Cognitive Radio Network Duality and Algorithms for Utility Maximization

Liang Zheng and Chee Wei Tan, *Senior Member, IEEE*

***Abstract*—We study a utility maximization framework for spectrum sharing among cognitive secondary users and licensed primary users in cognitive radio networks. All the users maximize the network utility by adapting their signal-to-interference-plus-noise ratio (SINR) assignment and transmit power subject to power budget constraints and additional interference temperature constraint for the secondary users. The utility maximization problem is challenging to solve optimally in a distributed manner due to the nonconvexity and the tight coupling between the power budget and interference temperature constraint sets. We first study a special case where egalitarian SINR fairness is the utility, and a tuning-free distributed algorithm with a geometric convergence rate is developed to solve it optimally. Then, we answer the general utility maximization question by developing a cognitive radio network duality to decouple the SINR assignment, the transmit power and the interference temperature allocation. This leads to a utility maximization algorithm that leverages the egalitarian fairness power control as a submodule to maintain a desirable separability in the SINR assignment between the secondary and primary users. This algorithm has the advantage that it can be distributively implemented, and the method converges relatively fast. Numerical results are presented to show that our proposed algorithms are theoretically sound and practically implementable.**

***Index Terms*—Optimization, network utility maximization, cognitive radio networks, spectrum allocation.**

I. Introduction

WE CONSIDER wireless cognitive radio networks with multiple users transmitting simultaneously on a shared spectrum, where the link transmission qualities are significantly influenced by multiuser interference. This interference, if uncontrolled or inadequately controlled, can degrade the utility of a user transmission and lead to adverse network operating points. In a cognitive radio network environment, the users are further governed by temperature constraints that must be satisfied in order to mitigate the interference coming from the cognitive secondary users from over-spilling into the licensed primary users. Thus, the total attainable utility of all the users depends on a joint Signal-to-Interference-plus-Noise Radio (SINR) assignment, the transmit power and the tolerable interference level (interference temperature), which in other words rely on the resource allocation policy. An optimal resource allocation can maximize the overall network utility without the cognitive secondary users causing overwhelming interference to the primary users. Secondary users should have a minimal impact on the operation of the primary users [1], which can be controlled by imposing an interference temperature constraint on the received interference at the primary users.

Given a set of power budget and interference temperature constraints, the SINR assignment of all users must be jointly coordinated, but there are two major hurdles that need to be overcome. First, algorithms that adapt the transmit power and interference temperature based on allocated SINR targets assume that the SINR targets are within the feasibility region, which however requires a centralized admission controller. Second, the algorithms have to be decentralized, practical to deploy and be fast enough with minimal or, preferably, no parameter tuning. This is especially important since secondary users can arrive and depart in a dynamic setting, and so resources have to be adapted fast enough to converge to a new optimal operating point whenever the network conditions change. This however is challenging due to the tight coupling between primary and secondary users in the SINR assignment. It is desirable that the resource allocation for primary and secondary users be distributed with minimal overhead.

There are several works that partially address these two challenges in the literature. The authors in [2] study the SINR assignment in the context of wireless cellular networks by a re-parameterization of the feasible SINR region, and propose a load and spillage algorithm that jointly updates the SINR assignment and power. This algorithm however confines the optimality by considering a reduced feasible SINR region. The authors in [3] propose algorithms to coordinate the secondary users by sensing a feedback signal from the primary users' interference temperature condition to reduce outage, but do not address the joint optimal SINR assignment and power allocation. The authors in [4] propose a cognitive radio admission control and scheduling policy, which however does not handle the interference temperature constraint. To coordinate the interference temperature by controlling the SINR assignment, the authors in [5] propose a belief propagation framework by wireless scheduling to solve a nonconvex utility maximization problem. Utility maximization algorithms using fractional frequency reuse have been proposed in [6], [7] for wireless cellular networks. The authors in [8] consider a joint SINR assignment and power control to maximize the worst-case SINR with an interference temperature constraint.

To overcome these two challenges, we propose and analyze distributed algorithms for jointly optimal SINR assignment,

Manuscript received 12 April 2012; revised 31 August 2012. The work in this paper was partially supported by grants from the Research Grants Council of Hong Kong Project No. RGC CityU 125212, Qualcomm Inc. and the Science, Technology and Innovation Commission of Shenzhen Municipality, Project No. JCYJ20120829161727318 on Green Communications in Small-cell Mobile Networks.

L. Zheng and C. W. Tan are with the College of Science and Engineering, City University of Hong Kong, Tat Chee Ave., Hong Kong (e-mail: {lianzheng2,cheewtan}@cityu.edu.hk).

Digital Object Identifier 10.1109/JSAC.2013.130315.

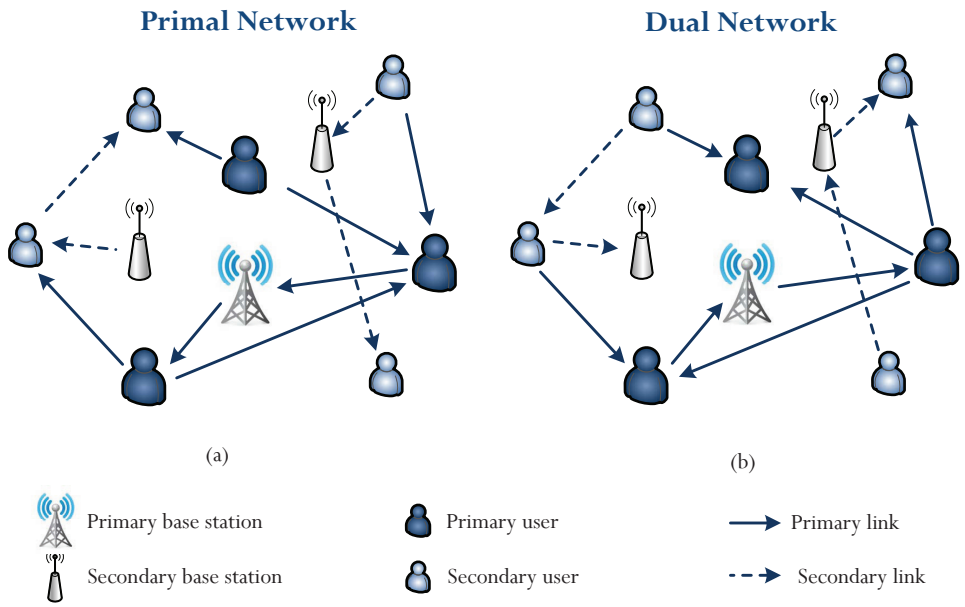


Fig. 1. An illustration of the cognitive radio network duality. It shows that the link directions are reversed in the primal network and the dual network. (a) The primal network. (b) The dual network.

power and interference temperature control. Our algorithms are based on a novel *cognitive radio network duality* that decouples SINR assignment, power and interference temperature. Network duality is a fundamental concept in multiuser communication [9], [10]. Network dualities for wireless cellular and ad-hoc networks were investigated in [11], [12] for a total power minimization problem (a convex problem) using linear programming duality. A network duality was later developed for a max-min weighted SINR problem (a nonconvex problem) in [13]–[15] using nonnegative matrix theory and geometric programming duality. These network dualities (including the one in this paper) assert that two different networks, respectively a primal network and a dual network, can be construed to attain an identical SINR performance. In particular, the dual network reverses the link directions in the primal network, as illustrated in Figure 1. Hence, a feasible SINR for one is also feasible for the other. Further, the power and interference temperature in the primal and the dual networks are jointly optimized to solve the utility maximization problem formulated for the primal network subject to the power budget and interference temperature constraints.

The main contributions of this paper are summarized as follows.

- We present an equivalent reformulation of the utility maximization problems as an optimization problem involving spectral radius constraint sets. We then characterize the global optimality by expressing the optimal solution analytically and the feasible SINR region. A special case that maximizes the egalitarian fairness of all the SINRs as the utility is solved optimally using a tuning-free geometrically fast convergent algorithm.
- We develop the cognitive radio network duality and characterize analytically the power and interference temperature of the primal and dual networks. The relationship between them and the gradients of the spectral radius constraints in the utility maximization problem is established.
- The utility maximization problem is solved using an optimization technique that can be interpreted as iteratively minimizing the *interference load* in the network. In particular, the egalitarian fairness SINR problem is adapted iteratively to solve the general utility maximization problem, which can further be solved distributively by leveraging the cognitive radio network duality.
- Our algorithms can be practically implemented in today's wireless networks (3GPP systems) as they reuse a power control submodule already widely implemented. Numerical evaluations show that our algorithms have good performance, often yielding the optimal solution in tens of iterations even for a large number of users.

The rest of this paper is organized as follows. We present the system model in Section II. In Section III, we reformulate our utility maximization problem in the SINR domain with spectral radius constraints. Then, using the nonlinear Perron-Frobenius theory, we first solve a special case of the utility maximization problem, the weighted max-min SINR, in Section IV. Next, using nonnegative matrix theory, we present the cognitive radio network duality that we use to design a distributed algorithm to solve the utility maximization problem. We evaluate the performance of our algorithms numerically in Section V. Finally, we conclude the paper in Section VI. All the proofs can be found in the Appendix.

We refer the readers to Figure 2 for an overview of the connection between the three key optimization problems in the paper. The following notation is used in our paper. Column vectors and matrices are denoted by boldfaced lowercase and uppercase respectively. Let $\rho(\mathbf{F})$ denote the Perron-Frobenius

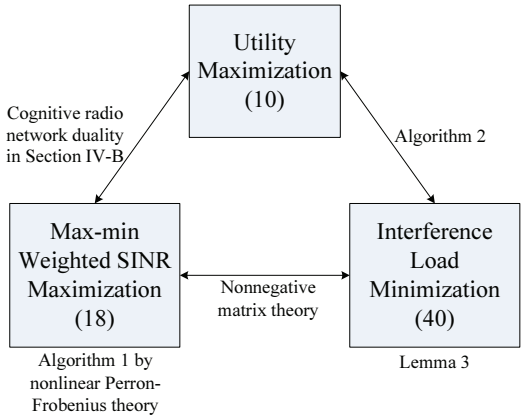


Fig. 2. Overview of the connection between the three optimization problems studied in this paper.

eigenvalue of an irreducible nonnegative matrix $\mathbf{F}$, and $\mathbf{x}(\mathbf{F})$ and $\mathbf{y}(\mathbf{F})$ denote the Perron right and left eigenvectors of $\mathbf{F}$ associated with $\rho(\mathbf{F})$. From the Perron-Frobenius Theorem (cf. [16]), the Perron-Frobenius eigenvalue is the largest positive eigenvalue of an irreducible nonnegative matrix, and the corresponding Perron right and left eigenvectors are both entry-wise positive. We let $\mathbf{e}_l$ denote the lth unit coordinate vector and $\mathbf{I}$ denote the identity matrix. The super-scripts $(\cdot)^\top$ denotes transpose. We denote $\mathbf{x} \circ \mathbf{y}$ as a Schur product of $\mathbf{x}$ and $\mathbf{y}$, i.e., $\mathbf{x} \circ \mathbf{y} = [x_1 y_1, \ldots, x_L y_L]^\top$, and denote $\mathbf{x}/\mathbf{y}$ as the component-wise division between $\mathbf{x}$ and $\mathbf{y}$, i.e., $\mathbf{x}/\mathbf{y} = [x_1/y_1, \ldots, x_L/y_L]^\top$. Let $\mathbf{1} = [1, \ldots, 1]^\top \in \mathbb{R}^L$. For a vector $\mathbf{x} = [x_1, \ldots, x_L]^\top$, $\mathrm{diag}(\mathbf{x})$ is its diagonal matrix $\mathrm{diag}(x_1, \ldots, x_L)$. Let $e^{\mathbf{x}}$ denote $e^{\mathbf{x}} = (e^{x_1}, \ldots, e^{x_L})^\top$, and $\log \mathbf{x}$ denote $\log \mathbf{x} = (\log x_1, \ldots, \log x_L)$. An equality involving eigenvectors is true up to a scaling constant.

II. System Model

We consider a cognitive radio network with a collection of primary users and secondary users, which we call the primal network. Assume that there are L secondary users (transmitter/receiver pairs) communicating simultaneously over a common frequency-flat fading channel. As illustrated in Figure 1(a), let $\mathbf{G} = [G_{lk}]_{l,k=1}^{L} > 0_{L \times L}$ represent the channel gain, where G_{lk} is the channel gain from the kth transmitter to the lth receiver, and $\mathbf{n} = [n_1, \ldots, n_L]^\top > \mathbf{0}$, where n_l is the noise power at the lth user. The vector $\mathbf{p} = [p_1, \ldots, p_L]^\top$ is the transmit power vector. The Signal-to-Interference-and-Noise Ratio (SINR) for the lth receiver is defined as the ratio of the received signal power to the sum of interference signal power and additive noise power. Now, the SINR of the lth user in the primal network can be given in terms of $\mathbf{p}$:

$$\mathsf{SINR}_l^P(\mathbf{p}) = \frac{p_l G_{ll}}{\sum_{k \neq l} p_k G_{lk} + n_l}. \tag{1}$$

We also define a nonnegative matrix $\mathbf{F}$ with entries:

$$F_{lk} = \begin{cases} 0, & \text{if } l = k \\ G_{lk}, & \text{if } l \neq k \end{cases} \tag{2}$$

and the vector

$$\mathbf{v} = \left(\frac{1}{G_{11}}, \frac{1}{G_{22}}, \ldots, \frac{1}{G_{LL}} \right)^\top. \tag{3}$$

Moreover, we assume that $\mathbf{F}$ is irreducible, i.e., each link has at least an interferer. For brevity, we denote vector $\boldsymbol{\gamma}$ as the SINR for all users. In this paper, we use the equivalent form of the SINR as:

$$\gamma_l = \frac{p_l}{\big(\mathrm{diag}(\mathbf{v})(\mathbf{F}\mathbf{p} + \mathbf{n})\big)_l}, \tag{4}$$

where we use $\mathbf{F}$ and $\mathbf{v}$ in (2) and (3) respectively.

Next, let $\mathbf{q}$ denote the vector containing the normalized total interference and noise for each user given by:

$$q_l = \sum_{k=1}^{L} p_k \frac{F_{lk}}{G_{ll}} + \frac{n_l}{G_{ll}}. \tag{5}$$

In this paper, for ease of presentation, we call $\mathbf{q}$ the interference temperature instead of the total interference plus noise. With the above notation, we have $\gamma_l = p_l/q_l$. Since $\mathbf{p} = \mathrm{diag}(\boldsymbol{\gamma})\mathbf{q}$ and $\mathbf{q} = \mathrm{diag}(\mathbf{v})(\mathbf{F}\mathbf{p} + \mathbf{n})$, the power and interference temperature can be written, respectively, as

$$\mathbf{p} = \mathrm{diag}(\boldsymbol{\gamma} \circ \mathbf{v})(\mathbf{F}\mathbf{p} + \mathbf{n}) \tag{6}$$

and

$$\mathbf{q} = \mathrm{diag}(\mathbf{v})\big(\mathbf{F}\,\mathrm{diag}(\boldsymbol{\gamma})\mathbf{q} + \mathbf{n}\big). \tag{7}$$

Furthermore, from (6) and (7), we can obtain a one-to-one mapping between $\boldsymbol{\gamma}$ and $\mathbf{p}$ and also one between $\boldsymbol{\gamma}$ and $\mathbf{q}$, given by:

$$\mathbf{p}(\boldsymbol{\gamma}) = \big(\mathbf{I} - \mathrm{diag}(\boldsymbol{\gamma} \circ \mathbf{v})\mathbf{F}\big)^{-1} \mathrm{diag}(\boldsymbol{\gamma} \circ \mathbf{v})\mathbf{n} \tag{8}$$

and

$$\mathbf{q}(\boldsymbol{\gamma}) = \big(\mathbf{I} - \mathrm{diag}(\mathbf{v})\mathbf{F}\,\mathrm{diag}(\boldsymbol{\gamma})\big)^{-1} \mathrm{diag}(\mathbf{v})\mathbf{n}, \tag{9}$$

respectively. Hence, given a feasible SINR assignment, a feasible power and interference temperature can be computed using (8) and (9) respectively.

III. Utility Maximization

In the following, we study a cognitive radio network utility maximization problem. Let us denote $U(\boldsymbol{\gamma})$ as the network utility, which is a network-wide quality-of-service (QoS) measure: for example, the max-min SINR fairness utility and the α-fair utility [2], [17]. By appropriately choosing the utility objective function subject to a set of resource constraints (upper bounds on the power and interference temperature), we maximize the network utility for both primary and secondary users. In a cognitive radio network, secondary users are allowed access to the network and share the spectrum with primary users. This can lead to significant interference to the primary users. Thus, it is important to avoid overwhelming interference from secondary users, i.e., constrain each primary user's interference temperature below a tolerable upper bound. Thus, the optimal utility has to be achieved subject to an additional interference management in the network [18]. In particular, the users are subject to a set of weighted power constraints and individual interference temperature constraints.

We consider the optimization of the utility in a feasible SINR domain subject to both the available power budget and the interference temperature constraints. Let us denote the upper bounds for the power budget and interference temperature constraints as $\bar{\mathbf{p}}$ and $\bar{\mathbf{q}}$ respectively. The general utility maximization problem is then given as:

$$\begin{array}{ll} \text{maximize} & U(\boldsymbol{\gamma}) \\ \text{subject to} & \mathbf{w}_l{}^{\top}\mathbf{p} \le \bar{p}_l, \quad l = 1, \ldots, L, \\ & \mathbf{e}_l{}^{\top}\mathbf{q} \le \bar{q}_l, \quad l = 1, \ldots, L, \\ & \mathbf{q} = \operatorname{diag}(\mathbf{v})(\mathbf{F}\mathbf{p} + \mathbf{n}), \\ & \gamma_l = p_l/q_l, \quad l = 1, \ldots, L, \\ \text{variables:} & \boldsymbol{\gamma}, \mathbf{p}, \mathbf{q}, \end{array} \tag{10}$$

where $\mathbf{w}_l$ is a nonnegative vector to model the lth power budget constraint. Obviously, we must have $\bar{q}_l > n_l v_l$ for all l for feasibility. The constraint set in (10) is general enough to model the primary and the secondary users, i.e., each index l can represent either a primary user or a secondary user. For example, for a primary user, we can set $\bar{q}_l$ to some finite value. For a secondary user, we can set $\bar{q}_l$ to infinity. Also, as special cases, when $\mathbf{w}_l = \mathbf{e}_l$, we have the individual power constraints $p_l \le \bar{p}_l$ for all l that model a cellular uplink system, and when $\mathbf{w}_l = \mathbf{1}$, we have a total power constraint that models a cellular downlink system. Since $\mathbf{w}_l$ can be chosen as any positive vector associated with the power allocation, the nonnegative vector $\mathbf{w}_l$ is general enough to model a wide range of practical power budget constraints. However, adding the interference temperature constraint to (10) imposes further limitations that make the problem in (10) more challenging to solve.

In general, (10) is nonconvex (due to the nonconvex constraint set) and thus may be difficult to solve. Denote the optimal $\mathbf{p}$ and $\mathbf{q}$ in (10) as $\mathbf{p}^\star$ and $\mathbf{q}^\star$. Next, we reformulate (10) as an equivalent optimization problem in the SINR domain with a set of spectral radius constraints that only have $\boldsymbol{\gamma}$ as the only variable. This is given by the following result.

Theorem 1: The optimal value in (10) is equal to the optimal value of the following problem:

$$\begin{array}{ll} \text{maximize} & U(\boldsymbol{\gamma}) \\ \text{subject to} & \rho\big(\operatorname{diag}(\boldsymbol{\gamma} \circ \mathbf{v})(\mathbf{F} + (1/\bar{p}_l)\mathbf{n}\mathbf{w}_l^{\top})\big) \le 1, \\ & \qquad\qquad l = 1, \ldots, L, \\ & \rho\big(\operatorname{diag}(\mathbf{v})(\mathbf{F}\operatorname{diag}(\boldsymbol{\gamma}) + (1/\bar{q}_l)\mathbf{n}\mathbf{e}_l^{\top})\big) \le 1, \\ & \qquad\qquad l = 1, \ldots, L, \\ \text{variables:} & \boldsymbol{\gamma}. \end{array} \tag{11}$$

Furthermore, $\boldsymbol{\gamma}^\star$ is an optimal solution to (11) if and only if $\mathbf{p}(\boldsymbol{\gamma}^\star)$ in (8) and $\mathbf{q}(\boldsymbol{\gamma}^\star)$ in (9) is an optimal solution to (10), i.e., $\mathbf{p}(\boldsymbol{\gamma}^\star) = \mathbf{p}^\star$ and $\mathbf{q}(\boldsymbol{\gamma}^\star) = \mathbf{q}^\star$. In addition, we have

$$\rho\big(\operatorname{diag}(\boldsymbol{\gamma}^\star \circ \mathbf{v})(\mathbf{F} + (1/\bar{p}_i)\mathbf{n}\mathbf{w}_i^{\top})\big) = 1, \tag{12}$$

where

$$i = \arg \max_{l=1,\ldots,L} \rho\big(\operatorname{diag}(\boldsymbol{\gamma}^\star \circ \mathbf{v})(\mathbf{F} + (1/\bar{p}_l)\mathbf{n}\mathbf{w}_l^{\top})\big) \tag{13}$$

or

$$\rho\big(\operatorname{diag}(\mathbf{v})(\mathbf{F}\operatorname{diag}(\boldsymbol{\gamma}^\star) + (1/\bar{q}_j)\mathbf{n}\mathbf{e}_j^{\top})\big) = 1, \tag{14}$$

where

$$j = \arg \max_{l=1,\ldots,L} \rho\big(\operatorname{diag}(\mathbf{v})(\mathbf{F}\operatorname{diag}(\boldsymbol{\gamma}^\star) + (1/\bar{q}_l)\mathbf{n}\mathbf{e}_l^{\top})\big). \tag{15}$$

Remark 1: Note that (12) and (14) imply that the optimal solution of (10) is such that $\mathbf{w}_i^{\top}\mathbf{p}^\star = \bar{p}_i$ for some i, or $q_j^\star = \bar{q}_j$ for some j respectively in (10) (cf. proof of Theorem 1). Moreover, the power $\mathbf{p}^\star$ and interference temperature $\mathbf{q}^\star$ are the Perron right eigenvectors of the nonnegative matrices in (12) and (14) respectively.

Theorem 1 indicates that the mappings that relate the SINR domain with the power and interference temperature domain in (8) and (9) respectively can be used to compute the optimal $\mathbf{p}^\star$ and $\mathbf{q}^\star$ respectively by using the optimal $\boldsymbol{\gamma}^\star$ in (11), but, more interestingly, it reveals that the optimal power and interference temperature can be regarded as eigenvectors of suitably constructed matrices that are functions of $\boldsymbol{\gamma}^\star$, which govern the resource allocation behavior.

Furthermore, using a logarithmic mapping of variable, (11) can be transformed to an optimization problem with a convex constraint set. For $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_L)^{\top} > \mathbf{0}$, let

$$\tilde{\gamma}_l = \log \gamma_l, \quad l = 1, \ldots, L,$$

i.e., $\boldsymbol{\gamma} = e^{\tilde{\boldsymbol{\gamma}}}$. Then, (11) is equivalent to:

$$\begin{array}{ll} \text{maximize} & U(e^{\tilde{\boldsymbol{\gamma}}}) \\ \text{subject to} & \log \rho(\operatorname{diag}(e^{\tilde{\boldsymbol{\gamma}}} \circ \mathbf{v})(\mathbf{F} + (1/\bar{p}_l)\mathbf{n}\mathbf{w}_l^{\top})) \le 0, \\ & \qquad\qquad l = 1, \ldots, L, \\ & \log \rho\big(\operatorname{diag}(\mathbf{v})(\mathbf{F}\operatorname{diag}(e^{\tilde{\boldsymbol{\gamma}}}) + (1/\bar{q}_l)\mathbf{n}\mathbf{e}_l^{\top})\big) \le 0, \\ & \qquad\qquad l = 1, \ldots, L, \\ \text{variables:} & \tilde{\boldsymbol{\gamma}}. \end{array} \tag{16}$$

Let us denote the optimal solution of (16) by $\tilde{\boldsymbol{\gamma}}^\star$. Note that $\tilde{\boldsymbol{\gamma}}^\star = \log \boldsymbol{\gamma}^\star$ for all l.

Remark 2: For an irreducible nonnegative matrix $\mathbf{F} \in \mathbb{R}_+^{L \times L}$, $\log \rho(\operatorname{diag}(e^{\tilde{\boldsymbol{\gamma}}})\mathbf{F})$ is a convex function [19], due to the log-convexity property of the Perron-Frobenius eigenvalue [20]. Therefore, the constraint set in (16) is convex.

In general, depending on $U(\boldsymbol{\gamma})$, (16) is nonconvex in $\tilde{\boldsymbol{\gamma}}$. However, we will study in the following interesting special cases of (16), equivalently (10), that can be solved optimally in a distributed manner. This is achieved by the cognitive radio network duality developed in the following.

IV. Cognitive Radio Network Duality and Algorithms

In this section, we consider (16) for both smooth and non-smooth utility functions. In particular, in Section IV-B, a max-min weighted SINR problem (for egalitarian SINR fairness) will be first solved using a nonlinear Perron-Frobenius theory.[1] This is then used together with the cognitive radio network duality in Section IV-C to solve (16). It will be shown that the transmit power and interference temperature can be analytically expressed as the Perron right eigenvectors of the specially constructed matrices associated with the spectral radius constraints in (16). This leads to the development of a *cognitive radio network duality* involving a dual network as illustrated in Figure 1(b), which enables a distributed algorithm to maximize the primal network utility.

A. Smooth and Non-smooth Utility

The reformulation introduced in Section III allows us to decompose the utility maximization problem in (10) into first optimizing $\boldsymbol{\gamma}$, i.e., optimizing the SINR assignment, and then optimizing the power $\mathbf{p}$ and interference temperature $\mathbf{q}$. In this section, we discuss the assumption for the objective function, which will also be useful in our proposed distributed algorithm

[1]There are several extensions to the classical (linear) Perron-Frobenius theorem in nonnegative matrix theory for classes of nonlinear maps. We use the finite dimensional nonlinear Perron-Frobenius theory developed in [21].

in Section IV-E to solve (10) for both smooth and non-smooth utility functions.

Assumption 1: The utility function $U(\boldsymbol{\gamma})$ is concave in $\log \gamma_l$ for all l.

For example, α-fairness utility [17] satisfies Assumption 1, given by:

$$U(\boldsymbol{\gamma}) = \begin{cases} \sum_{l=1}^{L} \log \gamma_l, & \text{if } \alpha = 1, \\ \sum_{l=1}^{L} (1-\alpha)^{-1} \gamma_l^{1-\alpha}, & \text{if } \alpha > 1. \end{cases}$$

Note that (16) includes the sum rate maximization problem studied in [22], [23], when $U(\boldsymbol{\gamma}) = \sum_{l=1}^{L} \log(1+\gamma_l)$, but this objective function does not satisfy Assumption 1, and henceforth it is a nonconvex problem that requires global optimization techniques, e.g., those studied in [22], [23].

If $U(\boldsymbol{\gamma})$ is differentiable and separable, i.e., $U(\boldsymbol{\gamma}) = \sum_{l=1}^{L} U_l(\gamma_l)$, let $\nabla U_l(\gamma_l) = \partial U_l / \partial \gamma_l$ and $\nabla^2 U_l(\gamma_l) = \partial^2 U_l / \partial^2 \gamma_l$ denote the first-order and second-order derivative of $U_l(\gamma_l)$ with respect to γ_l respectively. Then, $U_l(\gamma_l)$ is concave in $\log \gamma_l$ if and only if the curvature is sufficiently large [2]:

$$\nabla^2 U_l(\gamma_l) \le -\frac{\nabla U_l(\gamma_l)}{\gamma_l}. \tag{17}$$

In the general case (when $U(\boldsymbol{\gamma})$ can be non-smooth), we consider the subgradient of $U(\boldsymbol{\gamma})$, whose definition is given as follows.

Definition 1 (cf. [24]): The subgradient $\mathbf{g} \in \mathbb{R}^L$ of $U(\boldsymbol{\gamma})$ at $\hat{\boldsymbol{\gamma}}$ satisfies

$$U(\boldsymbol{\gamma}) \le U(\hat{\boldsymbol{\gamma}}) + \mathbf{g}^\top (\boldsymbol{\gamma} - \hat{\boldsymbol{\gamma}})$$

for any feasible $\boldsymbol{\gamma}$. If $U(\boldsymbol{\gamma})$ is concave and differentiable at $\boldsymbol{\gamma}$, the subgradient is unique and is given by its gradient $\mathbf{g} = \nabla U(\boldsymbol{\gamma})$. On the other hand, if $U(\boldsymbol{\gamma})$ is concave but not differentiable, its subgradient is in the set:

$$\bigcap \{\mathbf{g} \mid U(\boldsymbol{\gamma}) \le U(\hat{\boldsymbol{\gamma}}) + \mathbf{g}^\top (\boldsymbol{\gamma} - \hat{\boldsymbol{\gamma}})\}$$

for any feasible $\boldsymbol{\gamma}$.

B. Non-smooth Special Case: $U(\boldsymbol{\gamma}) = \min_{l=1,\ldots,L} \frac{\gamma_l}{\beta_l}$

In this section, let us consider the max-min weighted SINR problem (for egalitarian SINR fairness), which is a special case of (11) that has a *non-smooth* concave objective function:

$$\begin{array}{lll} \text{maximize} & \min\limits_{l=1,\ldots,L} \dfrac{\gamma_l}{\beta_l} & \\ \text{subject to} & \rho\big(\operatorname{diag}(\boldsymbol{\gamma} \circ \mathbf{v})(\mathbf{F} + (1/\bar{p}_l)\mathbf{n}\mathbf{w}_l^\top)\big) \le 1, & \\ & & l = 1, \ldots, L, \\ & \rho\big(\operatorname{diag}(\mathbf{v})(\mathbf{F}\operatorname{diag}(\boldsymbol{\gamma}) + (1/\bar{q}_l)\mathbf{n}\mathbf{e}_l^\top)\big) \le 1, & \\ & & l = 1, \ldots, L, \\ \text{variables:} & \boldsymbol{\gamma}, & \end{array} \tag{18}$$

where $\boldsymbol{\beta}$ is a positive vector with the entry β_l used to reflect a priority of the lth link. A larger β_l indicates a higher priority.

Let us define the following set of nonnegative matrices:

$$\mathbf{B}_l = \operatorname{diag}(\mathbf{v})(\mathbf{F} + \frac{1}{\bar{p}_l}\mathbf{n}\mathbf{w}_l^\top), \quad l = 1, \ldots, L, \tag{19}$$

$$\mathbf{D}_l = \Big(\mathbf{I} + \frac{1}{\bar{q}_l - n_l v_l} \operatorname{diag}(\mathbf{v})\mathbf{n}\mathbf{e}_l^\top\Big)\operatorname{diag}(\mathbf{v})\mathbf{F}, \quad l = 1, \ldots, L. \tag{20}$$

By applying the nonnegative matrix theory and the nonlinear Perron-Frobenius theory, we obtain a closed form solution to (18) as well as the corresponding optimal solution in (10), which is unique.

Lemma 1: The optimal value of (18) is given by

$$\frac{1}{\max\limits_{l=1,\ldots,L}\{\rho(\operatorname{diag}(\boldsymbol{\beta})\mathbf{B}_l), \rho(\mathbf{D}_l \operatorname{diag}(\boldsymbol{\beta}))\}}. \tag{21}$$

The optimal solution to (18) $\boldsymbol{\gamma}^\star$ is a vector with $\gamma_l^\star/\beta_l$ equal to a common value

$$\gamma_m^\star/\beta_m = 1/\max_{l=1,\ldots,L}\{\rho(\operatorname{diag}(\boldsymbol{\beta})\mathbf{B}_l), \rho(\mathbf{D}_l \operatorname{diag}(\boldsymbol{\beta}))\}$$

for all l, where

$$m = \arg\max_{l=1,\ldots,L} \big\{\rho\big(\operatorname{diag}(\boldsymbol{\beta})\mathbf{B}_l\big), \rho\big(\mathbf{D}_l \operatorname{diag}(\boldsymbol{\beta})\big)\big\}. \tag{22}$$

If the optimal value is $1/\rho(\operatorname{diag}(\boldsymbol{\beta})\mathbf{B}_m)$ for m in (22), the optimal power and interference temperature in (10) is, respectively, given by

$$\mathbf{p}^\star = \frac{\bar{p}_m}{\mathbf{w}_m^\top \mathbf{x}\big(\operatorname{diag}(\boldsymbol{\beta})\mathbf{B}_m\big)}\mathbf{x}\big(\operatorname{diag}(\boldsymbol{\beta})\mathbf{B}_m\big) \tag{23}$$

and

$$\mathbf{q}^\star = \operatorname{diag}(\boldsymbol{\gamma}^\star)^{-1}\mathbf{p}^\star,$$

and if the optimal value is $1/\rho(\mathbf{D}_m \operatorname{diag}(\boldsymbol{\beta}))$ for m in (22), the optimal interference temperature and power in (10) is, respectively, given by

$$\mathbf{q}^\star = \frac{\bar{q}_m}{\mathbf{e}_m^\top \mathbf{x}\big(\mathbf{D}_m \operatorname{diag}(\boldsymbol{\beta})\big)}\mathbf{x}\big(\mathbf{D}_m \operatorname{diag}(\boldsymbol{\beta})\big) \tag{24}$$

and

$$\mathbf{p}^\star = \operatorname{diag}(\boldsymbol{\gamma}^\star)\mathbf{q}^\star.$$

We next give an intriguingly simple algorithm to compute the analytical solution in Lemma 1. In particular, by applying the nonlinear Perron-Frobenius theory in [21], the following algorithm computes $\mathbf{p}^\star$ given in Lemma 1.

Algorithm 1 (Max-min weighted SINR Algorithm):

Initialize $\mathbf{p}(0)$.

1) Each lth user updates its power $p_l(k+1)$ as follows:
$p_l(k+1) = \dfrac{\beta_l}{\mathsf{SINR}_l^P(\mathbf{p}(k))} p_l(k)$.
2) Normalize $\mathbf{p}(k+1)$:
$$\mathbf{p}(k+1) \leftarrow \frac{\mathbf{p}(k+1)}{\max\limits_{l=1,\ldots,L}\left\{\dfrac{\mathbf{w}_l^\top \mathbf{p}(k+1)}{\bar{p}_l}, \dfrac{\mathbf{e}_l^\top \operatorname{diag}(\mathbf{v})\mathbf{F}\mathbf{p}(k+1)}{\bar{q}_l - n_l v_l}\right\}}.$$

Theorem 2: Starting from any initial point $\mathbf{p}(0)$, $\mathbf{p}(k)$ converges geometrically fast to the power $\mathbf{p}^\star$ given in Lemma 1.

Remark 3: At Step 1, the Foschini-Miljanic power control algorithm update in [25] (a power control submodule widely implemented in 3GPP systems) is used. At Step 2, the computation of $\mathbf{w}_l^\top \mathbf{p}(k+1)$ and the normalization of $\mathbf{p}(k+1)$ can be computed by a gossip algorithm in a distributed manner [26]. Notice that $\mathbf{e}_l^\top \operatorname{diag}(\mathbf{v})(\mathbf{F}\mathbf{p}(k+1)+\mathbf{n})$ and n_l are respectively the interference and the noise power at lth receiver, which can be measured from the interference temperature. Thus, $\mathbf{e}_l^\top \operatorname{diag}(\mathbf{v})\mathbf{F}\mathbf{p}(k+1)$ can be locally obtained.

Interestingly, using the Friedland-Karlin inequalities in [23], [27], (18) is equivalent to (11) with a smooth objective

function given by

$$U(\boldsymbol{\gamma}) = \sum_{l=1}^{L} \big(\mathbf{x}(\boldsymbol{\Omega}) \circ \mathbf{y}(\boldsymbol{\Omega})\big)_l \log \gamma_l, \tag{25}$$

where $\boldsymbol{\Omega} \in \{\operatorname{diag}(\boldsymbol{\beta})\mathbf{B}_l, \mathbf{D}_l \operatorname{diag}(\boldsymbol{\beta})\}$ is the matrix defined for the mth user with m given in (22): If the optimal value for (18) is $1/\rho(\operatorname{diag}(\boldsymbol{\beta})\mathbf{B}_m)$ for m in (22), then $\boldsymbol{\Omega} = \operatorname{diag}(\boldsymbol{\beta})\mathbf{B}_m$; If the optimal value for (18) is $1/\rho(\mathbf{D}_m \operatorname{diag}(\boldsymbol{\beta}))$ for m in (22), then $\boldsymbol{\Omega} = \mathbf{D}_m \operatorname{diag}(\boldsymbol{\beta})$.

C. Cognitive Radio Network Duality

In this section, we develop a cognitive radio dual network that achieves an identical SINR performance as the primal network but with all the link directions reversed. The dual network is constructed using the same topology as the primal network but with the links reversed. Thus, we use $\mathbf{F}^\top$ as the channel fading matrix in the dual network (due to the reversed link directions). As shown in Figure 1(b), the dual network reverses the link direction of the primal network, i.e., the channel gain from the kth transmitter to the lth receiver is G_{kl} in the dual network. The primal and dual networks are used for the development of a *cognitive radio network duality* that is used for distributed algorithm design.

After establishing this network duality, we will turn to solve (16) for general utility functions that satisfy Assumption 1 using a projected subgradient method [24]. Interestingly, this method can be made distributed by connecting the gradients of the spectral radius functions in (16) with the power and the interference temperature in both the primal and the dual networks (cf. Figure 1). This is achieved by applying the cognitive radio network duality that exploits the structure of the spectral radius functions in (16).

Recall that we have already defined the nonnegative matrices $\mathbf{B}_l$ and $\mathbf{D}_l$ given in (19) and (20) respectively. The gradients $\mathbf{g} \in \mathbb{R}^L$ of $\log \rho(\operatorname{diag}(\boldsymbol{\gamma})\mathbf{B}_l)$ and $\log \rho(\mathbf{D}_l \operatorname{diag}(\boldsymbol{\gamma}))$ at $\boldsymbol{\gamma}$ are given respectively by [23], [27]:

$$\mathbf{g} = \mathbf{x}\big(\operatorname{diag}(\boldsymbol{\gamma})\mathbf{B}_l\big) \circ \mathbf{y}\big(\operatorname{diag}(\boldsymbol{\gamma})\mathbf{B}_l\big) \tag{26}$$

and

$$\mathbf{g} = \mathbf{x}\big(\mathbf{D}_l \operatorname{diag}(\boldsymbol{\gamma})\big) \circ \mathbf{y}\big(\mathbf{D}_l \operatorname{diag}(\boldsymbol{\gamma})\big), \tag{27}$$

normalized such that $\mathbf{1}^\top \mathbf{g} = 1$. We already know from the reformulation in Section III that the Perron right eigenvectors of the matrices in (12) and (14) are the optimal transmit power $\mathbf{p}^\star$ and the optimal interference temperature $\mathbf{q}^\star$ of the primal network respectively, which also appear in (26) for $l = i$ in (13) and (27) for $l = j$ in (15) respectively. Observe that the gradient for the spectral radius function is the Schur product of the Perron right and left eigenvectors (recall that the Schur product is the component-wise product of two vectors). This interesting characterization leads naturally to the development of a dual network, where a physical interpretation can be given to the Perron left eigenvectors of the matrices in (12) and (14). This interesting fact enables a distributed method to compute $\mathbf{g}$ in (26) and (27).

Definition 2: Let the power $\mathbf{s}$ and the interference temperature $\mathbf{t}$ in the dual network be given respectively by

$$\mathbf{s} = \operatorname{diag}(\boldsymbol{\gamma})\mathbf{t} \tag{28}$$

and

$$\mathbf{t} = \operatorname{diag}(\mathbf{v})(\mathbf{F}^\top \mathbf{s} + \mathbf{w}_\iota), \tag{29}$$

where the weight vector $\mathbf{w}_\iota$ for the transmit power constraint in the primal network is assumed to be a virtual received noise in the dual network, and the index ι corresponds to either i in (13) or j in (15), which indicates the power or interference temperature constraint that is tight at optimality of (16). Now, the SINR of the lth user in the dual network can be given in terms of $\mathbf{s}$:

$$\mathsf{SINR}_l^D(\mathbf{s}) = \frac{s_l}{\big(\operatorname{diag}(\mathbf{v})(\mathbf{F}^\top \mathbf{s} + \mathbf{w}_\iota)\big)_l}. \tag{30}$$

Observe that the channel matrix in (1) is replaced by its transpose in (30). However, this dual network must achieve all possible SINR values that are feasible in the primal network. Combining (28) and (29), we have

$$\mathbf{s} = \operatorname{diag}(\boldsymbol{\gamma} \circ \mathbf{v})(\mathbf{F}^\top \mathbf{s} + \mathbf{w}_\iota) \tag{31}$$

and

$$\mathbf{t} = \operatorname{diag}(\mathbf{v})\big(\mathbf{F}^\top \operatorname{diag}(\boldsymbol{\gamma})\mathbf{t} + \mathbf{w}_\iota\big). \tag{32}$$

Similar to the derivation of (8) and (9), we can also write (31) and (32), respectively, as functions of $\boldsymbol{\gamma}$:

$$\mathbf{s}(\boldsymbol{\gamma}) = (\mathbf{I} - \operatorname{diag}(\boldsymbol{\gamma} \circ \mathbf{v})\mathbf{F}^\top)^{-1} \operatorname{diag}(\boldsymbol{\gamma} \circ \mathbf{v})\mathbf{w}_\iota \tag{33}$$

and

$$\mathbf{t}(\boldsymbol{\gamma}) = (\mathbf{I} - \operatorname{diag}(\mathbf{v})\mathbf{F}^\top \operatorname{diag}(\boldsymbol{\gamma}))^{-1} \operatorname{diag}(\mathbf{v})\mathbf{w}_\iota. \tag{34}$$

Now, (33) and (34) can be used as one-to-one mappings between $\boldsymbol{\gamma}^\star$ and $\mathbf{s}^\star$ as well as between $\boldsymbol{\gamma}^\star$ and $\mathbf{t}^\star$ respectively.

Corresponding to the primal network power and interference temperature constraints, the dual network power and interference temperature constraints can be given respectively by (note the use of ι, where ι can be i in (13) or ι can be j in (15)):

$$\mathbf{n}^\top \mathbf{s} \le \bar{p}_\iota,$$

and

$$\mathbf{n}^\top \mathbf{t} \le \bar{q}_\iota.$$

Moreover, it is still true that either the dual network power constraint or the dual network interference temperature constraint is tight at optimality. We then have the following result that connects the transmit powers and interference temperatures of both the primal and dual networks, which is used in a distributed algorithm design in Section IV-E.

Lemma 2: The Perron right and left eigenvector of the nonnegative matrix $\operatorname{diag}(\boldsymbol{\gamma}^\star)\mathbf{B}_i$, where $\mathbf{B}_i$ is given in (19) with i in (13), satisfy

$$\mathbf{x}\big(\operatorname{diag}(\boldsymbol{\gamma}^\star)\mathbf{B}_i\big) = \mathbf{p}^\star \tag{35}$$

and

$$\mathbf{y}\big(\operatorname{diag}(\boldsymbol{\gamma}^\star)\mathbf{B}_i\big) = \operatorname{diag}(\boldsymbol{\gamma}^\star \circ \mathbf{v})^{-1}\mathbf{s}^\star, \tag{36}$$

respectively. The Perron right and left eigenvector of the nonnegative matrix $\mathbf{D}_j \operatorname{diag}(\boldsymbol{\gamma}^\star)$, where $\mathbf{D}_j$ is given in (20) with j in (15), satisfy, respectively,

$$\mathbf{x}\big(\mathbf{D}_j \operatorname{diag}(\boldsymbol{\gamma}^\star)\big) = \mathbf{q}^\star \tag{37}$$

and

$$\mathbf{y}\big(\mathbf{D}_j \operatorname{diag}(\boldsymbol{\gamma}^\star)\big) = \operatorname{diag}(\boldsymbol{\gamma}^\star/\mathbf{v})\mathbf{t}^\star. \tag{38}$$

Finally, Figure 3 summarizes the cognitive radio network duality that characterizes the transmit power and interference temperature in the primal and the dual networks as the Perron right and left eigenvectors of appropriately constructed nonnegative matrices.

Cognitive Radio Network Duality

	Primal network		Dual network
Power Budget Constraint	$\mathbf{p}^\star = \mathbf{x}\big(\operatorname{diag}(\boldsymbol{\gamma}^\star \circ \mathbf{v})(\mathbf{F} + (1/\bar{p}_i)\mathbf{n}\mathbf{w}_i^\top)\big)$ $= \big(\mathbf{I} - \operatorname{diag}(\boldsymbol{\gamma}^\star \circ \mathbf{v})\mathbf{F}\big)^{-1} \operatorname{diag}(\boldsymbol{\gamma}^\star \circ \mathbf{v})\mathbf{n}$ $\mathbf{w}_i^\top \mathbf{p}^\star \le \bar{p}_i$	$\leftrightarrow$	$\mathbf{s}^\star = \operatorname{diag}(\boldsymbol{\gamma}^\star \circ \mathbf{v})\mathbf{y}\big(\operatorname{diag}(\boldsymbol{\gamma}^\star \circ \mathbf{v})(\mathbf{F} + (1/\bar{p}_i)\mathbf{n}\mathbf{w}_i^\top)\big)$ $= \big(\mathbf{I} - \operatorname{diag}(\boldsymbol{\gamma}^\star \circ \mathbf{v})\mathbf{F}^\top\big)^{-1} \operatorname{diag}(\boldsymbol{\gamma}^\star \circ \mathbf{v})\mathbf{w}_i$ $\mathbf{n}^\top \mathbf{s}^\star \le \bar{p}_i$
Interference Temperature Constraint	$\mathbf{q}^\star = \mathbf{x}\big((\mathbf{I} + \frac{1}{\bar{q}_j - n_j v_j} \operatorname{diag}(\mathbf{v})\mathbf{n}\mathbf{e}_j^\top) \operatorname{diag}(\mathbf{v})\mathbf{F} \operatorname{diag}(\boldsymbol{\gamma}^\star)\big)$ $= \big(\mathbf{I} - \operatorname{diag}(\mathbf{v})\mathbf{F} \operatorname{diag}(\boldsymbol{\gamma}^\star)\big)^{-1} \operatorname{diag}(\mathbf{v})\mathbf{n}$ $\mathbf{e}_j^\top \mathbf{q}^\star \le \bar{q}_j$	$\leftrightarrow$	$\mathbf{t}^\star = \operatorname{diag}(\boldsymbol{\gamma}^\star/\mathbf{v})^{-1}\mathbf{y}\big((\mathbf{I} + \frac{1}{\bar{q}_j - n_j v_j} \operatorname{diag}(\mathbf{v})\mathbf{n}\mathbf{e}_j^\top) \operatorname{diag}(\mathbf{v})\mathbf{F} \operatorname{diag}(\boldsymbol{\gamma}^\star)\big)$ $= \big(\mathbf{I} - \operatorname{diag}(\mathbf{v})\mathbf{F}^\top \operatorname{diag}(\boldsymbol{\gamma}^\star)\big)^{-1} \operatorname{diag}(\mathbf{v})\mathbf{w}_j$ $\mathbf{n}^\top \mathbf{t}^\star \le \bar{q}_j$

Fig. 3. The cognitive radio network duality illustrates the connection between the primal and the dual networks in terms of both the Perron right and left eigenvectors of the nonnegative matrices associated with the spectral radius constraints in (16).

D. Interference Load Minimization

In Section III, the constraints of the utility maximization problem in (10) can be succinctly reformulated as spectral radius constraints, i.e., $\rho\big(\operatorname{diag}(\boldsymbol{\gamma})\mathbf{B}_l\big) \le 1 \ \ \forall l$ and $\rho\big(\mathbf{D}_l \operatorname{diag}(\boldsymbol{\gamma})\big) \le 1 \ \ \forall l$, where $\mathbf{B}_l$ and $\mathbf{D}_l$ are given in (19) and (20) respectively. These spectral radius functions capture the effect of interference on the feasibility of SINR assignment. The spectral radius thus plays the role of a useful measure for interference, which we call the *interference load*. This means that a smaller $\rho\big(\operatorname{diag}(\boldsymbol{\gamma})\mathbf{B}_l\big)$ or $\rho\big(\mathbf{D}_l \operatorname{diag}(\boldsymbol{\gamma})\big)$ indicates a smaller interference load on the network, which leads to a larger feasible SINR region that can be optimized. On the other hand, the interference load increases with interference and therefore reduces the feasible SINR region. This connection between the interference load and our utility maximization problem in Section III will be made precise in the following. By leveraging both the cognitive radio network duality in Section IV-C and the *interference load minimization* problem (to be introduced below), a distributed algorithm is then proposed to solve the utility maximization problem in (11).

First, let us consider the following convex optimization problem given by:

$$\begin{array}{ll} \text{maximize} & \boldsymbol{\alpha}^\top \tilde{\boldsymbol{\gamma}} \\ \text{subject to} & \rho\big(\operatorname{diag}(e^{\tilde{\boldsymbol{\gamma}}})\mathbf{B}_l\big) \le 1, \quad l = 1, \ldots, L, \\ & \rho\big(\mathbf{D}_l \operatorname{diag}(e^{\tilde{\boldsymbol{\gamma}}})\big) \le 1, \quad l = 1, \ldots, L, \\ \text{variables:} & \tilde{\boldsymbol{\gamma}}, \end{array} \tag{39}$$

where $\mathbf{B}_l$ and $\mathbf{D}_l$ are given in (19) and (20) respectively, $\tilde{\boldsymbol{\gamma}}$ is the logarithmic mapping by $\tilde{\boldsymbol{\gamma}} = \log \boldsymbol{\gamma}$, and $\boldsymbol{\alpha} \in \mathbb{R}_+^L$ is a given probability vector that is used to approximate $U(e^{\tilde{\boldsymbol{\gamma}}})$ using its Taylor series approximation up to the first order terms (cf. proof of Theorem 3).

Below, it is fruitful to consider the *interference load minimization* problem that is intimately related to (39) and instead minimizes a spectral radius function subject to a single linear constraint:

$$\begin{array}{ll} \text{minimize} & \max\limits_{l=1,\ldots,L} \{\rho(\operatorname{diag}(e^{\tilde{\boldsymbol{\eta}}})\mathbf{B}_l), \rho(\mathbf{D}_l \operatorname{diag}(e^{\tilde{\boldsymbol{\eta}}}))\} \\ \text{subject to} & \boldsymbol{\alpha}^\top \tilde{\boldsymbol{\eta}} \ge 0, \\ \text{variables:} & \tilde{\boldsymbol{\eta}}, \end{array} \tag{40}$$

where $\tilde{\boldsymbol{\eta}}$ is the logarithmic transformation: $\tilde{\boldsymbol{\eta}} = \log \boldsymbol{\eta}$. The following result connects (39) and (40).

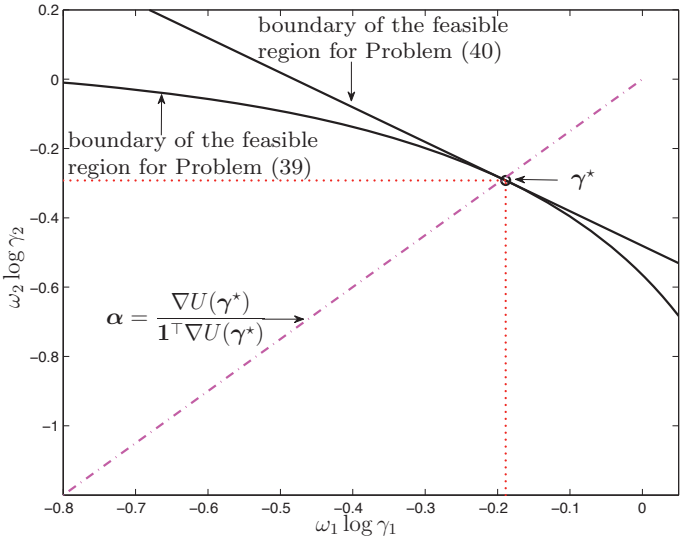


Fig. 4. Illustration of the connection between (39) and (40). Achievable region for a 2-user example with objective function $\sum_l \omega_l \log \gamma_l$. The channel gains are given by $G_{11} = 0.69$, $G_{12} = 0.12$, $G_{21} = 0.13$, $G_{22} = 0.70$ and the weight is $\omega = [0.40,\ 0.60]^\top$. The maximum power and interference temperature for users are $\bar{\mathbf{p}} = [1.50,\ 1.00]^\top$ W and $\bar{\mathbf{q}} = [2.50,\ 3.00]^\top$ W respectively. The noise powers for both users are 1 W. The intersection point of the direction $\boldsymbol{\alpha}$ and achievable region is the optimal solution. Moreover, the minimization of (40) also intersects with the optimal solution of (39) at the boundary of the feasible region.

Lemma 3: Let $\boldsymbol{\gamma}^\star$ and $\boldsymbol{\eta}^\star$ be the optimal solution of (39) and (40) respectively, and let $\xi^\star$ and $\zeta^\star$ be the optimal value of (39) and (40) respectively. Then, $\boldsymbol{\gamma}^\star$ and $\boldsymbol{\eta}^\star$ satisfy

$$\boldsymbol{\gamma}^\star = \frac{1}{\zeta^\star} \boldsymbol{\eta}^\star. \tag{41}$$

Furthermore, since $\boldsymbol{\alpha}$ is a probability vector, $\xi^\star$ and $\zeta^\star$ satisfy

$$\xi^\star = -\log \zeta^\star.$$

The formulation of (40) that minimizes the interference load thus provides a connection (by choosing $\boldsymbol{\alpha}$ to be proportional to the subgradient of the utility function) between the general utility maximization problem in (10) and its special case of egalitarian SINR fairness optimization in (18). An interesting interpretation of Lemma 3 is that the optimal SINR in the general utility maximization can be scaled relative to the optimal SINR achieved under the egalitarian SINR fairness. Figure 4 illustrates the geometric interpretation of the connection between (39) and (40) by an example using

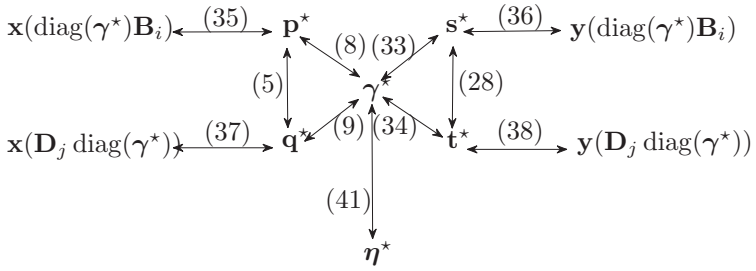


Fig. 5. A summary of the transformation and mapping between the optimal primal power $\mathbf{p}^\star$, the optimal primal interference temperature $\mathbf{q}^\star$, the optimal dual power $\mathbf{s}^\star$, the optimal dual interference temperature $\mathbf{t}^\star$, the optimal SINR assignment $\boldsymbol{\gamma}^\star$ and the optimal SINR scaling $\boldsymbol{\eta}^\star$. The index i and j are given in (13) and (15) respectively.

$U(\boldsymbol{\gamma}) = \sum_{l=1}^{L} \omega_l \log \gamma_l$ with a positive $\boldsymbol{\omega}$. The relationship between the various transformation and mappings of the optimization variables are shown in Figure 5.

E. Utility Maximization Algorithm

In (40), we have shown that it is possible to find a scaling factor that connects the SINR assignment in (39) (which is related to (16) through a Taylor's series first order approximation) with the egalitarian SINR fairness. This means that (39) can be first solved by considering (40) to find the scaling factor, and then to scale the optimal solution of (40) to finally obtain the solution of (16).

We use the projected subgradient method to solve (40). The parameter $\boldsymbol{\alpha}$ in (40) is updated iteratively. In particular, at the $(k+1)$th iteration, we update $\boldsymbol{\alpha}(k)$ as the subgradient of $U(e^{\tilde{\boldsymbol{\gamma}}})$ at $\tilde{\boldsymbol{\gamma}}(k)$. Thus, instead of solving (16) directly, we replace the objective function of (16) in a neighborhood of a feasible point $\tilde{\boldsymbol{\gamma}}(k)$ by its Taylor series approximation (cf. proof of Theorem 3), which is a successive convex approximation technique. Meanwhile, the algorithm iterates according to the subgradient of both the objective function and the constraint functions. However, computing the gradient of the spectral radius functions, i.e., the Schur product of the Perron right and left eigenvectors, requires centralized computation in general.

However, by exploiting the cognitive radio network duality, this task can be made distributed. Making use of the results in Section IV-C, we can then obtain a distributed algorithm to solve (16). Observe that the gradient $\mathbf{g} \in \mathbb{R}^L$ of $\rho\big(\mathrm{diag}(\boldsymbol{\eta})\mathbf{B}_l\big)$ and $\rho\big(\mathbf{D}_l\,\mathrm{diag}(\boldsymbol{\eta})\big)$ in terms of $\mathbf{p}$, $\mathbf{q}$, $\mathbf{s}$ and $\mathbf{t}$ are, respectively, given by (cf. Lemma 2):

$$\mathbf{g} = \mathbf{p} \circ \big(\mathrm{diag}(\boldsymbol{\eta} \circ \mathbf{v})^{-1}\mathbf{s}\big) \tag{42}$$

and

$$\mathbf{g} = \mathbf{q} \circ \big(\mathrm{diag}(\boldsymbol{\eta}/\mathbf{v})\mathbf{t}\big), \tag{43}$$

normalized such that $\mathbf{1}^\top\mathbf{g} = 1$. Furthermore, (42) and (43) can be rewritten, respectively, as:

$$g_l = p_l\Big(\frac{1}{\eta_l v_l}\Big)s_l \tag{44}$$

and

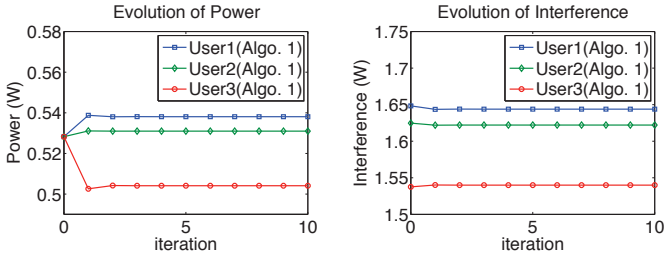
$$g_l = q_l\Big(\frac{\eta_l}{v_l}\Big)t_l. \tag{45}$$

Now, in (44) and (45), the respective variables p_l, q_l, s_l, t_l and η_l can be locally obtained, thus making the gradient computation distributed. We next use (44) and (45) to obtain a distributed algorithm based on the projected subgradient method to solve (10).

Algorithm 2 (Utility Maximization Algorithm):

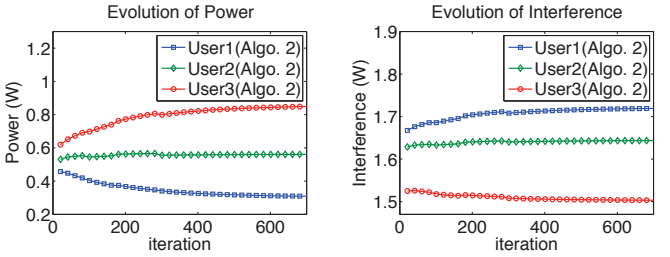
Initialize $\boldsymbol{\eta}(0)$, set the step size $\nu(0) \in (0,1)$.

1) Compute the weight $\boldsymbol{\alpha}(k)$:
   if $U(\boldsymbol{\eta})$ is smooth,
   $$\boldsymbol{\alpha}(k) = \frac{\nabla U(\boldsymbol{\eta}(k))}{\mathbf{1}^\top \nabla U(\boldsymbol{\eta}(k))},$$
   else
   $$\boldsymbol{\alpha}(k) = \frac{\hat{\mathbf{g}}}{\mathbf{1}^\top\hat{\mathbf{g}}}, \text{ where } \hat{\mathbf{g}} \text{ satisfies}$$
   $U(\boldsymbol{\eta}) \le U(\boldsymbol{\eta}(k)) + \hat{\mathbf{g}}^\top(\boldsymbol{\eta} - \boldsymbol{\eta}(k))$ for any feasible $\boldsymbol{\eta}$.
   end if
2) In the primal network, set the power and interference temperature output of Algorithm 1 with $\boldsymbol{\beta} = \boldsymbol{\eta}(k)$, which upon its convergence solves the primal network optimization problem:
$$\begin{array}{lll} \text{maximize} & \min\limits_{l=1,\ldots,L} \dfrac{\mathsf{SINR}_l^P(\mathbf{p})}{\eta_l(k)} & \\ \text{subject to} & \mathbf{w}_l^\top\mathbf{p} \le \bar{p}_l, & l = 1,\ldots,L, \\ & \mathbf{e}_l^\top\mathbf{q} \le \bar{q}_l, & l = 1,\ldots,L, \\ & \mathbf{q} = \mathrm{diag}(\mathbf{v})(\mathbf{F}\mathbf{p} + \mathbf{n}), & \\ \text{variables:} & \mathbf{p}, \mathbf{q} & \end{array} \tag{46}$$
   as $\mathbf{p}(k)$ and $\mathbf{q}(k)$ respectively.
   The computation of this step also provides in addition the value of ι_k, i.e., the i_kth power or j_kth interference temperature constraint that is tight in (46), where $i_k = \arg\max\limits_{l=1,\ldots,L} \rho\big(\mathrm{diag}(\boldsymbol{\eta}(k) \circ \mathbf{v})(\mathbf{F} + (1/\bar{p}_l)\mathbf{n}\mathbf{w}_l^\top)\big)$ and $j_k = \arg\max\limits_{l=1,\ldots,L} \rho\big(\mathrm{diag}(\mathbf{v})(\mathbf{F}\,\mathrm{diag}(\boldsymbol{\eta}(k)) + (1/\bar{q}_l)\mathbf{n}\mathbf{e}_l^\top)\big)$.
   In the dual network, set the power and interference temperature output of Algorithm 1 with $\boldsymbol{\beta} = \boldsymbol{\eta}(k)$ and with $\mathsf{SINR}^P(\mathbf{p})$ replaced by $\mathsf{SINR}^D(\mathbf{s})$ at Algorithm 1's Step 1 and the normalization at Step 2 of Algorithm 1 $\max\limits_{l=1,\ldots,L}\{\mathbf{w}_l^\top\mathbf{p}(k+1)/\bar{p}_l, \mathbf{e}_l^\top\,\mathrm{diag}(\mathbf{v})\mathbf{F}\mathbf{p}(k+1)/(\bar{q}_l - n_l v_l)\}$ replaced by $\max\{\mathbf{n}^\top\mathbf{s}(k+1)/\bar{p}_{\iota_k}, \mathbf{n}^\top\,\mathrm{diag}(\mathbf{v})\mathbf{F}^\top\mathbf{s}(k+1)/(\bar{q}_{\iota_k} - \mathbf{n}^\top\,\mathrm{diag}(\mathbf{v})\mathbf{w}_{\iota_k})\}$, which upon its convergence solves the dual network optimization problem:
$$\begin{array}{lll} \text{maximize} & \min\limits_{l=1,\ldots,L} \dfrac{\mathsf{SINR}_l^D(\mathbf{s})}{\eta_l(k)} & \\ \text{subject to} & \mathbf{n}^\top\mathbf{s} \le \bar{p}_{\iota_k}, & \\ & \mathbf{n}^\top\mathbf{t} \le \bar{q}_{\iota_k}, & \\ & \mathbf{t} = \mathrm{diag}(\mathbf{v})(\mathbf{F}^\top\mathbf{s} + \mathbf{w}_{\iota_k}), & \\ \text{variables:} & \mathbf{s}, \mathbf{t} & \end{array} \tag{47}$$



(a) (b)

Fig. 6. Illustration of the convergence of Algorithm 1 with randomly chosen initial $\mathbf{p}(0)$. We plot the trajectory of power for each of the three users in (a), and plot the trajectory of interference temperature for each of the three users in (b). We can observe from the figures that the convergence of Algorithm 1 is geometrically fast.



(a) (b)

Fig. 7. Illustration of the convergence of Algorithm 2 with randomly chosen initial $\boldsymbol{\eta}(0)$. Algorithm 2 is a distributed algorithm. $\mathbf{p}(k)$ and $\mathbf{q}(k)$ are obtained from Step 2, which terminates when $\epsilon = 10^{-13}$. The number of iterations in the figure corresponds to the outer loop (slower time-scale). We plot the trajectory of power for each of the three users in (a), and plot the trajectory of interference temperature for each of the three users in (b).

as $\mathbf{s}(k)$ and $\mathbf{t}(k)$ respectively.

3) Each lth user updates its gradient $g_l(k)$ as follows:
   if $\boldsymbol{\alpha}(k)^\top \log \boldsymbol{\eta}(k) < 0$
   $g_l(k) = \alpha_l(k)$,
   else
   if $\iota_k = i_k$
   $g_l(k) = p_l(k)s_l(k)/(\eta_l(k)v_l), \quad l = 1, \ldots, L,$
   if $\iota_k = j_k$
   $g_l(k) = q_l(k)t_l(k)\eta_l(k)/v_l, \quad l = 1, \ldots, L.$
   end if
4) Update $\boldsymbol{\eta}(k+1)$:
   $\boldsymbol{\eta}(k+1) \leftarrow \boldsymbol{\eta}(k) e^{-\nu(k)(\mathbf{g}(k)/\mathbf{1}^\top \mathbf{g}(k))}$.
5) Update $\nu(k+1)$ according to Theorem 3 below and go to Step 1.

Theorem 3: Starting from any initial point $\boldsymbol{\eta}(0)$, if the step size $\nu(k)$ satisfies $\sum_{k=0}^{\infty} \nu(k) = \infty$, $\sum_{k=0}^{\infty} (\nu(k))^2 < \infty$, then $\mathbf{p}(k)$ and $\mathbf{q}(k)$ in Algorithm 2 converge to the optimal solution $\mathbf{p}^\star$ and $\mathbf{q}^\star$ of (10) respectively.

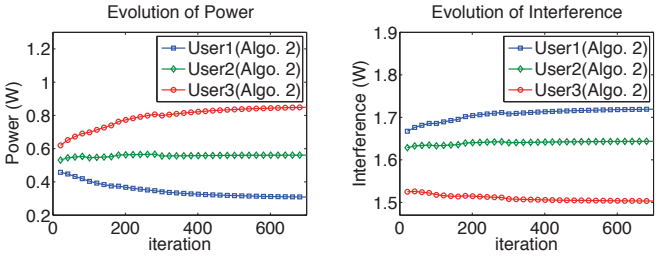
Furthermore, if a constant step size is used in Step 4, Algorithm 2 is guaranteed to converge to a neighborhood of the optimal solution.

Remark 4: If $U(\boldsymbol{\gamma})$ is smooth, Algorithm 2 solves (40) with $\boldsymbol{\alpha}$ given by $\frac{\nabla U(\boldsymbol{\gamma}^\star)}{\mathbf{1}^\top \nabla U(\boldsymbol{\gamma}^\star)}$, where $\boldsymbol{\gamma}^\star$ is the optimal solution to (10) (cf. Figure 4).

Remark 5: Since we run Algorithm 1 at each iteration of Algorithm 2 as an inner loop, Algorithm 2 is a two time-scale algorithm. At Step 2, we obtain the primal network power $\mathbf{p}(k)$ and the primal network interference temperature $\mathbf{q}(k)$ from the output of Algorithm 1 by using the input weight parameter $\boldsymbol{\eta}(k)$. Similarly, we can also obtain the dual network power $\mathbf{s}(k)$ and the dual network interference temperature $\mathbf{t}(k)$ in the same way. This means that $\mathbf{p}(k)$, $\mathbf{q}(k)$, $\mathbf{s}(k)$, and $\mathbf{t}(k)$ are the optimal solutions of Algorithm 1 for a given $\boldsymbol{\eta}(k)$. The computation of $\mathbf{1}^\top \nabla U(\boldsymbol{\eta}(k))$, $\boldsymbol{\alpha}(k)^\top \log \boldsymbol{\eta}(k)$ and $\mathbf{1}^\top \mathbf{g}(k)$ can be made distributed by a gossip algorithm [26]. Practical stopping criterions for solving (46) and (47) can be $|p_l(k+1) - p_l(k)| \leq \epsilon$ and $|s_l(k+1) - s_l(k)| \leq \epsilon$ respectively for a given small ϵ.

V. Numerical Examples

In this section, we provide numerical examples to illustrate the performance of Algorithm 1 in Section IV-B and Algorithm 2 in Section IV-E, illustrating the convergence properties



(a) (b)

Fig. 8. An illustration of Algorithm 2. All the other parameters are the same as Figure 7 except that $\epsilon = 10^{-3}$.

of Algorithms 1 and 2. Recall that an optimization problem having two different utility functions can be constructed to yield the same optimal solution for egalitarian SINR fairness (cf. (25) in Section IV-B). By setting the weight of (25) appropriately, and applying Algorithm 1 and Algorithm 2 to solve the max-min SINR and the weighted sum of logarithmic SINR utility respectively, we evaluate the numerical convergence of the power and interference temperature iterates.

We use the following channel gain matrix:

$$\mathbf{G} = \begin{bmatrix} 0.69 & 0.12 & 0.14 \\ 0.13 & 0.70 & 0.13 \\ 0.14 & 0.15 & 0.75 \end{bmatrix},$$

and the following weights for the power constraints:

$$\mathbf{w}_1 = [\ 0.8491\ \ 0.9340\ \ 0.6781\ ]^\top$$
$$\mathbf{w}_2 = [\ 0.7577\ \ 0.7431\ \ 0.3922\ ]^\top$$
$$\mathbf{w}_3 = [\ 0.6555\ \ 0.1712\ \ 0.7060\ ]^\top.$$

We set $\bar{\mathbf{p}} = [1.50\ 1.00\ 1.20]^\top$ W and $\bar{\mathbf{q}} = [2.50\ 3.00\ 2.20]^\top$ W. The noise power of each user is 1 W.

Figure 6 plots the evolution of the power and interference temperature for three users that run Algorithm 1 with $\boldsymbol{\beta}$ equal to $\mathbf{1}$. We set the initial power vector to $\mathbf{p}(0) = [0.5\ 0.5\ 0.5]^\top$ W, and run Algorithm 1 for 10 iterations before it terminates. Figure 6 shows that Algorithm 1 converges geometrically fast to the optimal solution (verifying Theorem 2). The optimal $\boldsymbol{\gamma}^\star$ is $[0.3273\ 0.3273\ 0.3273]^\top$ (verifying Lemma 1).

Figure 7 plots the evolution of the power and interference temperature for three users that run Algorithm 2. The objective function that we use in this numerical example is $U(\boldsymbol{\gamma}) = \sum_{l=1}^{L} \omega_l \log \gamma_l$, where $\boldsymbol{\omega}$ is $[0.23\ \ 0.41\ \ 0.36]^\top$. The sum of weighted logarithmic SINR satisfies Assumption 1. We set the initial $\boldsymbol{\eta}$ to $\boldsymbol{\eta}(0) = [0.82\ 0.90\ 1.11]^\top$, and run Algorithm

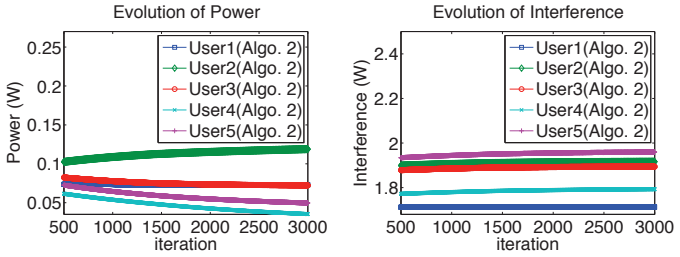


Fig. 9. An illustration of Algorithm 1 for 30 users. In this figure, we show the power and interference temperature evolution for only 5 users.

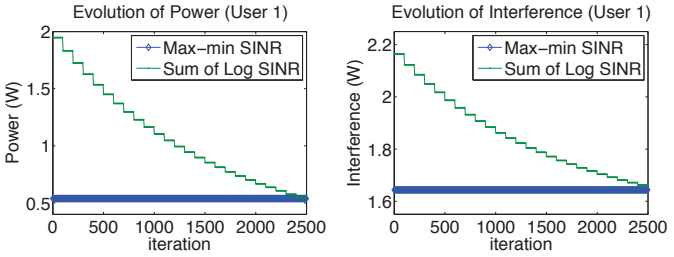


Fig. 10. Performance comparison of Algorithm 1 and Algorithm 2 by solving two equivalent problems, namely (18) and (11) with a utility given by (25). The plots in (a) and (b) show the trajectories of the power and the interference temperature iterates respectively for User 1. The number of iterations in the figure corresponds to the inner loop (faster time-scale). With the weight $\boldsymbol{\omega}$ chosen as in (25), these two algorithms converge to the same solution. As shown, Algorithm 1 converges much faster than Algorithm 2.

2 for 700 iterations and run Algorithm 1 as an inner loop. The stopping criterion for the inner loop is $\epsilon = 10^{-13}$. We use a constant step size, i.e., $\nu(k) = 0.04$ for all k. Figure 7 shows that Algorithm 7 converges to the optimal solution (verifying Theorem 3). The optimal $\mathbf{p}^\star$ is $[0.31\ 0.57\ 0.86]^\top$ W and the optimal $\mathbf{q}^\star$ is $[1.72\ 1.65\ 1.51]^\top$ W so that $\mathbf{w}_2^\top \mathbf{p}^\star$ is equal to $\bar{p}_2$ (verifying Theorem 1). We also adjust the value of ϵ to verify the robustness of our algorithm. Figure 8 shows the evolution of the power and interference temperature when the inner loop of Algorithm 2, i.e., Algorithm 1, terminates with $\epsilon = 10^{-3}$ keeping all the other parameters the same as those used in Figure 7. Furthermore, we also run Algorithm 2 with a suitably large number of primary and secondary users. Figure 9 shows the evolution of the power and interference temperature for five users out of thirty users.

Next, we compare the convergence of the power and interference temperature in two different problems that have the same optimal solution. From the second numerical example, we already know in advance that the second power constraint is tight at optimality. Thus, we can set $U(\boldsymbol{\gamma}) = \sum_{l=1}^{L} \big(\mathbf{x}(\mathbf{B}_2) \circ \mathbf{y}(\mathbf{B}_2)\big)_l \log \gamma_l$. Although the solution of these two different problems are the same, Figure 10 shows that Algorithm 1 converges faster than Algorithm 2.

VI. Conclusion

We studied the network utility maximization in a cognitive radio network with both power budget constraints and interference temperature constraints in this paper. We first reformulated the problem as one in the SINR domain that has appropriately constructed spectral radius constraints. The advantages of our reformulation were that, firstly, it captured the entire feasible SINR region and, secondly, it decoupled the SINR assignment for primary and secondary users from power and interference temperature control. We also studied a special case of egalitarian fairness utility, which is the max-min weighted SINR problem using the nonlinear Perron-Frobenius theory, and a geometrically fast convergent (with no parameter tuning) algorithm was proposed to solve the egalitarian fairness problem. We developed a cognitive radio network duality for the general utility maximization problem that can be used to decouple the SINR assignment, the transmit power and the interference temperature allocation. We leveraged this cognitive radio network duality together with an interference load minimization problem that is related to the egalitarian fairness problem to develop a distributed algorithm to solve the utility maximization problem. Extensive numerical examples demonstrated the good performance of our algorithms, particularly the flexibility, the fast convergence time and the robustness to different numbers of primary and secondary users.

Appendix

A. Proof of Theorem 1

Proof: Our proof is based on the Perron-Frobenius theorem [16]. The inequality constraint set in (10) can be written as:

$$(1/\bar{p}_l)\mathbf{w}_l^\top \mathbf{p} \le 1, \quad l = 1, \ldots, L \tag{48}$$

and

$$(1/\bar{q}_l)\mathbf{e}_l^\top \mathbf{q} \le 1, \quad l = 1, \ldots, L. \tag{49}$$

Combining (48) and (6) yields

$$\operatorname{diag}(\boldsymbol{\gamma} \circ \mathbf{v})(\mathbf{F} + (1/\bar{p}_l)\mathbf{n}\mathbf{w}_l^\top)\mathbf{p} \le \mathbf{p}, \quad l = 1, \ldots, L. \tag{50}$$

Likewise, combining (49) and (7) yields

$$\operatorname{diag}(\mathbf{v})(\mathbf{F}\operatorname{diag}(\boldsymbol{\gamma}) + (1/\bar{q}_l)\mathbf{n}\mathbf{e}_l^\top)\mathbf{q} \le \mathbf{q}, \quad l = 1, \ldots, L. \tag{51}$$

We now state the following lemma:

Lemma 4 (The Subinvariance Theorem [16]): Let $\mathbf{A}$ be an irreducible nonnegative matrix, Λ be a positive number, and $\boldsymbol{v}$ be a nonnegative vector satisfying

$$\mathbf{A}\boldsymbol{v} \le \Lambda \boldsymbol{v}.$$

Then $\boldsymbol{v} > 0$ and $\Lambda > \rho(\mathbf{A})$. Moreover, $\Lambda = \rho(\mathbf{A})$ if and only if $\mathbf{A}\boldsymbol{v} = \Lambda \boldsymbol{v}$.

Letting $\Lambda = 1$, $\mathbf{A} = \operatorname{diag}(\boldsymbol{\gamma} \circ \mathbf{v})(\mathbf{F} + (1/\bar{p}_l)\mathbf{n}\mathbf{w}_l^\top)$ or $\mathbf{A} = \operatorname{diag}(\mathbf{v})(\mathbf{F}\operatorname{diag}(\boldsymbol{\gamma}) + (1/\bar{q}_l)\mathbf{n}\mathbf{e}_l^\top)$ in Lemma 4, the inequalities (50) and (51) can be written as, respectively,

$$\rho\big(\operatorname{diag}(\boldsymbol{\gamma} \circ \mathbf{v})(\mathbf{F} + (1/\bar{p}_l)\mathbf{n}\mathbf{w}_l^\top)\big) \le 1, \quad l = 1, \ldots, L \tag{52}$$

and

$$\rho\big(\operatorname{diag}(\mathbf{v})(\mathbf{F}\operatorname{diag}(\boldsymbol{\gamma}) + (1/\bar{q}_l)\mathbf{n}\mathbf{e}_l^\top)\big) \le 1, \quad l = 1, \ldots, L, \tag{53}$$

which form the constraint set in (11). The optimal $\mathbf{p}^\star$ and $\mathbf{q}^\star$ can simply be obtained from (8) and (9) respectively.

Next, we show that at least one of the constraints in (52) and (53) become tight at optimality. Observe that the spectral radius functions in (52) and (53) are monotonically increasing functions of $\boldsymbol{\gamma}$. Since the utility function is increasing in $\boldsymbol{\gamma}$, maximizing the utility function in (11) leads to at least one of the constraints in (52) and (53) becoming tight at optimality. This means that the corresponding constraints in (48) and (49) become tight, i.e., the tight constraint in (48) has the index i

given in (13), and the tight constraint in (49) has the index j given in (15).

Furthermore, the optimal solutions of (11) and (10) are both found at the boundary of the constraint set. This can be verified as follows. If we have $\rho\big(\operatorname{diag}(\boldsymbol{\gamma}^\star \circ \mathbf{v})(\mathbf{F}+(1/\bar{p}_i)\mathbf{n}\mathbf{w}_i^\top)\big)=1$ at the optimality of (11), it can be rewritten as $\big(\operatorname{diag}(\boldsymbol{\gamma}^\star \circ \mathbf{v})(\mathbf{F}+(1/\bar{p}_i)\mathbf{n}\mathbf{w}_i^\top)\big)\mathbf{p}^\star=\mathbf{p}^\star$. By substituting (6), we obtain $\big(\operatorname{diag}(\boldsymbol{\gamma}^\star \circ \mathbf{v})(\mathbf{F}+(1/\bar{p}_i)\mathbf{n}\mathbf{w}_i^\top)\big)\mathbf{p}^\star=\operatorname{diag}(\boldsymbol{\gamma}^\star \circ \mathbf{v})(\mathbf{F}\mathbf{p}^\star+\mathbf{n})$. Thus, we have $\mathbf{w}_i^\top \mathbf{p}^\star=\bar{p}_i$. ■

B. Proof of Lemma 1

Proof: We first rewrite (51) as $(\mathbf{I}-\frac{1}{\bar{q}_l}\operatorname{diag}(\mathbf{v})\mathbf{n}\mathbf{e}_l^\top)^{-1}\operatorname{diag}(\mathbf{v})\mathbf{F}\operatorname{diag}(\boldsymbol{\gamma})\mathbf{q} \leq \mathbf{q}$, and then show that this is equivalent to $\mathbf{D}_l\operatorname{diag}(\boldsymbol{\gamma})\mathbf{q}\leq\mathbf{q}$, where $\mathbf{D}_l=(\mathbf{I}+\dfrac{1}{\bar{q}_l-n_lv_l}\operatorname{diag}(\mathbf{v})\mathbf{n}\mathbf{e}_l^\top)\operatorname{diag}(\mathbf{v})\mathbf{F}$.

Note that $(\operatorname{diag}(\mathbf{v})\mathbf{n}\mathbf{e}_l^\top)^n=n_lv_l(\operatorname{diag}(\mathbf{v})\mathbf{n}\mathbf{e}_l^\top)^{n-1}$, where n is a positive integer. Applying this recursively, we have $(\operatorname{diag}(\mathbf{v})\mathbf{n}\mathbf{e}_l^\top)^n=(n_lv_l)^{n-1}\big(\operatorname{diag}(\mathbf{v})\mathbf{n}\mathbf{e}_l^\top\big)$.

Next, we use the von Neumann's series expansion to obtain

$$\begin{aligned}
&\left(\mathbf{I}-\frac{1}{\bar{q}_l}\operatorname{diag}(\mathbf{v})\mathbf{n}\mathbf{e}_l^\top\right)^{-1}\\
&= \mathbf{I}+\lim_{N\to\infty}\sum_{n=1}^{N}\Big(\frac{1}{\bar{q}_l}\operatorname{diag}(\mathbf{v})\mathbf{n}\mathbf{e}_l^\top\Big)^n\\
&= \mathbf{I}+\lim_{N\to\infty}\sum_{n=1}^{N}\frac{1}{\bar{q}_l}\Big(\frac{n_lv_l}{\bar{q}_l}\Big)^{n-1}\operatorname{diag}(\mathbf{v})\mathbf{n}\mathbf{e}_l^\top\\
&= \mathbf{I}+\lim_{N\to\infty}\frac{1}{\bar{q}_l}\Big(\frac{1-(n_lv_l/\bar{q}_l)^{N-1}}{1-n_lv_l/\bar{q}_l}\Big)\operatorname{diag}(\mathbf{v})\mathbf{n}\mathbf{e}_l^\top,
\end{aligned}$$

where we use the geometric series formula in the last equality. But, $\lim_{N\to\infty}\left(\dfrac{n_lv_l}{\bar{q}_l}\right)^{N-1}=0$, thus

$$\left(\mathbf{I}-\frac{1}{\bar{q}_l}\operatorname{diag}(\mathbf{v})\mathbf{n}\mathbf{e}_l^\top\right)^{-1}=\mathbf{I}+\frac{1}{\bar{q}_l-n_lv_l}\operatorname{diag}(\mathbf{v})\mathbf{n}\mathbf{e}_l^\top.$$

Hence, by the Subinvariance Theorem (Lemma 4 in the proof of Theorem 1), $\mathbf{D}_l\operatorname{diag}(\boldsymbol{\gamma})\mathbf{q}\leq\mathbf{q}\quad\forall l$ implies that $\rho(\mathbf{D}_l\operatorname{diag}(\boldsymbol{\gamma}))\leq 1\quad\forall l$. We thus have the following reformulation of (18) given by:

$$\begin{array}{ll}\text{maximize} & \min\limits_{l=1,\ldots,L}\dfrac{\gamma_l}{\beta_l}\\ \text{subject to} & \log\rho\big(\operatorname{diag}(\boldsymbol{\gamma})\mathbf{B}_l\big)\leq 0,\quad l=1,\ldots,L,\\ & \log\rho\big(\mathbf{D}_l\operatorname{diag}(\boldsymbol{\gamma})\big)\leq 0,\quad l=1,\ldots,L,\\ \text{variables:} & \boldsymbol{\gamma}.\end{array}\tag{54}$$

Using a logarithmic mapping of variable, i.e., $\tilde{\boldsymbol{\gamma}}=\log\boldsymbol{\gamma}$, we write the epigraph form of (54) as:

$$\begin{array}{ll}\text{maximize} & \tau\\ \text{subject to} & \tau\leq e^{\tilde{\gamma}_l}/\beta_l\quad\forall l,\\ & \log\rho\big(\operatorname{diag}(e^{\tilde{\boldsymbol{\gamma}}})\mathbf{B}_l\big)\leq 0,\quad l=1,\ldots,L,\\ & \log\rho\big(\mathbf{D}_l\operatorname{diag}(e^{\tilde{\boldsymbol{\gamma}}})\big)\leq 0,\quad l=1,\ldots,L,\\ \text{variables:} & \tilde{\boldsymbol{\gamma}},\tau.\end{array}\tag{55}$$

Let $\tau^\star$ be the optimal solution, equivalently the optimal value, of (55). We form the full Lagrangian for (55) by introducing the dual variables $\boldsymbol{\lambda}\in\mathbb{R}_+^L$ for the L inequality constraints $\tau\leq e^{\tilde{\gamma}_l}/\beta_l\quad\forall l$, $\boldsymbol{\mu}\in\mathbb{R}_+^L$ for the L inequality constraints $\log\rho\big(\operatorname{diag}(e^{\tilde{\boldsymbol{\gamma}}})\mathbf{B}_l\big)\leq 0\quad\forall l$, and $\boldsymbol{\nu}\in\mathbb{R}_+^L$ for the L inequality constraints $\log\rho\big(\mathbf{D}_l\operatorname{diag}(e^{\tilde{\boldsymbol{\gamma}}})\big)\leq 0\quad\forall l$ to obtain:

$$\begin{aligned}
L(\tau,\tilde{\boldsymbol{\gamma}},\boldsymbol{\lambda},\boldsymbol{\mu},\boldsymbol{\nu}) &= \tau-\boldsymbol{\lambda}^\top(\tau\mathbf{1}-\operatorname{diag}(\boldsymbol{\beta})^{-1}e^{\tilde{\boldsymbol{\gamma}}})\\
&\quad-\sum_{l=1}^{L}\mu_l\log\rho(\operatorname{diag}(e^{\tilde{\boldsymbol{\gamma}}})\mathbf{B}_l)-\sum_{l=1}^{L}\nu_l\log\rho(\mathbf{D}_l\operatorname{diag}(e^{\tilde{\boldsymbol{\gamma}}})).
\end{aligned}\tag{56}$$

Taking the partial derivative of (56) with respect to $\tilde{\boldsymbol{\gamma}}$, we have

$$\begin{aligned}
\partial L/\partial\tilde{\boldsymbol{\gamma}} &= (\operatorname{diag}(\boldsymbol{\beta})^{-1}\boldsymbol{\lambda})\circ e^{\tilde{\boldsymbol{\gamma}}}\\
&\quad-\sum_{l=1}^{L}\mu_l\big(\mathbf{x}(\operatorname{diag}(e^{\tilde{\boldsymbol{\gamma}}})\mathbf{B}_l)\circ\mathbf{y}(\operatorname{diag}(e^{\tilde{\boldsymbol{\gamma}}})\mathbf{B}_l)\big)\\
&\quad-\sum_{l=1}^{L}\nu_l\big(\mathbf{x}(\mathbf{D}_l\operatorname{diag}(e^{\tilde{\boldsymbol{\gamma}}}))\circ\mathbf{y}(\mathbf{D}_l\operatorname{diag}(e^{\tilde{\boldsymbol{\gamma}}}))\big),
\end{aligned}$$

and, setting it to zero, we have at optimality:

$$\begin{aligned}
&(\operatorname{diag}(\boldsymbol{\beta})^{-1}\boldsymbol{\lambda}^\star)\circ e^{\tilde{\boldsymbol{\gamma}}^\star}=\\
&\qquad\sum_{l=1}^{L}\mu_l^\star\big(\mathbf{x}(\operatorname{diag}(e^{\tilde{\boldsymbol{\gamma}}^\star})\mathbf{B}_l)\circ\mathbf{y}(\operatorname{diag}(e^{\tilde{\boldsymbol{\gamma}}^\star})\mathbf{B}_l)\big)\\
&\qquad+\sum_{l=1}^{L}\nu_l^\star\big(\mathbf{x}(\mathbf{D}_l\operatorname{diag}(e^{\tilde{\boldsymbol{\gamma}}^\star}))\circ\mathbf{y}(\mathbf{D}_l\operatorname{diag}(e^{\tilde{\boldsymbol{\gamma}}^\star}))\big).
\end{aligned}\tag{57}$$

According to the Perron-Frobenius theorem, the Perron right and left eigenvectors of an irreducible nonnegative matrix are nonnegative [16]. Since at least one constraint is tight at optimality and its corresponding dual variable $\mu_l^\star$ or $\nu_l^\star$ is positive for some l, we know from (57) that $\boldsymbol{\lambda}^\star$ is a positive vector. Combining the fact that $\boldsymbol{\lambda}^\star>0$ with complementary slackness, we have $\tau^\star=\gamma_l^\star/\beta_l$ for all l. This means that

$$\operatorname{diag}(\boldsymbol{\beta})^{-1}\boldsymbol{\gamma}^\star=\tau^\star\mathbf{1}.\tag{58}$$

Then, (55) can be further written as

$$\begin{array}{ll}\text{maximize} & \tau\\ \text{subject to} & \tau\rho\big(\operatorname{diag}(\boldsymbol{\beta})\mathbf{B}_l\big)\leq 1,\quad l=1,\ldots,L,\\ & \tau\rho\big(\mathbf{D}_l\operatorname{diag}(\boldsymbol{\beta})\big)\leq 1,\quad l=1,\ldots,L,\\ \text{variables:} & \tau.\end{array}\tag{59}$$

Using (58), the constraints in (59) can be further written as $\tau^\star\leq 1/\rho(\operatorname{diag}(\boldsymbol{\beta})\mathbf{B}_l)$ and $\tau^\star\leq 1/\rho(\mathbf{D}_l\operatorname{diag}(\boldsymbol{\beta}))$ for all l. As such, τ is maximal when $\tau=\min\limits_{l=1,\ldots,L}\{1/\rho(\operatorname{diag}(\boldsymbol{\beta})\mathbf{B}_l),1/\rho(\mathbf{D}_l\operatorname{diag}(\boldsymbol{\beta}))\}$.

By the Perron-Frobenius theorem, if $\tau^\star=1/\rho(\operatorname{diag}(\boldsymbol{\beta})\mathbf{B}_l)$ for some l, i.e., $\operatorname{diag}(\boldsymbol{\beta})\mathbf{B}_l\mathbf{p}^\star=\rho(\operatorname{diag}(\boldsymbol{\beta})\mathbf{B}_l)\mathbf{p}^\star$ for some l, then $\mathbf{p}^\star$ is the Perron right eigenvector of $\operatorname{diag}(\boldsymbol{\beta})\mathbf{B}_l$ that is normalized to be tight for the lth power constraint, and $\mathbf{q}^\star$ can be computed in terms of $\mathbf{p}^\star$ and $\tau^\star$ as in (23); Likewise, if $\tau^\star=1/\rho(\mathbf{D}_l\operatorname{diag}(\boldsymbol{\beta}))$ for some l, i.e., $\mathbf{D}_l\operatorname{diag}(\boldsymbol{\beta})\mathbf{q}^\star=\rho(\mathbf{D}_l\operatorname{diag}(\boldsymbol{\beta}))\mathbf{q}^\star$ for some l, then $\mathbf{q}^\star$ is the Perron right eigenvector of $\mathbf{D}_l\operatorname{diag}(\boldsymbol{\beta})$ that is normalized to be tight for this lth interference temperature constraint, and $\mathbf{p}^\star$ can be computed in terms of $\mathbf{q}^\star$ and $\tau^\star$ as in (24). ■

C. Proof of Theorem 2

Proof: Combining $\operatorname{diag}(\boldsymbol{\beta})^{-1}\boldsymbol{\gamma}^\star=\tau^\star\mathbf{1}$ in (58) and $\gamma_l^\star=\dfrac{p_l^\star}{\big(\operatorname{diag}(\mathbf{v})(\mathbf{F}\mathbf{p}^\star+\mathbf{n})\big)_l}$ in (4), we have

$$(1/\tau^\star)\mathbf{p}^\star=\operatorname{diag}(\boldsymbol{\beta}\circ\mathbf{v})(\mathbf{F}\mathbf{p}^\star+\mathbf{n}),\tag{60}$$

Now, we will show that $\mathbf{p}^\star$ in (60) is further constrained by a monotone norm. Let us rewrite the interference temperature constraints in (49) as follows. By substituting $\mathbf{q}=\operatorname{diag}(\mathbf{v})(\mathbf{F}\mathbf{p}+\mathbf{n})$ into $\mathbf{e}_l^\top\mathbf{q}\leq\bar{q}_l$, we have:

$$\mathbf{e}_l^\top\operatorname{diag}(\mathbf{v})(\mathbf{F}\mathbf{p}+\mathbf{n})\leq\bar{q}_l,\quad l=1,\ldots,L$$

$$\Rightarrow \frac{1}{q_l - n_l v_l}\mathbf{e}_l^\top \operatorname{diag}(\mathbf{v})\mathbf{F}\mathbf{p} \le 1, \quad l = 1, \ldots, L.$$

Combining this with (48), $\mathbf{p}^\star$ must satisfy

$$\max_{l=1,\ldots,L}\left\{\frac{\mathbf{w}_l^\top \mathbf{p}^\star}{\bar{p}_l}, \frac{\mathbf{e}_l^\top \operatorname{diag}(\mathbf{v})\mathbf{F}\mathbf{p}^\star}{\bar{q}_l - n_l v_l}\right\} = 1, \tag{61}$$

which is a monotone norm constraint in $\mathbf{p}^\star$. Next, we state the following nonlinear Perron-Frobenius theory result in [21].

Theorem 4 (Krause's theorem [21]): Let $\|\cdot\|$ be a monotone norm on $\mathbb{R}^L$. For a concave mapping $f : \mathbb{R}_+^L \to \mathbb{R}_+^L$ with $f(\mathbf{z}) > 0$ for $\mathbf{z} \ge \mathbf{0}$, the following statements hold. The conditional eigenvalue problem $f(\mathbf{z}) = \lambda\mathbf{z}$, $\lambda \in \mathbb{R}$, $\mathbf{z} \ge \mathbf{0}$, $\|\mathbf{z}\| = 1$ has a unique solution $(\lambda^*, \mathbf{z}^*)$, where $\lambda^* > 0, \mathbf{z}^* > \mathbf{0}$. Furthermore, $\lim_{k\to\infty} \tilde{f}(\mathbf{z}(k))$ converges geometrically fast to $\mathbf{z}^*$, where $\tilde{f}(\mathbf{z}) = f(\mathbf{z})/\|f(\mathbf{z})\|$.

Let $f(\mathbf{p}) = \operatorname{diag}(\boldsymbol{\beta}\circ\mathbf{v})(\mathbf{F}\mathbf{p}+\mathbf{n})$, $\lambda = 1/\tau$, and we constrain $\mathbf{p}$ by the monotone norm $\|\cdot\|$ given on the lefthand side of (61). Then, by Theorem 4, $\mathbf{p}^\star$ and $\tau^\star$ are unique, and the convergence of the iteration

$$\mathbf{p}(k+1) = \frac{\operatorname{diag}(\boldsymbol{\beta}\circ\mathbf{v})(\mathbf{F}\mathbf{p}(k)+\mathbf{n})}{\max_{l=1,\ldots,L}\left\{\frac{\mathbf{w}_l^\top \mathbf{p}(k)}{\bar{p}_l}, \frac{\mathbf{e}_l^\top \operatorname{diag}(\mathbf{v})\mathbf{F}\mathbf{p}(k)}{\bar{q}_l - n_l v_l}\right\}}$$

to the unique fixed point $\mathbf{p}^\star = f(\mathbf{p}^\star)/\|f(\mathbf{p}^\star)\|$ is geometrically fast, regardless of the initial point. ■

D. Proof of Lemma 2

Proof: From the proof of Theorem 1, we know that at least one of the inequality constraints in (10) will be tight at optimality. Thus, if one of the power constraints is tight, we have equalities that ${\mathbf{w}_i}^\top \mathbf{p}^\star = \bar{p}_i$ and $\mathbf{v}^\top \mathbf{s}^\star = \bar{p}_i$ at optimality. Now, ${\mathbf{w}_i}^\top \mathbf{p}^\star = \bar{p}_i$ is equivalent to $\operatorname{diag}(\boldsymbol{\gamma}^\star)\mathbf{B}_i\mathbf{p}^\star = \mathbf{p}^\star$. We then have $\mathbf{p}^\star = \mathbf{x}(\operatorname{diag}(\boldsymbol{\gamma}^\star)\mathbf{B}_i)$ which yields (35).

Next, by combining $(1/\bar{p}_i)\mathbf{v}^\top\mathbf{s}^\star = 1$ and (31), we have:

$$\begin{aligned}
&\operatorname{diag}(\boldsymbol{\gamma}^\star\circ\mathbf{v})\big(\mathbf{F}^\top + (1/\bar{p}_i)\mathbf{w}_i\mathbf{n}^\top\big)\mathbf{s}^\star = \mathbf{s}^\star \\
&\Rightarrow {\mathbf{s}^\star}^\top\big(\mathbf{F} + (1/\bar{p}_i)\mathbf{n}\mathbf{w}_i^\top\big) = \operatorname{diag}(\boldsymbol{\gamma}^\star\circ\mathbf{v})^{-1}{\mathbf{s}^\star}^\top \\
&\Rightarrow \big(\operatorname{diag}(\boldsymbol{\gamma}^\star\circ\mathbf{v})^{-1}\mathbf{s}^\star\big)^\top \operatorname{diag}(\boldsymbol{\gamma}^\star)\mathbf{B}_i \\
&\quad = (\operatorname{diag}(\boldsymbol{\gamma}^\star\circ\mathbf{v})^{-1}\mathbf{s}^\star)^\top.
\end{aligned}$$

Thus, we have $\operatorname{diag}(\boldsymbol{\gamma}^\star\circ\mathbf{v})^{-1}\mathbf{s}^\star = \mathbf{y}(\operatorname{diag}(\boldsymbol{\gamma}^\star)\mathbf{B}_i)$ which yields (36).

From Lemma 1, we know that $\mathbf{e}_j^\top\mathbf{q}^\star = \bar{q}_j$ is equivalent to $\Big(\mathbf{I} + \frac{1}{\bar{q}_j - n_j v_j}\operatorname{diag}(\mathbf{v})\mathbf{n}\mathbf{e}_j^\top\Big)\operatorname{diag}(\mathbf{v})\mathbf{F}\operatorname{diag}(\boldsymbol{\gamma}^\star)\mathbf{q}^\star = \mathbf{q}^\star$. We then have $\mathbf{q}^\star = \mathbf{x}\big(\mathbf{D}_j\operatorname{diag}(\boldsymbol{\gamma}^\star)\big)$ which yields (37).

Next, we develop the connection between $\mathbf{t}^\star$ and $\mathbf{y}\Big(\big(\mathbf{I} + \frac{1}{\bar{q}_j - n_j v_j}\operatorname{diag}(\mathbf{v})\mathbf{n}\mathbf{e}_j^\top\big)\operatorname{diag}(\mathbf{v})\mathbf{F}\operatorname{diag}(\boldsymbol{\gamma}^\star)\Big)$.

Set $\boldsymbol{\varphi}$ as the Perron left eigenvector of $\big(\mathbf{I} + \frac{1}{\bar{q}_j - n_j v_j}\operatorname{diag}(\mathbf{v})\mathbf{n}\mathbf{e}_j^\top\big)\operatorname{diag}(\mathbf{v})\mathbf{F}\operatorname{diag}(\boldsymbol{\gamma}^\star)$ corresponding to the eigenvalue 1, i.e., $\boldsymbol{\varphi} = \mathbf{y}\Big(\big(\mathbf{I} + \frac{1}{\bar{q}_j - n_j v_j}\operatorname{diag}(\mathbf{v})\mathbf{n}\mathbf{e}_j^\top\big)\operatorname{diag}(\mathbf{v})\mathbf{F}\operatorname{diag}(\boldsymbol{\gamma}^\star)\Big)$. We have the following equations:

$$\begin{cases}
\mathbf{t}^\star = \operatorname{diag}(\mathbf{v})\mathbf{F}^\top\operatorname{diag}(\boldsymbol{\gamma}^\star)\mathbf{t}^\star + \mathbf{w}_j, \\
\boldsymbol{\varphi}^\top = \boldsymbol{\varphi}^\top\Big(\big(\mathbf{I} + \frac{1}{\bar{q}_j - n_j v_j}\operatorname{diag}(\mathbf{v})\mathbf{n}\mathbf{e}_j^\top\big)\operatorname{diag}(\mathbf{v})\mathbf{F}\operatorname{diag}(\boldsymbol{\gamma}^\star)\Big), \\
\mathbf{n}^\top\mathbf{t}^\star = \bar{q}_j.
\end{cases} \tag{62}$$

From $\boldsymbol{\varphi}^\top = \boldsymbol{\varphi}^\top\Big(\big(\mathbf{I} + \frac{1}{\bar{q}_j - n_j v_j}\operatorname{diag}(\mathbf{v})\mathbf{n}\mathbf{e}_j^\top\big)\operatorname{diag}(\mathbf{v})\mathbf{F}\operatorname{diag}(\boldsymbol{\gamma}^\star)\Big)$ in (62), we can obtain:

$$\boldsymbol{\varphi} = \Big(\frac{\mathbf{n}^\top\operatorname{diag}(\mathbf{v})\boldsymbol{\varphi}}{\bar{q}_j - n_j v_j}\Big)\Big(\mathbf{I} - \operatorname{diag}(\boldsymbol{\gamma}^\star)\mathbf{F}^\top\operatorname{diag}(\mathbf{v})\Big)^{-1}\mathbf{e}_j,$$

where $\frac{\mathbf{n}^\top\operatorname{diag}(\mathbf{v})\boldsymbol{\varphi}}{\bar{q}_j - n_j v_j}$ is a scalar. Since $\boldsymbol{\varphi}$ is an eigenvector that can be scaled, we have

$$\boldsymbol{\varphi} = \Big(\mathbf{I} - \operatorname{diag}(\boldsymbol{\gamma}^\star)\mathbf{F}^\top\operatorname{diag}(\mathbf{v})\Big)^{-1}\mathbf{e}_j.$$

$$\Rightarrow \boldsymbol{\varphi} = \operatorname{diag}(\mathbf{v})^{-1}\Big(\operatorname{diag}(\mathbf{v})^{-1} - \operatorname{diag}(\boldsymbol{\gamma}^\star)\mathbf{F}^\top\Big)^{-1}\mathbf{e}_j.$$

On the other hand, the expression for the dual network power is

$$\begin{aligned}
\mathbf{s}^\star &= \Big(\mathbf{I} - \operatorname{diag}(\boldsymbol{\gamma}^\star\circ\mathbf{v})\mathbf{F}^\top\Big)^{-1}\operatorname{diag}(\boldsymbol{\gamma}^\star\circ\mathbf{v})\mathbf{w}_j \\
&= \Big(\operatorname{diag}(\mathbf{v})^{-1} - \operatorname{diag}(\boldsymbol{\gamma}^\star)\mathbf{F}^\top\Big)^{-1}\operatorname{diag}(\boldsymbol{\gamma}^\star)\mathbf{w}_j.
\end{aligned}$$

Letting $\mathbf{M} = \big(\operatorname{diag}(\mathbf{v})^{-1} - \operatorname{diag}(\boldsymbol{\gamma}^\star)\mathbf{F}^\top\big)^{-1}$, we then have:

$$\begin{aligned}
\mathbf{s}^\star &= \sum_{l\ne j}\mathbf{M}\operatorname{diag}(\boldsymbol{\gamma}^\star\circ\mathbf{w}_j)\mathbf{e}_l + \operatorname{diag}(\mathbf{v})\boldsymbol{\varphi} \\
\Rightarrow \sum_{l=1}^{L}\mathbf{s}^\star &= \sum_{l=1}^{L}\Big(\sum_{l\ne j}\mathbf{M}\operatorname{diag}(\boldsymbol{\gamma}^\star\circ\mathbf{w}_j)\mathbf{e}_l + \operatorname{diag}(\mathbf{v})\boldsymbol{\varphi}\Big) \\
\Rightarrow L\mathbf{s}^\star &= (L-1)\mathbf{s}^\star + L\operatorname{diag}(\mathbf{v})\boldsymbol{\varphi} \\
\Rightarrow \mathbf{s}^\star &= \operatorname{diag}(\mathbf{v})\boldsymbol{\varphi}.
\end{aligned}$$

Since, from (28), $\mathbf{s} = \operatorname{diag}(\gamma)\mathbf{t}$, thus $\boldsymbol{\varphi} = \operatorname{diag}(\boldsymbol{\gamma}^\star/\mathbf{v})\mathbf{t}^\star$. We thus obtain: $\mathbf{y}\Big(\big(\mathbf{I} + \frac{1}{\bar{q}_j - n_j v_j}\operatorname{diag}(\mathbf{v})\mathbf{n}\mathbf{e}_j^\top\big)\operatorname{diag}(\mathbf{v})\mathbf{F}\operatorname{diag}(\boldsymbol{\gamma}^\star)\Big) = \operatorname{diag}(\boldsymbol{\gamma}^\star/\mathbf{v})\mathbf{t}^\star$ which yields (38). ■

E. Proof of Lemma 3

Proof: We first write (40) in the epigraph form as

$$\begin{aligned}
\text{minimize} \quad & \zeta \\
\text{subject to} \quad & \log\rho(\operatorname{diag}(e^{\tilde{\boldsymbol{\eta}}})\mathbf{B}_l) \le \log\zeta, \quad l = 1, \ldots, L, \\
& \log\rho(\mathbf{D}_l\operatorname{diag}(e^{\tilde{\boldsymbol{\eta}}})) \le \log\zeta, \quad l = 1, \ldots, L, \\
& \boldsymbol{\alpha}^\top\tilde{\boldsymbol{\eta}} \ge 0, \\
\text{variables:} \quad & \zeta, \tilde{\boldsymbol{\eta}}.
\end{aligned} \tag{63}$$

We use the Karush-Kuhn-Tucker (KKT) conditions to connect the optimality conditions of (39) and (63). To form the Lagrangian for (39) and (63), we introduce the dual variables $\boldsymbol{\mu} \in \mathbb{R}_+^L$, $\boldsymbol{\nu} \in \mathbb{R}_+^L$ and $\lambda \in \mathbb{R}_+$. The Lagrangian for (39) and (63) are, respectively,

$$\begin{aligned}
L(\tilde{\boldsymbol{\gamma}}, \boldsymbol{\mu}, \boldsymbol{\nu}) &= -\boldsymbol{\alpha}^\top\tilde{\boldsymbol{\gamma}} + \sum_{l=1}^{L}\mu_l\log\rho\big(\operatorname{diag}(e^{\tilde{\gamma}})\mathbf{B}_l\big) \\
&\quad + \sum_{l=1}^{L}\nu_l\log\rho\big(\mathbf{D}_l\operatorname{diag}(e^{\tilde{\gamma}})\big),
\end{aligned}$$

and

$$\begin{aligned}
L(\tilde{\boldsymbol{\eta}}, \zeta, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\nu}}, \lambda) &= \zeta + \sum_{l=1}^{L}\hat{\mu}_l\Big(\log\rho\big(\operatorname{diag}(e^{\tilde{\boldsymbol{\eta}}})\mathbf{B}_l\big) - \log\zeta\Big) \\
&\quad + \sum_{l=1}^{L}\hat{\nu}_l\Big(\log\rho\big(\mathbf{D}_l\operatorname{diag}(e^{\tilde{\boldsymbol{\eta}}})\big) - \log\zeta\Big) - \lambda\boldsymbol{\alpha}^\top\tilde{\boldsymbol{\eta}}.
\end{aligned}$$

Taking the first-order derivative of $L(\tilde{\boldsymbol{\gamma}}, \boldsymbol{\mu}, \boldsymbol{\nu})$ and $L(\tilde{\boldsymbol{\eta}}, \zeta, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\nu}}, \lambda)$ with respect to $\tilde{\boldsymbol{\gamma}}$ and $\tilde{\boldsymbol{\eta}}$ respectively and

setting them to zero, we have respectively, at optimality:

$$\sum_{l=1}^{L}\mu_l^\star\Big(\mathbf{x}\big(\operatorname{diag}(e^{\tilde{\boldsymbol{\gamma}}^\star})\mathbf{B}_l\big)\circ\mathbf{y}\big(\operatorname{diag}(e^{\tilde{\boldsymbol{\gamma}}^\star})\mathbf{B}_l\big)\Big) + \sum_{l=1}^{L}\nu_l^\star\Big(\mathbf{x}\big(\mathbf{D}_l\operatorname{diag}(e^{\tilde{\boldsymbol{\gamma}}^\star})\big)\circ\mathbf{y}\big(\mathbf{D}_l\operatorname{diag}(e^{\tilde{\boldsymbol{\gamma}}^\star})\big)\Big)=\boldsymbol{\alpha}, \tag{64}$$

and

$$\sum_{l=1}^{L}\hat{\mu}_l^\star\Big(\mathbf{x}\big(\operatorname{diag}(e^{\tilde{\boldsymbol{\eta}}^\star})\mathbf{B}_l\big)\circ\mathbf{y}\big(\operatorname{diag}(e^{\tilde{\boldsymbol{\eta}}^\star})\mathbf{B}_l\big)\Big) + \sum_{l=1}^{L}\hat{\nu}_l^\star\Big(\mathbf{x}\big(\mathbf{D}_l\operatorname{diag}(e^{\tilde{\boldsymbol{\eta}}^\star})\big)\circ\mathbf{y}\big(\mathbf{D}_l\operatorname{diag}(e^{\tilde{\boldsymbol{\eta}}^\star})\big)\Big)=\lambda^\star\boldsymbol{\alpha}. \tag{65}$$

Taking the first-order derivative of $L(\tilde{\boldsymbol{\eta}},\zeta,\hat{\boldsymbol{\mu}},\hat{\boldsymbol{\nu}},\lambda)$ with respect to ζ, we have at optimality: $\sum_{l=1}^{L}\hat{\mu}_l^\star+\sum_{l=1}^{L}\hat{\nu}_l^\star=\zeta^\star$.

Since the Schur product of the Perron right and left eigenvectors can be normalized as a probability vector, we sum both the lefthand-side and righthand-side entries of the vector equations in (64) and (65) to obtain respectively $\sum_{l=1}^{L}\mu_l^\star+\sum_{l=1}^{L}\nu_l^\star=1$ and $\sum_{l=1}^{L}\hat{\mu}_l^\star+\sum_{l=1}^{L}\hat{\nu}_l^\star=\lambda^\star$.

Thus, we have $\zeta^\star=\lambda^\star$ and $\zeta\big(\sum_{l=1}^{L}\mu_l^\star+\sum_{l=1}^{L}\nu_l^\star\big)=\sum_{l=1}^{L}\hat{\mu}_l^\star+\sum_{l=1}^{L}\hat{\nu}_l^\star$, respectively.

Leveraging the equations that $\mathbf{x}\big(\operatorname{diag}(e^{\tilde{\boldsymbol{\eta}}^\star})\mathbf{B}_l\big)=\mathbf{x}\big(\operatorname{diag}(\zeta^\star e^{\tilde{\boldsymbol{\eta}}^\star})\mathbf{B}_l\big)$, $\mathbf{x}\big(\mathbf{D}_l\operatorname{diag}(e^{\tilde{\boldsymbol{\eta}}^\star})\big)=\mathbf{x}\big(\mathbf{D}_l\operatorname{diag}(\zeta^\star e^{\tilde{\boldsymbol{\eta}}^\star})\big)$, $\mathbf{y}\big(\operatorname{diag}(e^{\tilde{\boldsymbol{\eta}}^\star})\mathbf{B}_l\big)=\mathbf{y}\big(\operatorname{diag}(\zeta^\star e^{\tilde{\boldsymbol{\eta}}^\star})\mathbf{B}_l\big)$, and $\mathbf{y}\big(\mathbf{D}_l\operatorname{diag}(e^{\tilde{\boldsymbol{\eta}}^\star})\big)=\mathbf{y}\big(\mathbf{D}_l\operatorname{diag}(\zeta^\star e^{\tilde{\boldsymbol{\eta}}^\star})\big)$ for all l, we have $e^{\tilde{\boldsymbol{\gamma}}^\star}=\frac{1}{\zeta^\star}e^{\tilde{\boldsymbol{\eta}}^\star}$, i.e., $\boldsymbol{\gamma}^\star=\frac{1}{\zeta^\star}\boldsymbol{\eta}^\star$.

Now, we prove the connection between the optimal value of (39) and that of (40) by solving the following equations: $e^{\tilde{\boldsymbol{\gamma}}^\star}=\dfrac{1}{\zeta^\star}e^{\tilde{\boldsymbol{\eta}}^\star},\boldsymbol{\alpha}^\top\tilde{\boldsymbol{\gamma}}^\star=\xi^\star,\boldsymbol{\alpha}^\top\tilde{\boldsymbol{\eta}}^\star=0,\sum_{l=1}^{L}\alpha_l=1$. We thus deduce that $\xi^\star=-\log\zeta^\star$. ∎

F. Proof of Theorem 3

Proof: Since, under Assumption 1, (16) is a convex optimization problem, we can solve it by a successive convex approximation technique. In particular, if $U(e^{\tilde{\boldsymbol{\gamma}}})$ is smooth, we replace the objective function of (16) by its Taylor series approximation (up to the first order terms):

$$U(e^{\tilde{\boldsymbol{\gamma}}})\approx U(e^{\hat{\boldsymbol{\gamma}}})+\nabla U(e^{\hat{\boldsymbol{\gamma}}})^\top(\tilde{\boldsymbol{\gamma}}-\hat{\boldsymbol{\gamma}}),$$

where $\hat{\boldsymbol{\gamma}}$ is any feasible point. At the $(k+1)$th iteration, we compute a feasible $\tilde{\boldsymbol{\gamma}}(k+1)$ by solving the $(k+1)$th approximation problem:

$$\begin{array}{ll}\text{minimize} & \nabla U(e^{\hat{\boldsymbol{\gamma}}(k)})^\top(\tilde{\boldsymbol{\gamma}}-\tilde{\boldsymbol{\gamma}}(k))\\ \text{subject to} & \log\rho(\operatorname{diag}(e^{\tilde{\boldsymbol{\gamma}}})\mathbf{B}_l)\le 0,\quad l=1,\ldots,L,\\ & \log\rho(\mathbf{D}_l\operatorname{diag}(e^{\tilde{\boldsymbol{\gamma}}}))\le 0,\quad l=1,\ldots,L,\\ \text{variables:} & \tilde{\boldsymbol{\gamma}},\end{array} \tag{66}$$

where $\tilde{\boldsymbol{\gamma}}(k)$ is the optimal solution of the kth approximation problem. This $(k+1)$th approximation problem has a problem structure similar to (39). We can then set $\boldsymbol{\alpha}:=\boldsymbol{\alpha}(k)$ to be $\nabla U(e^{\hat{\boldsymbol{\gamma}}(k)})$ in (39) and set the objective function in (39) as $\boldsymbol{\alpha}(k)^\top\tilde{\boldsymbol{\gamma}}$.

In addition, from Lemma 3, solving (39) is equivalent to solving (40). An interpretation of Lemma 3 is as follows: $\boldsymbol{\alpha}^\top\log\boldsymbol{\eta}(k)$ gives the directional derivative of $U(\boldsymbol{\eta}(k))$ at $\log\boldsymbol{\eta}(k)$ in the direction $\log\boldsymbol{\eta}(k)$, which is orthogonal at the optimal point. Therefore, we let the lower bound of $\boldsymbol{\alpha}^\top\boldsymbol{\eta}$ be 0. On the other hand, if $U(e^{\tilde{\boldsymbol{\gamma}}})$ is non-smooth, then $\boldsymbol{\alpha}$ is its subgradient according to Definition 1.

Next, by making use of the projected subgradient method to solve (66), Algorithm 2 computes the gradient in a distributed way. As stated in Lemma 2, by replacing $\boldsymbol{\gamma}^\star$ with $\boldsymbol{\eta}(k)$, we have $\mathbf{p}(k)=\mathbf{x}(\operatorname{diag}(\boldsymbol{\eta}(k))\mathbf{B}_l)$, $\mathbf{s}(k)=\operatorname{diag}(\boldsymbol{\eta}(k))\mathbf{y}(\operatorname{diag}(\boldsymbol{\eta}(k))\mathbf{B}_l)$, $\mathbf{q}(k)=\mathbf{x}\big(\mathbf{D}_l\operatorname{diag}(\boldsymbol{\eta}(k))\big)$, and $\mathbf{t}(k)=\operatorname{diag}(\boldsymbol{\eta}(k))^{-1}\mathbf{y}\big(\mathbf{D}_l\operatorname{diag}(\boldsymbol{\eta}(k))\big)$. We also have that the gradient $\mathbf{g}\in\mathbb{R}^L$ of the constraint set in (66) can be given respectively by (42) and (43) in terms of $\mathbf{p}(k)$, $\mathbf{q}(k)$, $\mathbf{s}(k)$ and $\mathbf{t}(k)$. Then the gradient can be written as in Step 3 of Algorithm 2 as $g_l(k)=p_l(k)s_l(k)/(\eta_l(k)v_l)$ or $g_l(k)=q_l(k)t_l(k)\eta_l(k)/v_l$ for the lth user if its power (corresponding to the index i_k) or interference temperature constraints (corresponding to the index j_k) are violated respectively. After the gradient update, the scaling factor is updated by $\boldsymbol{\eta}(k+1)=e^{\tilde{\boldsymbol{\eta}}(k+1)}=\boldsymbol{\eta}(k)e^{-\nu(k)(\mathbf{g}(k)/\mathbf{1}^\top\mathbf{g}(k))}$.

In particular, when we use a diminishing step size rule in [24], i.e., $\sum_{k=0}^{\infty}\nu(k)=\infty$, $\sum_{k=0}^{\infty}(\nu(k))^2<\infty$, the projected subgradient method is guaranteed to converge to the optimal solution of (40). This proves the convergence of Algorithm 2. ∎

REFERENCES

[1] R. Etkin, A. Parekh, and D. Tse, "Spectrum sharing for unlicensed bands," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 3, pp. 517–528, 2007.

[2] P. Hande, S. Rangan, M. Chiang, and X. Wu, "Distributed uplink power control for optimal SIR assignment in cellular data networks," *IEEE/ACM Trans. Netw.*, vol. 16, no. 6, pp. 1420–1433, 2008.

[3] S. Huang, X. Liu, and Z. Ding, "Decentralized cognitive radio control based on inference from primary link control information," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 2, pp. 394–406, 2011.

[4] M. Lotfinezhad, L. Ben, and E. S. Sousa, "Optimal control of constrained cognitive radio networks with dynamic population size," *Proc. IEEE Infocom*, 2010.

[5] S. Rangan and R. Madan, "Belief propagation methods for intercell interference coordination in femtocell networks," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 3, pp. 631–640, 2012.

[6] A. L. Stolyar and H. Viswanathan, "Self-organizing dynamic fractional frequency reuse for best-effort traffic through distributed inter-cell coordination," *Proc. IEEE Infocom*, 2009.

[7] A. L. Stolyar and H. Viswanathan, "Self-organizing dynamic fractional frequency reuse in OFDMA systems," *Proc. IEEE Infocom*, 2008.

[8] K. R. Krishnan and H. Luss, "Power selection for maximizing SINR in femtocells for specified SINR in macrocell," *Proc. IEEE WCNC*, 2011.

[9] P. Viswanath and D. N. C. Tse, "Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality," *IEEE Trans. Inf. Theory*, vol. 49, pp. 1912–1921, Aug. 2003.

[10] W. Yu, "Uplink-downlink duality via minimax duality," *IEEE Trans. Inf. Theory*, vol. 52, pp. 361–374, Feb. 2006.

[11] F. Rashid-Farrokhi, K. J. R. Liu, and L. Tassiulas, "Transmit beamforming and power control for cellular wireless systems," *IEEE J. Sel. Areas Commun.*, vol. 16, pp. 1437–1449, Oct. 1998.

[12] B. Song, R. L. Cruz, and B. D. Rao, "Network duality for multiuser MIMO beamforming networks and applications," *IEEE Trans. Commun.*, vol. 55, no. 3, pp. 618–630, 2007.

[13] C. W. Tan, M. Chiang, and R. Srikant, "Maximizing sum rate and minimizing MSE on multiuser downlink: Optimality, fast algorithms and equivalence via max-min SINR," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 6127–6143, 2011.

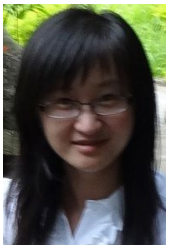
[14] D. W. H. Cai, T. Quek, and C. W. Tan, "A unified analysis of max-min weighted SINR for MIMO downlink system," *IEEE Trans. Signal Process.*, vol. 59, no. 8, pp. 3850–3862, 2011.

[15] D. W. H. Cai, T. Quek, C. W. Tan, and S. H. Low, "Max-min SINR coordinated multipoint downlink transmission - duality and algorithms," *IEEE Trans. Signal Process.*, vol. 60, no. 10, pp. 5384–5395, 2012.

[16] E. Seneta, *Non-Negative Matrices and Markov Chains*. New York: Springer-Verlag, 2nd ed., 1981.

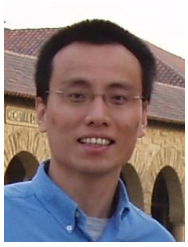
[17] J. Mo and J. Walrand, "Fair end-to-end window-based congestion control," *IEEE/ACM Trans. Netw.*, vol. 8, no. 5, pp. 556–567, 2000.

[18] J. S. Pang, G. Scutari, D. P. Palomar, and F. Facchinei, "Design of cognitive radio systems under temperature-interference constraints: a variational inequality approach," *IEEE Trans. Signal Process.*, vol. 58, no. 6, pp. 3251–3271, 2007.
[19] J. F. C. Kingman, "A convexity property of positive matrices," *Proc. American Mathematical Society*, vol. 12, no. 2, pp. 283–284, 1961.
[20] S. Boyd and L. Vanderberghe, *Convex Optimization*. Cambridge University Press, 2004.
[21] U. Krause, "Concave Perron-Frobenius theory and applications," *Nonlinear analysis*, vol. 47, no. 2001, pp. 1457–1466, 2001.
[22] C. W. Tan, S. Friedland, and S. H. Low, "Spectrum management in multiuser cognitive wireless networks: optimality and algorithm," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 2, pp. 421–430, 2011.
[23] C. W. Tan, S. Friedland, and S. H. Low, "Nonnegative matrix inequalities and their application to nonconvex power control optimization," *SIAM Journal on Matrix Analysis and Applications*, vol. 32, no. 3, pp. 1030–1055, 2011.
[24] D. P. Bertsekas, *Nonlinear Programming*. Belmont, MA, USA: Athena Scientific, 2nd ed., 2003.
[25] G. J. Foschini and Z. Miljanic, "A simple distributed autonomous power control algorithm and its convergence," *IEEE Trans. Veh. Technol.*, vol. 42, no. 4, pp. 641–646, 1993.
[26] S. Boyd, A. Ghosh, and D. Shah, "Randomized gossip algorithms," *IEEE Trans. Inf. Theory*, vol. 52, no. 6, pp. 2508–2530, 2006.
[27] S. Friedland and S. Karlin, "Some inequalities for the spectral radius of non-negative matrices and applications," *Duke Mathematical Journal*, vol. 42, no. 3, pp. 459–490, 1975.



Liang Zheng received the Bachelor degree in Software Engineering from Sichuan University, Chengdu, China, in 2011. She is pursuing her Ph.D. degree at City University of Hong Kong.

Her research interests are in wireless networks, mobile computing, nonlinear optimization and its applications.



Chee Wei Tan (M '08, SM '12) received the M.A. and Ph.D. degree in electrical engineering from Princeton University, Princeton, NJ, in 2006 and 2008, respectively.

He is an Assistant Professor at City University of Hong Kong. Previously, he was a Postdoctoral Scholar at the California Institute of Technology (Caltech), Pasadena, CA. He was a Visiting Faculty at Qualcomm R&D, San Diego, CA, in 2011. His research interests are in wireless and broadband communications, signal processing and nonlinear optimization.

Dr. Tan currently serves as an Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS. He was the recipient of the 2008 Princeton University Wu Prize for Excellence and the 2011 IEEE Communications Society AP Outstanding Young Researcher Award.