \forall i$
Initialization: $z_i = \ln n_i,\ \lambda_i = 0,\ \nu_i = 0,\ \mu_i = 1\ \forall\, i \in \mathcal{M}$
$\gamma_i^{\min} = 140, \gamma_i^{\max} = 20000,\ \ i \in \{1, 6\}$
$\gamma_i^{\min} = 8, \gamma_i^{\max} = 20,\ \ i \in \{2, 3, 4\}$
$\gamma_i^{\min} = 20, \gamma_i^{\max} = 140,\ \ i \in \{5, 7, 8\}$

TABLE III
COORDINATES OF 8 TX-RX PAIRS (SHOWN IN 2 COLUMNS). TX ARE DEPLOYED OVER A SQUARE AREA OF SIDE 10M. EACH RX IS LOCATED BETWEEN 1 AND 3 METERS AWAY FROM ITS PEER TRANSMITTER. POSITIONS ARE RANDOMLY SELECTED.

Tx$_i$; Rx$_i$ $(i = 1, 2, 3, 4)$	Tx$_i$; Rx$_i$ $(i = 5, 6, 7, 8)$
(4.80,5.15);(4.92,3.67)	(6.17,3.18);(6.95,4.40)
(5.61,6.06);(6.11,7.51)	(6.85,5.88);(8.07,6.70)
(6.16,9.67);(4.70,10.93)	(5.10,1.30);(4.45,0.12)
(6.62,8.22);(5.17,9.39)	(7.14,2.54);(5.83,1.05)

TABLE IV
UNCONSTRAINED OPTIMIZATION IN SINGLE-CHANNEL NETWORKS: SUM-UTILITY (TOP) AND SINR PER USER (BOTTOM).

	Lagrangian	ADP[a]	Gradient projection alg.[b]
$\sum_i u_i$	33.676	33.676	33.676
γ_1	81.16	81.07	81.03
γ_2	43.35	43.34	43.34
γ_3	191.03	191.08	191.09
γ_4	6.24	6.24	6.24
γ_5	55.22	55.28	55.30
γ_6	443.06	443.00	443.00
γ_7	542.09	546.16	547.54
γ_8	7.59	7.53	7.51

All algorithms initialized randomly within the power constraints.
[a] $p_i^{\max}/p_i^{\min} = 40$ dB; all prices initialized randomly in $(0, 1/(n_i B))$.
[b] Stepsize = 0.2.

TABLE V
OPTIMIZATION WITH DIVERSE QOS CONSTRAINTS IN SINGLE-CHANNEL NETWORKS: SUM-UTILITY (TOP) AND SINR PER USER (BOTTOM).

	Lagrangian	QoS-ps-DSA[a]	QoSe-DSA[a]	Standard power control alg.
$\sum_i u_i$	32.4	23.6	23.6	
γ_1	140.0	**0.0137911**	**0.0137911**	70.4
γ_2	20.0	20.0	20.0	20.0
γ_3	20.0	20.0	20.0	20.0
γ_4	20.0	20.0	20.0	20.0
γ_5	32.9	52.5.0	52.5	81.4
γ_6	786.1	655.3	655.3	734.2
γ_7	140.0	140.0	140.0	140.0
γ_8	30.0	32.2	32.2	24.1

[a] $p_i^{\max}/p_i^{\min} = 40\text{dB}$; all powers initialized at $p_i^{\max}$; all prices initialized at 10^{-4}. Powers took (continuous) values in $[p_i^{\min}, p_i^{\max}]$.

TABLE VI
SIMULATION PARAMETERS FOR TEST CASE 2.

$M = 8,\ F = 16,\ \beta = 0.025$
$u_i(\gamma_{i,f}) = U_{i,f}(\gamma_{i,f}) = V_{i,f}(\gamma_{i,f}) = \ln \gamma_{i,f} \quad \forall\, i \in \mathcal{M},\ f \in \mathcal{F}$
$p_i^{\max}/n_{i,f} = 40\ \text{dB} \quad \forall\, i \in \mathcal{M},\ f \in \mathcal{F}$
Initialization: $y_{i,f} = \ln(p_i^{\max}/M),\ z_{i,f} = \ln n_{i,f},$ $\lambda_i = 0,\ \nu_i = 0,\ \mu_{i,f} = 1 \quad \forall\, i \in \mathcal{M},\ f \in \mathcal{F}$
Projection onto $\mathcal{Y}_i$ via MATLAB's `fmincon`
$U_i^{\min} = -50,\ i \in \{1,5,6,7,8\}$
$U_i^{\min} = -40,\ i \in \{2,3\}$
$U_i^{\min} = -30,\ i = 4$
$V_i^{\max} = 50,\ i \in \{1,2,3,4,5,6\}$
$V_i^{\max} = 10,\ i \in \{7,8\}$

TABLE VII
UNCONSTRAINED OPTIMIZATION IN MULTI-CHANNEL NETWORKS: SUM-UTILITY (TOP) AND INDIVIDUAL UTILITIES PER USER (BOTTOM).

	Lagrangian	DADP[a]
$\sum_{i,f} u_i^f$	-149.96	-149.96
1	-21.09	-21.11
2	-52.99	-52.99
3	-8.12	-8.12
4	-38.05	-38.05
5	-6.12	-6.10
6	9.10	9.10
7	18.26	18.35
8	-50.95	-51.05

[a] $p_i^{\max}/p_i^{\min} = 40$ dB; stepsize = 0.05; 30 inner iterations per dual. All powers initialized randomly in $(p_i^{\min}, p_i^{\max})$ and all prices in $(0, 1/n_{i,f})$.

TABLE VIII
OPTIMIZATION WITH DIVERSE QOS CONSTRAINTS IN MULTI-CHANNEL NETWORKS: SUM-UTILITY (TOP) AND INDIVIDUAL UTILITIES PER USER (BOTTOM).

	Lagrangian	MC-QoS-ps-DSA[a]	MC-QoSe-DSA[a]
$\sum_{i,f} u_i^f$	-162.38	-317.50	-688.46
1	-16.90	**-68.85**	**-82.67**
2	-40.00	**-114.50**	**-102.75**
3	-38.13	-6.69	**-104.85**
4	-30.00	**-48.81**	**-89.13**
5	-8.76	5.63	**-63.15**
6	4.23	-3.43	**-79.98**
7	8.20	10.14	**-70.13**
8	-41.03	**-91.14**	**-95.80**

[a] $p_i^{\max}/p_i^{\min} = 40$ dB; all powers initialized randomly in $(p_i^{\min}, p_i^{\max})$ and all prices in $(0, 1/n_{i,f})$. Powers took (continuous) values so that $p_i^{\min} \le \sum_f p_{i,f} \le p_i^{\max}$ for all users. Projection onto power constraints via MATLAB's `fmincon`.