Beamforming Duality and Algorithms for Weighted Sum Rate Maximization in Cognitive Radio Networks

I-Wei Lai, *Member, IEEE*, Liang Zheng, *Student Member, IEEE*,
Chia-Han Lee, *Member, IEEE*, and Chee Wei Tan, *Senior Member, IEEE*

Abstract

In this paper, we investigate the joint design of transmit beamforming and power control to maximize the weighted sum rate in the multiple-input single-output (MISO) cognitive radio network constrained by arbitrary power budgets and interference temperatures. The nonnegativity of the physical quantities, e.g., channel parameters, powers, and rates, is exploited to enable key tools in nonnegative matrix theory, such as the (linear and nonlinear) Perron-Frobenius theory, quasi-invertibility, and Friedland-Karlin inequalities, to tackle this nonconvex problem. Under certain (quasi-invertibility) sufficient condition, we propose a tight convex relaxation technique that relaxes multiple constraints to bound the global optimal value in a systematic way. Then, a single-input multiple-output (SIMO)-MISO duality is established through a virtual dual SIMO network and Lagrange duality. This SIMO-MISO duality is equivalent to the zero Lagrange duality gap condition that connects the optimality conditions of the primal MISO network and the virtual dual SIMO network. Moreover, by exploiting the SIMO-MISO duality, an algorithm is developed to solve the sum rate maximization problem optimally. Numerical examples demonstrate the computational efficiency of our algorithm when the number of transmit antennas is large.

Index Terms

Optimization, convex relaxation, cognitive radio network, nonnegative matrix theory, quasi-invertibility, Karush-Kuhn-Tucker conditions, Perron-Frobenius theorem.

I. Introduction

Cognitive radio has been widely regarded as a key technology to enhance the spectrum utilization in wireless communications [1]–[5]. While a primary user has exclusive privilege to access its licensed channel, a cognitive radio (secondary user) is capable of dynamically adjusting its parameters to make best use of the available spectrum without interfering with the primary users. Formed by primary and secondary users, cognitive radio networks generally apply the interference temperature constraints to guarantee the quality of service (QoS) of the primary users [6], [7], regardless of the number of secondary users. The single-input single-output (SISO) cognitive radio network duality for utility maximization has been investigated in [8], but how to realize dynamic spectrum or power allocation to maximize the total throughput or sum rate in response to various QoS demands remains an open issue. In this paper, we focus on the joint design of transmit beamforming and power control, in order to maximize the weighted sum rate in the multiple-input single-output (MISO) cognitive radio network.

The problem of sum rate maximization for the cognitive radio network is relatively new and challenging. The sum rate maximization for the secondary users under non-interfering channels is studied in [9], but this problem becomes nonconvex under interference channels. A message-passing algorithm is proposed to suboptimally solve such nonconvex problem in [10]. The work in [11] characterizes the solution for a special case, i.e., two-user sum rate maximization with equal weights, which is then proved to be NP-hard [12]. The standard Lagrange dual relaxation is used to tackle the nonconvexity [12]–[14]. However, due to the positive duality gap, the optimality of dual algorithms cannot be guaranteed in general. By using geometric programming, successive convex approximation algorithms have been developed [15]–[17]. Nevertheless, these works only solve the problem suboptimally.

Recently, global optimization algorithms have been proposed in the literature, e.g., algorithms using difference of convex function programming [18], [19]. In [20], [21], this power-domain optimization problem is reformulated to an equivalent problem in the signal-to-interference-plus-noise ratio (SINR) domain. Based on this reformulation, this problem then becomes a convex maximization problem with an unbounded convex constraint set that can be solved optimally by an outer approximation algorithm [20] or a branch-and-bound method [21]. However, the unbounded constraint set in the SINR domain has the drawback of numerical artifacts. The authors in [22] derive sufficient conditions to optimally solve the problem with only a single total power constraint. Their results are then used to provide useful performance bounds for the problem with individual power constraints [23]. In [24], an equivalent problem in the rate domain is formulated to allow convex relaxation to be used to bound the optimal value, and a branch-and-bound algorithm is developed to approach the optimal solution. Although many researches have been devoted to the weighted sum rate maximization, most existing algorithms have serious drawbacks from a theoretical viewpoint as they cannot guarantee the optimality of numerical solutions. For example, algorithms based on Lagrange duality suffer from the positive duality gap. Additionally, previously-proposed global optimization algorithms generally have exponential-time complexity in the worst case.

In this paper, we investigate the joint design of transmit beamforming and power control to maximize the weighted sum rate in MISO cognitive radio networks. We propose a novel convex relaxation technique that transforms multiple power and/or interference temperature constraints to a single power constraint. As proved in [24], the sum rate maximization can be convexified if certain sufficient conditions can be induced, i.e., all the constraints satisfy the *quasi-invertibility* of certain nonnegative matrices. Since our relaxation reduces the number of constraints, this sufficient condition is easier to satisfy.

Then, the single-input multiple-output (SIMO)-MISO duality is developed by linking a virtually constructed dual SIMO network to the primal MISO network using the Perron-Frobenius theorem. Conventionally, the uplink-downlink duality is constructed through the standard Lagrange duality, which cannot be applied in our work due to the positive duality gap resulted from the nonconvexity of the power-domain sum rate problem. Instead, we establish the proposed SIMO-MISO duality by virtually creating a dual SIMO network with the reversed link directions compared to the primal MISO network. Then, by appropriately setting the noise, coefficients, and upper bounds for the constraints in the virtual dual SIMO network, we prove the identical optimal solutions and values of the virtual dual SIMO network to that of the primal MISO network through the Perron-Frobenius theory. The SIMO-MISO duality is proved to be equivalent to a Lagrange duality with zero duality gap, which is further utilized to prove the tightness of the proposed relaxation technique and the optimality of the linear minimum mean square error (LMMSE) transmit beamforming. Furthermore, this duality allows the representation of the optimality condition with respect to the optimal powers in the primal MISO network and the optimal powers in the virtual dual SIMO network. Based on this representation of the optimality condition, an algorithm for the joint design is developed which comprises a fixed-point algorithm and a subgradient projection method. Numerical examples demonstrate that our algorithm is computationally efficient, especially when the number of transmit antennas is large.

We refer the readers to Fig. 1 for an overview of the connections among the optimization problems and algorithms investigated in this work. The key contributions of this paper are listed as follows:

- A new convex relaxation to solve the weighted sum rate maximization with arbitrary power and interference temperature constraint set in MISO cognitive radio networks. This convex relaxation is different from and provably tighter than that in a related work in [24].
- The SIMO-MISO duality and its equivalence to a Lagrange duality with zero duality gap.
- The optimal algorithm for the joint design of the transmit beamforming and power control which, as a byproduct for the SISO case, improves the algorithm proposed in [24] (rectify a limit cycle condition due to fixed-point iteration).

The paper is organized as follows. In Section II, we describe the MISO cognitive radio network and the weighted sum rate maximization. We propose the convex relaxation technique in Section III, derive the virtual dual SIMO network, and prove the tightness of the proposed relaxation technique by using the

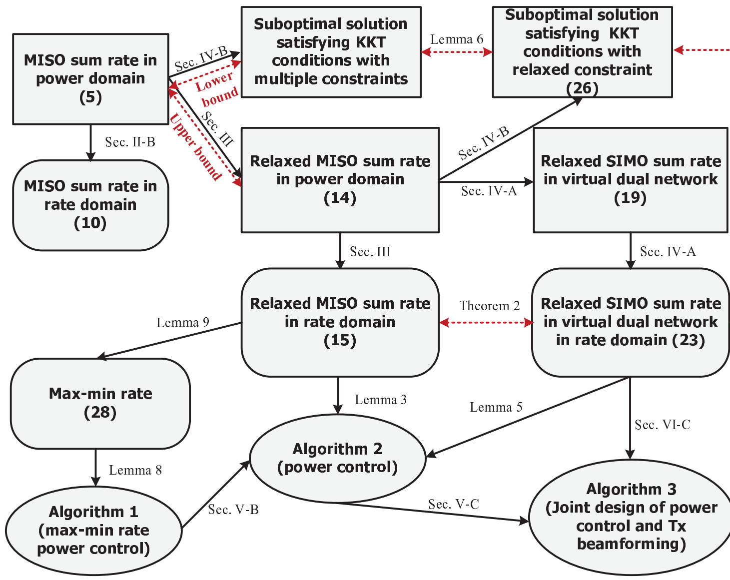
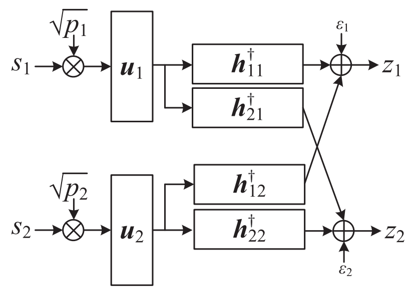


Fig. 1. Overview of the connections among the optimization problems and algorithms investigated in this work. The rectangular boxes are for nonconvex problems, rectangular boxes with rounded corners are for convex problems, and the ellipses are for algorithms. The red dashed lines indicate the connections studied in Section IV-B.

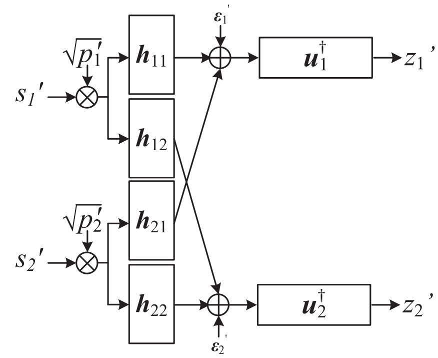
SIMO-MISO duality in Section IV. We then propose the algorithm for the joint design of the transmit beamforming and power control in Section V. In Section VI, we provide numerical examples to evaluate the performance of the proposed algorithms. Finally, conclusions are drawn in Section VII.

Definition

The following notations are used in this paper: Let $\rho(\mathbf{A})$ denote the spectral radius function which outputs the Perron-Frobenius eigenvalue of a nonnegative matrix $\mathbf{A}$. The Perron right and left eigenvectors of $\mathbf{A}$ are respectively defined as $\mathbf{x}(\mathbf{A})$ and $\mathbf{y}(\mathbf{A})$ associated with $\rho(\mathbf{A})$. The superscripts $(\cdot)^*$, $(\cdot)^\top$, and $(\cdot)^\dagger$ respectively denote the conjugate, transpose, and Hermitian operations, $\mathbf{E}[\cdot]$ indicates the expectation operation, $|\cdot|$ and $\|\cdot\|_2$ respectively represent the absolute value and l^2 norm, $\mathbf{I}$ denotes the identity matrix, and $\mathbf{1} = [1, \ldots, 1]^\top \in \mathbb{R}^L$. For vectors $\mathbf{x} = [x_1, \ldots, x_L]^\top \in \mathbb{C}^L$ and $\mathbf{y} = [y_1, \ldots, y_L]^\top \in \mathbb{C}^L$, let $(\mathbf{x})_l = x_l$ be the operation of extracting the lth entry of $\mathbf{x}$. Define $\mathrm{diag}(\mathbf{x})$ as a diagonal matrix with diagonal entries being $[x_1, \ldots, x_L]$. Let $e^{\mathbf{x}} = [e^{x_1}, \ldots, e^{x_L}]^\top$, $\log \mathbf{x} = [\log x_1, \ldots, \log x_L]^\top$, $\mathbf{x}/\mathbf{y} = [x_1/y_1, \ldots, x_L/y_L]^\top$, $\sqrt{\mathbf{x}} = [\sqrt{x_1}, \ldots, \sqrt{x_L}]^\top$, and $\mathbf{x} \circ \mathbf{y}$ be the Schur product of $\mathbf{x}$ and $\mathbf{y}$, i.e., $\mathbf{x} \circ \mathbf{y} = [x_1 y_1, \ldots, x_L y_L]^\top$. Define $\mathbf{x} \geq \mathbf{y}$ as the component-wise inequality between those two vectors, and $\mathbb{R}_{\geq 0}$ as the set of nonnegative real values.



(a) The primal MISO network



(b) The virtual dual SIMO network

Fig. 2. The cognitive radio network for a two-user case.

II. MISO System Model and Optimization

Consider a cognitive radio network with L users sharing a common frequency band and time slot, with every transmitter and receiver pair being a MISO system. As shown in Fig. 2(a), let us define the channel matrix from the ith transmitter to the lth receiver as $\mathbf{h}_{li} \in \mathbb{C}^{N_{\mathrm{T}}}$, where N_{T} is the number of transmit antennas, and the normalized transmit beamforming vectors as $\mathbf{u}_l \in \mathbb{C}^{N_{\mathrm{T}}}$ with $\|\mathbf{u}_l\|_2 = 1$, for $l = 1, \ldots, L$. The received symbol z_l takes the form of

$$z_l = \sum_{i=1}^{L} \mathbf{h}_{li}^{\dagger} \mathbf{u}_i \sqrt{p_i} s_i + \varepsilon_l, \; l = 1, \ldots, L, \tag{1}$$

where p_i and s_i, respectively, denote the transmit power and the transmit symbol of the ith transmitter. The complex additive noise ε_l is assumed to be zero mean with covariance matrix $\mathbf{E}[\varepsilon_l^* \varepsilon_l] = n_l$, where n_l is the noise variance at the lth receiver.

A. System Formulation

Define $\mathbf{z} = [z_1, \ldots, z_L]^\top$, $\mathbf{p} = [p_1, \ldots, p_L]^\top$, $\mathbf{s} = [s_1, \ldots, s_L]^\top$, and $\boldsymbol{\varepsilon} = [\varepsilon_1, \ldots, \varepsilon_L]^\top$. The matrix-vector form of (1) is given by

$$\mathbf{z} = \begin{bmatrix} \mathbf{h}_{11}^{\dagger}\mathbf{u}_1 & \ldots & \mathbf{h}_{1L}^{\dagger}\mathbf{u}_L \\ \vdots & \ddots & \vdots \\ \mathbf{h}_{L1}^{\dagger}\mathbf{u}_1 & \ldots & \mathbf{h}_{LL}^{\dagger}\mathbf{u}_L \end{bmatrix} \mathrm{diag}(\sqrt{\mathbf{p}})\mathbf{s} + \boldsymbol{\varepsilon}. \tag{2}$$

Now, let $\mathbf{U} = [\mathbf{u}_1, \ldots, \mathbf{u}_L]$ and $\mathbf{G}(\mathbf{U}) \in \mathbb{C}^{L\times L}$ be the effective channel gain matrix whose (l, i)th element $G_{li}(\mathbf{U})$ indicates the effective gain between the ith transmitter and the lth receiver, i.e.,

$$G_{li}(\mathbf{U}) = |\mathbf{h}_{li}^{\dagger}\mathbf{u}_i|^2. \tag{3}$$

The SINR of the lth transmitter-receiver pair is then given by

$$\gamma_l = \frac{G_{ll}(\mathbf{U})p_l}{\sum_{i\neq l} G_{li}(\mathbf{U})p_i + n_l} = \frac{p_l}{q_l(\mathbf{p}, \mathbf{U})}, \tag{4}$$

where we denote $q_l(\mathbf{p},\mathbf{U})$ as the *interference temperature*, i.e., the normalized interference plus noise, given as

$$q_l(\mathbf{p},\mathbf{U}) = \begin{cases} \dfrac{\sum_{i\neq l} G_{li}(\mathbf{U})p_i + n_l}{G_{ll}(\mathbf{U})}, & \text{if } p_l \neq 0, \\ 0, & \text{if } p_l = 0. \end{cases}$$

This interference temperature is used to model regulatory constraints for interference management imposed on secondary users in the cognitive radio network. In particular, as the QoS of the primary users in the cognitive radio network has to be guaranteed, the interference level they suffer should be confined. Note that when the lth user is inactive, i.e., $p_l = 0$, its interference temperature should be set to zero, since the lth user is not transmitting in the network. With the definition of SINR in (4), the achievable rate of the lth user is given by $r_l = \log(1+\gamma_l)$.

B. Problem Formulation

The MISO weighted sum rate maximization with a weight vector $\mathbf{w}$ is formulated as follows:

$$\begin{array}{ll} \underset{\mathbf{p}\geq\mathbf{0},\ \|\mathbf{u}_l\|_2=\mathbf{1},\ l=1,\ldots,L}{\text{maximize}} & \mathbf{w}^\top\mathbf{r} = \displaystyle\sum_{l=1}^{L} w_l \log\left(1+\frac{p_l}{q_l(\mathbf{p},\mathbf{U})}\right) \\ \text{subject to} & \mathbf{a}_k^\top\mathbf{p} \leq \bar{p}_k,\ k=1,\ldots,K, \\ & \mathbf{b}_m^\top\mathbf{q}(\mathbf{p},\mathbf{U}) \leq \bar{q}_m\ m=1,\ldots,M, \end{array} \tag{5}$$

where $\mathbf{q}(\mathbf{p},\mathbf{U}) = [q_1(\mathbf{p},\mathbf{U}),\ldots,q_L(\mathbf{p},\mathbf{U})]^\top$. K general weighted power constraints are imposed with nonnegative weights $\mathbf{a}_k = [a_{k1},\ldots,a_{kL}]^\top \in \mathbb{R}^L_{\geq 0}$ and upper bounds $\bar{p}_k$, $k=1,\ldots,K$. Likewise, M interference temperature constraints have the nonnegative coefficients $\mathbf{b}_m \in \mathbb{R}^L_{\geq 0}$ and upper bounds $\bar{q}_m$, $m = 1,\ldots,M$. To find the optimal values and optimal solutions in (5), we first investigate the MISO weighted sum rate maximization for cognitive radio networks with fixed transmit beamforming, and then design the transmit beamforming in Sec. IV-C. Specifically, by utilizing the proposed SIMO-MISO duality which will be expounded later, we can derive the optimal transmit beamforming which is a function of optimal power in the dual SIMO network. The MISO weighted sum rate maximization with fixed transmit beamforming can be regarded as the SISO weighted sum rate maximization and has been studied in [24]. Due to the fixed transmit beamforming and for notation brevity, we respectively simplify $\mathbf{G}(\mathbf{U})$ and $\mathbf{q}(\mathbf{p},\mathbf{U})$ to $\mathbf{G}$ and $\mathbf{q}$.

It should be emphasized that we consider a general formulation that cover different cases of cognitive radio networks. The general formulation of the constraint set is useful to impose the weighted power and interference temperature constraints on all the users by setting the weight appropriately, e.g., giving more weights to assigning higher priority to primary users' share of power resources. Considering an example of S primary users and the individual interference temperature constraints, i.e., $b_{mm} = 1$ and $b_{ml} = 0$

for $l \neq m$, we can set $\bar{q}_m$ to some finite value for a primary user (when $m \leq S$), while we can set $\bar{q}_m$ to infinity for a secondary user (when $m > S$). To illustrate this, let us consider a simple case of two users, namely primary user 1 and secondary user 2 ($L = 2$) with the interference temperatures respectively given by $q_1 = \frac{G_{12}}{G_{11}} p_2 + \frac{n_1}{G_{11}}$ and $q_2 = \frac{G_{21}}{G_{22}} p_1 + \frac{n_2}{G_{22}}$. According to our formulation, we set $\mathbf{b}_1 = [1, 0]^\top$, $\mathbf{b}_2 = [0, 1]^\top$, $\bar{q}_1$ to some finite value, and $\bar{q}_2$ to infinity such that the interference to the secondary user, i.e., q_2, is unconstrained. Since we have $\frac{G_{12}}{G_{11}} p_2 + \frac{n_1}{G_{11}} \leq \bar{q}_1$, we know that the power of the secondary user p_2 is also constrained, while the power of the primary user is unconstrained. Thus, the primary user is protected from the excessive interference from the secondary user.

Let us define the vector $\boldsymbol{\beta}$ and the matrix $\mathbf{F}$ with β_l and F_{li} being their lth and (l, i)th elements respectively given by $\beta_l = \frac{1}{G_{ll}}$ and $F_{ll} = 1/G_{ll}$ and $F_{li} = 0$ for $l \neq i$. Now, the interference temperature can be represented by $\mathbf{q} = \mathrm{diag}(\boldsymbol{\beta})(\mathbf{F}\mathbf{p} + \mathbf{n})$, where $\mathbf{n} = [n_1, \ldots, n_L]^\top$. By using the relationship between $\mathbf{p}$ and $\mathbf{q}$ in (4), we have the following expression for the power vector $\mathbf{p} = \mathrm{diag}(\boldsymbol{\gamma})\mathbf{q} = \mathrm{diag}(\boldsymbol{\gamma} \circ \boldsymbol{\beta})(\mathbf{F}\mathbf{p} + \mathbf{n})$, where $\boldsymbol{\gamma} = [\gamma_1, \ldots, \gamma_L]^\top$. By expressing the SINR $\boldsymbol{\gamma}$ in terms of the rate $\mathbf{r} = [r_1, \ldots, r_L]^\top$, $\mathbf{p}$ can be written as a function of the rate vector:

$$\mathbf{p} = \mathrm{diag}((e^{\mathbf{r}} - \mathbf{1}) \circ \boldsymbol{\beta})(\mathbf{F}\mathbf{p} + \mathbf{n}). \tag{6}$$

Likewise, the expression of the interference temperature can be represented by $\mathbf{q} = \mathrm{diag}(\boldsymbol{\beta})\big(\mathbf{F}\mathrm{diag}(e^{\mathbf{r}} - \mathbf{1})\mathbf{q} + \mathbf{n}\big)$. With the above representations, we can formulate the constraints in (5) in a matrix-vector form and replace the variable $\mathbf{p}$ by the rate $\mathbf{r}$ and interference temperature $\mathbf{q}$.

Theorem 1: The optimal value in (5) is equal to the optimal value of the following problem [24]

$$\begin{aligned} \underset{\mathbf{r} \geq \mathbf{0}, \mathbf{q} \geq \mathbf{0}}{\text{maximize}} \quad & \mathbf{w}^\top \mathbf{r} \\ \text{subject to} \quad & \mathbf{B}_k \mathrm{diag}(e^{\mathbf{r}})\mathbf{q} \leq (\mathbf{B}_k + \mathbf{I})\mathbf{q}, \;\; k = 1, \ldots, K, \\ & \mathbf{D}_m \mathrm{diag}(e^{\mathbf{r}})\mathbf{q} \leq (\mathbf{D}_m + \mathbf{I})\mathbf{q}, \;\; m = 1, \ldots, M \end{aligned} \tag{7}$$

with the nonnegative matrices $\mathbf{B}_k = \mathrm{diag}(\boldsymbol{\beta}) \left(\mathbf{F} + \frac{1}{\bar{p}_k} \mathbf{n} \mathbf{a}_k^\top \right)$ for $k = 1, \ldots, K$, and the nonnegative matrices $\mathbf{D}_m = \left(\mathbf{I} + \frac{1}{\bar{q} - \mathbf{b}_m^\top \mathrm{diag}(\boldsymbol{\beta})\mathbf{n}} \mathrm{diag}(\boldsymbol{\beta}) \mathbf{n} \mathbf{b}_m^\top \right) \mathrm{diag}(\boldsymbol{\beta})\mathbf{F}$ for $m = 1, \ldots, M$. The proof is given in Appendix A.

The Perron-Frobenius theorem allows us to further reformulate the constraints in (7) in the form of spectral radius functions, given the existence of the *quasi-invertibility* of certain nonnegative matrix.

Definition 1 (Quasi-Invertibility [25]): A *quasi-inverse* matrix $\tilde{\mathbf{A}}$ of a square nonnegative matrix $\mathbf{A}$ is a square nonnegative matrix satisfying $\mathbf{A} - \tilde{\mathbf{A}} = \mathbf{A}\tilde{\mathbf{A}} = \tilde{\mathbf{A}}\mathbf{A}$.

With Definition 1, the quasi-inverse matrices of $\mathbf{B}_k$ and $\mathbf{D}_m$ exist if

$$\tilde{\mathbf{B}}_k = (\mathbf{I} + \mathbf{B}_k)^{-1}\mathbf{B}_k \geq \mathbf{0}, \quad k = 1, \ldots, K, \tag{8}$$

$$\tilde{\mathbf{D}}_m = (\mathbf{I} + \mathbf{D}_m)^{-1}\mathbf{D}_m \geq \mathbf{0}, \quad m = 1, \ldots, M. \tag{9}$$

The following lemma shows that the quasi-invertibility is true when the effective cross-channel gain is sufficiently weak.

Lemma 1: If $\mathbf{F} = 0$, we have $\mathbf{B}_k = \mathrm{diag}(\boldsymbol{\beta})\frac{\mathbf{n}\mathbf{a}_k^\top}{\bar{p}_k}$. Then, we always have the quasi-inverse matrix $\tilde{\mathbf{B}}_k \geq 0$ with $\tilde{\mathbf{B}}_k = \frac{1}{\bar{p}_k + \mathbf{a}_k^\top \mathrm{diag}(\boldsymbol{\beta})\mathbf{n}}\mathrm{diag}(\boldsymbol{\beta})\mathbf{n}\mathbf{a}_k^\top$. The proof is given in Appendix B.

It should be emphasized that while Lemma 1 theoretically shows that the quasi-invertibility is true under the condition of weak effective cross-channel gains, this quasi-invertibility is also true when $\bar{\mathbf{q}}$ is large [24] or when $\bar{\mathbf{p}}$ is small [22], which respectively represent the practical cases of large interference level among users and the practical cases of low SNR.

The following result shows that when the quasi-invertibility conditions in (8) and (9) are true, the optimization problem (7) can be convexified.

Corollary 1: When (8) and (9) are true, the sum rate maximization problem in (7) is equivalent to the following problem whose constraints are represented by the spectral radius functions [24]

$$\begin{aligned} \underset{\mathbf{r} \geq \mathbf{0}}{\text{maximize}} \quad & \mathbf{w}^\top \mathbf{r} \\ \text{subject to} \quad & \log \rho(\tilde{\mathbf{B}}_k \mathrm{diag}(e^{\mathbf{r}})) \leq 0, \; k = 1, \ldots, K, \\ & \log \rho(\tilde{\mathbf{D}}_m \mathrm{diag}(e^{\mathbf{r}})) \leq 0, \; m = 1, \ldots, M, \end{aligned} \tag{10}$$

where $\mathbf{r}$ is the only remaining optimization variable. Since $\tilde{\mathbf{B}}_k$ and $\tilde{\mathbf{D}}_m$ are irreducible nonnegative matrices, the corresponding constraints in (10) are convex in $\mathbf{r}$ due to the *log-convexity property* of the spectral radius function [26]. That is, (10) is a convex optimization problem in the rate domain. The proof is given in Appendix C.

By applying the nonnegative matrix theory, the optimality conditions of (10) are given below.

Lemma 2: The optimal rate $\mathbf{r}^\star$ in (10) and the associated optimal power $\mathbf{p}^\star$ satisfy the following optimality conditions [24]:

$$\rho(\boldsymbol{\Omega}\mathrm{diag}(e^{\mathbf{r}^\star})) = 1, \qquad \mathbf{p}^\star = \mathrm{diag}(e^{\mathbf{r}^\star} - \mathbf{1})\mathbf{x}(\boldsymbol{\Omega}\mathrm{diag}(e^{\mathbf{r}^\star})) \tag{11}$$

with $\boldsymbol{\Omega} = \mathrm{argmax}_{\boldsymbol{\Omega}' = \{\tilde{\mathbf{B}}_{k^\star}, \tilde{\mathbf{D}}_{m^\star}\}} \rho(\boldsymbol{\Omega}'\mathrm{diag}(e^{r^\star}))$, where $k^\star = \mathrm{argmax}_{k=1,\ldots,K}\, \rho(\tilde{\mathbf{B}}_k \mathrm{diag}(e^{\mathbf{r}^\star}))$ and $m^\star = \mathrm{argmax}_{m=1,\ldots,M}\, \rho(\tilde{\mathbf{D}}_m \mathrm{diag}(e^{\mathbf{r}^\star}))$.

Note that (11) is true up to a scaling constant so that the optimal power in (11) should be normalized appropriately depending on the constraint. For instance, given that $\boldsymbol{\Omega} = \tilde{\mathbf{B}}_{k^\star}$, which implies the tightness of the $k^\star$th power constraint, the optimal power vector is normalized as $\mathbf{p}^\star \leftarrow \frac{\bar{p}_{k^\star}}{\mathbf{a}_{k^\star}^\top \mathbf{p}^\star}\mathbf{p}^\star$. The proof is given in Appendix D.

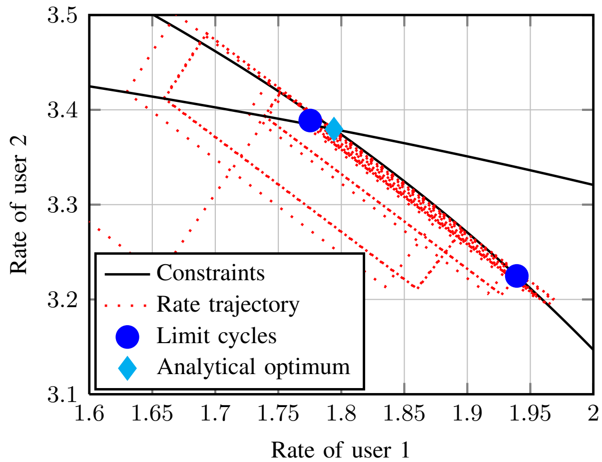


Fig. 3. Example of the rate trajectory of the polynomial-time algorithm in [24] for a two-user case. $\mathbf{a}_1 = [0.3131,\ 0.7725]$, $\bar{p}_1 = 4.8840$, $\mathbf{a}_2 = [0.9458,\ 0.3979]$, and $\bar{p}_2 = 3.8981$.

III. Convex Relaxation for Multiple-Constrained Weighted Sum Rate Maximization

The review in the previous section shows that the SISO sum rate maximization can be convexified when all the reformulated spectral radius constraints possess the quasi-invertibility condition. A polynomial-time algorithm has been proposed in [24] to solve such special cases of the sum rate maximization, but it is under the implicit assumption of a single tight constraint at optimality. If multiple constraints are tight, we illustrate that this polynomial-time algorithm [24] may result in limit cycle oscillations as shown by the rate trajectory depicted in Fig. 3 for a two-user case. This is because, for multiple tight constraints, the optimum may happen at the crosspoints of the constraints as shown in Fig. 3. However, since this polynomial-time algorithm implicitly assumes only a single bounding constraint, its bounding constraint alternates with each iteration. Due to this phenomenon, the rates may oscillate in limit cycles near the optimum and fail to converge. This issue is resolved in this section by the proposed relaxation technique. Moreover, our technique can be regarded as a novel convex relaxation technique since it convexifies the sum rate maximization when some of the constraints do not satisfy the quasi-invertibility sufficient condition.

The proposed convex relaxation technique starts with reformulating the interference temperature constraints to the power constraints

$$\mathbf{b}_m^\top \mathrm{diag}(\boldsymbol{\beta})(\mathbf{F}\mathbf{p} + \mathbf{n}) \leq \bar{q}_m, \Rightarrow (\mathbf{F}^\top \mathrm{diag}(\boldsymbol{\beta})\mathbf{b}_m)^\top \mathbf{p} \leq \bar{q}_m - \mathbf{b}_m^\top \mathrm{diag}(\boldsymbol{\beta})\mathbf{n}, \tag{12}$$

and then the multiple constraints in (5) can be condensed into a single power constraint:

$$\mathbf{a}_\theta^\top \mathbf{p} \leq \bar{p}_\theta, \tag{13}$$

$$\mathbf{a}_\theta = \sum_{k=1}^{K} \mathbf{a}_k \theta_k + \sum_{m=1}^{M} \left(\mathbf{F}^\top \mathrm{diag}(\boldsymbol{\beta})\mathbf{b}_m\right) \theta_{m+K},$$

$$\bar{p}_\theta = \sum_{k=1}^{K} \bar{p}_k \theta_k + \sum_{m=1}^{M} \left(\bar{q}_m - \mathbf{b}_m^\top \mathrm{diag}(\boldsymbol{\beta})\mathbf{n} \right) \theta_{m+K},$$

where $\boldsymbol{\theta} = [\theta_1, \dots, \theta_{K+M}] \in \mathbb{R}_{\geq 0}^{K+M}$ is the convex combination weight vector of individual constraints satisfying $\boldsymbol{\theta}^\top \mathbf{1} = 1$. Any feasible point in the intersection of all the constraints in (5) is also a feasible point in the unified constraint $\mathbf{a}_\theta^\top \mathbf{p} \leq \bar{p}_\theta$ so that (13) is a relaxed constraint. This modulation by $\boldsymbol{\theta}$ is exploited as a relaxation technique with associated relaxed weighted sum rate maximization given below:

$$\begin{array}{ll} \underset{\mathbf{p} \geq \mathbf{0}}{\text{maximize}} & \mathbf{w}^\top \mathbf{r} \\ \text{subject to} & \mathbf{a}_\theta^\top \mathbf{p} \leq \bar{p}_\theta. \end{array} \tag{14}$$

Following the approach of the optimization reformulation in Section II, when the associated nonnegative quasi-invertible matrix of the relaxed single constraint exists, i.e., $\tilde{\mathbf{B}}_\theta = (\mathbf{I} + \mathbf{B}_\theta)^{-1}\mathbf{B}_\theta \geq 0$ with the definition of $\mathbf{B}_\theta = \mathrm{diag}(\boldsymbol{\beta})(\mathbf{F} + \frac{1}{\bar{p}_\theta}\mathbf{n}\mathbf{a}_\theta^\top)$, the problem in (14) can be convexified as, owing to the log-convexity property of the spectral radius function [26],

$$\begin{array}{ll} \underset{\mathbf{r} \geq \mathbf{0}}{\text{maximize}} & \mathbf{w}^\top \mathbf{r} \\ \text{subject to} & \log \rho(\tilde{\mathbf{B}}_\theta \mathrm{diag}(e^{\mathbf{r}})) \leq 0. \end{array} \tag{15}$$

Remark 1: The problem in (15) is convex if $\tilde{\mathbf{B}}_\theta \geq 0$, but the multiple-constrained sum rate maximization (7) can be convexified only if $\tilde{\mathbf{B}}_k \geq 0$, $k = 1, \dots K$, and $\tilde{\mathbf{D}}_m \geq 0$, $m = 1, \dots M$. This implies that the relaxed single-constrained sum rate maximization (14) is more likely to be convexified than the multi-constrained problem in (7), meaning that the proposed technique is a convex relaxation.

To elaborate the statement in Remark 1, we use a two-user cognitive radio network with individual power constraints as an example to demonstrate the effectiveness of the proposed relaxation technique. For a 2×2 matrix $\mathbf{Z}$ with z_{ij} being its (i,j)th element, we can derive the closed-form expression of $(\mathbf{I} + \mathbf{Z})^{-1}\mathbf{Z}$ as

$$(\mathbf{I} + \mathbf{Z})^{-1}\mathbf{Z} = \frac{1}{\det(\mathbf{I} + \mathbf{Z})} \begin{bmatrix} z_{11} + z_{11}z_{22} - z_{12}z_{21} & z_{12} \\ z_{21} & z_{22} + z_{11}z_{22} - z_{12}z_{21} \end{bmatrix}, \tag{16}$$

where $\det(\cdot)$ is the determinant operation. For illustration purpose, let us assume $\mathbf{a}_i = [1,\ 0]^\top$, $\bar{p}_i = 1$, $i = 1,\ 2$, $\boldsymbol{\beta} = \mathbf{1}$, and $\mathbf{n} = \mathbf{1}$ for the sum rate maximization with individual power constraints. We have $\mathbf{B}_1 = [1, G_{12}\ ;\ G_{21} + 1, 0]$ and

$$(\mathbf{I} + \mathbf{B}_1)^{-1}\mathbf{B}_1 = \frac{1}{\det(\mathbf{I} + \mathbf{B}_1)} \begin{bmatrix} 1 - G_{12}(G_{21} + 1) & G_{12} \\ G_{21} + 1 & -G_{12}(G_{21} + 1) \end{bmatrix} \tag{17}$$

by using (16). Apparently, the matrix in (17) is not nonnegative. However, by applying the relaxation technique with $\boldsymbol{\theta} = [\theta_1,\ \theta_2]^\top$, we have $\mathbf{B}_\theta = [\theta_1, G_{12} + \theta_2\ ;\ G_{21} + \theta_1, \theta_2]$ and

$$(\mathbf{I} + \mathbf{B}_\theta)^{-1}\mathbf{B}_\theta = \frac{1}{\det(\mathbf{I} + \mathbf{B}_\theta)} \begin{bmatrix} \theta_1 - G_{12}\theta_1 - G_{21}\theta_2 - G_{12}G_{21} & G_{12} + \theta_2 \\ G_{21} + \theta_1 & \theta_2 - G_{12}\theta_1 - G_{21}\theta_2 - G_{12}G_{21} \end{bmatrix}.$$

In this case, a nonnegative $\tilde{\mathbf{B}}_\theta$ may exist, i.e., $\tilde{\mathbf{B}}_\theta = (\mathbf{I} + \mathbf{B}_\theta)^{-1}\mathbf{B}_\theta \geq 0$. For example, given that the equal weight $\boldsymbol{\theta} = [0.5,\ 0.5]^\top$ is used, we have $\tilde{\mathbf{B}}_\theta \geq 0$ when $G_{12} + G_{21} + 2G_{12}G_{21} \leq 1$, implying the existence of the quasi-invertibility sufficient condition when the cross-channel gain is sufficiently small.

Last, the following result states the quasiconvexity of the relaxed sum rate maximization with respect to the relaxation modulation parameter $\boldsymbol{\theta}$.

Lemma 3: Given $\boldsymbol{\theta}_1,\ \boldsymbol{\theta}_2 \in \mathbb{R}^{K+M}_{\geq 0}$ and $\lambda \in \mathbb{R}_{\geq 0},\ \lambda \leq 1$, the optimal value $g(\boldsymbol{\theta}) = \mathbf{w}^\top \mathbf{r}(\theta)$ is quasiconvex, i.e., $g(\lambda\boldsymbol{\theta}_1 + (1-\lambda)\boldsymbol{\theta}_2) \leq \max\big(g(\boldsymbol{\theta}_1), g(\boldsymbol{\theta}_2)\big)$. The proof is given in Appendix E.

With the quasiconvexity property, the optimal modulation value $\boldsymbol{\theta}^\star$ can be obtained by classical approaches such as the subgradient projection method [27], [28].

IV. Virtual Dual SIMO Network

In this section, we study the sum rate maximization in the virtual dual SIMO cognitive radio network. Under the sufficient condition of $\tilde{\mathbf{B}}_\theta \geq 0$, these two optimization problems (in the rate domain) lead to identical optimal values and optimal solutions, which can be interpreted as the *SIMO-MISO duality*. This duality is equivalent to a Lagrange duality with zero duality gap, which can be used to prove the tightness of the proposed convex relaxation technique in (15).

A. Virtual Dual SIMO Network by the Perron-Frobenius Theorem

As shown in Fig. 2(b), the virtual dual SIMO network comprises the channels which are the transpose of the channels in the primal MISO network

dual receive beamforming vectors. After the receive beamforming, the receive symbol z'_l polluted by the noise vector $\boldsymbol{\varepsilon}'_l \in \mathbb{C}^L$ is given by

$$z'_l = \mathbf{u}_l^\dagger \sum_{i=1}^{L} \mathbf{h}_{il} \sqrt{p'_i} s'_i + \mathbf{u}_l^\dagger \boldsymbol{\varepsilon}'_l,\ l = 1, \ldots, L,$$

where p'_i and s'_i are respectively the transmit power and transmit symbol of the ith user in the virtual dual SIMO network. With the definition of $\mathbf{G}$ in (3), the SINR of the virtual dual SIMO network takes the form of

$$\gamma'_l = \frac{G_{ll} p'_l}{\sum_{i \neq l} G_{il} p'_i + n'_l} = \frac{|\mathbf{h}_{ll}^\dagger \mathbf{u}_l|^2 p'_l}{\sum_{i \neq l} |\mathbf{h}_{il}^\dagger \mathbf{u}_l|^2 p'_i + n'_l}, \tag{18}$$

where $\mathbf{E}[|\mathbf{u}_l^\top \boldsymbol{\varepsilon}'_l|^2] = n'_l$ is due to $\|\mathbf{u}_l\|_2 = 1$ and the covariance matrix $\mathbf{E}[\boldsymbol{\varepsilon}'_l \boldsymbol{\varepsilon}'^\dagger_l] = n'_l \mathbf{I}$ of the zero-mean independent and identically distributed (i.i.d.) Gaussian vector $\boldsymbol{\varepsilon}'_l$.

Based on this virtual dual SIMO network, we study a power-domain weighted sum rate maximization problem with single constraint that has the same weight vector $\mathbf{w}$ as that in the primal MISO network:

$$\begin{aligned} \underset{\mathbf{p}' \geq \mathbf{0}}{\text{maximize}} \quad & \mathbf{w}^\top \mathbf{r} \\ \text{subject to} \quad & \mathbf{a}'^\top \mathbf{p}' \leq \bar{p}'. \end{aligned} \tag{19}$$

Let us consider the following nonnegative matrix, which is similar to the role of the nonnegative matrix $\mathbf{B}_\theta$:

$$\mathbf{C} = \mathrm{diag}(\boldsymbol{\beta}) \left(\mathbf{F}^\top + \frac{1}{\bar{p}'} \mathbf{n}' \mathbf{a}'^\top \right). \tag{20}$$

By the association between the problem parameters in the virtual dual SIMO network and the relaxed sum rate maximization (14) in the primal MISO network, i.e.,

$$(\mathbf{a}', \mathbf{n}', \bar{\mathbf{p}}') = (\mathbf{n}, \mathbf{a}_\theta, \bar{\mathbf{p}}_\theta), \tag{21}$$

we can represent $\mathbf{C}$ in (20) as

$$\mathbf{C} = \mathrm{diag}(\boldsymbol{\beta}) \left(\mathbf{F}^\top + \frac{1}{\bar{p}_\theta} \mathbf{a}_\theta \mathbf{n}^\top \right). \tag{22}$$

Using this new representation, the sum rate maximization in the virtual dual SIMO network can be convexified and proved to be equivalent to the relaxed sum rate maximization (15), if $\tilde{\mathbf{B}}_\theta \geq 0$, as stated in the following theorem.

Theorem 2: If $\tilde{\mathbf{B}}_\theta \geq 0$, the optimization problem in (19) with the parameter relationship (21) is equivalent to

$$\begin{array}{ll} \underset{\mathbf{r} \geq \mathbf{0}}{\text{maximize}} & \mathbf{w}^\top \mathbf{r} \\ \text{subject to} & \log \rho(\tilde{\mathbf{C}} \mathrm{diag}(e^{\mathbf{r}})) \leq 0, \end{array} \tag{23}$$

where $\tilde{\mathbf{C}} = (\mathbf{C} + \mathbf{I})^{-1} \mathbf{C}$ is the quasi-inverse matrix of $\mathbf{C}$ in (22). Due to the log-convexity of the spectral radius function, the problem (23) is convex with its optimality condition given by

$$\rho(\tilde{\mathbf{C}} \mathrm{diag}(e^{\mathbf{r}^\star})) = \rho(\tilde{\mathbf{B}}_\theta \mathrm{diag}(e^{\mathbf{r}^\star})) = 1,$$

and has the same optimal value and optimal solutions (rates) as that in (15). The proof is given in Appendix F.

The optimal power vector of (19) is given by $\mathbf{p}'^\star = \mathrm{diag}(e^{\mathbf{r}^\star} - \mathbf{1}) \mathbf{x}(\tilde{\mathbf{C}} \mathrm{diag}(e^{\mathbf{r}^\star}))$ by the same approach as in Lemma 2, which notably can be regarded as a function of the Perron right eigenvector in the virtual dual SIMO network. The following Lemma gives an alternative representation of $\mathbf{p}'$.

Lemma 4: The optimal power vector in the virtual dual SIMO network is a function of the Perron left eigenvector in the primal MISO network, i.e.,

$$\mathbf{p}'^\star = \mathrm{diag}(\boldsymbol{\beta} \circ (\mathbf{1} - e^{-\mathbf{r}^\star})) \mathbf{y}(\tilde{\mathbf{B}}_\theta \mathrm{diag}(e^{\mathbf{r}^\star})), \tag{24}$$

which can be obtained as a by-product from solving (15). The proof is given in Appendix G.

With the aid of Lemma 4, the optimality condition can be represented by the Schur product of the optimal powers, respectively, in the primal MISO network and the virtual dual SIMO network, as shown in the following lemma.

Lemma 5: In addition to Lemma 2, an alternative optimality condition can be obtained that is represented by the Schur product of the optimal powers in the primal MISO and virtual dual SIMO networks:

$$\frac{1}{\lambda^\star}(\mathbf{w}+\boldsymbol{\mu}^\star)=\mathrm{diag}(\boldsymbol{\beta}\circ(e^{\mathbf{r}^\star}-2\cdot\mathbf{1}+e^{-\mathbf{r}^\star}))^{-1}\mathbf{p}^\star\circ\mathbf{p}'^\star, \tag{25}$$

where $\boldsymbol{\mu}^\star \in \mathbb{R}^L_{\geq 0}$ and $\lambda^\star > 0$ are the Lagrange dual variables associated with the nonnegative constraint $\mathbf{r} \geq \mathbf{0}$ and the spectral radius constraint of (23), respectively. The proof is given in Appendix H. As will be expounded later, this alternative representation of the optimality condition facilitates the design of the distributed power control algorithm.

B. Proof of Tight Relaxation

To prove that the relaxation in (14) is tight and equivalent to solving (5) when $\tilde{\mathbf{B}}_\theta \geq 0$, we first prove the optimality of the solution obtained by applying the Karush-Kuhn-Tucker (KKT) conditions to the nonconvex problem (14).

Theorem 3: For the nonconvex problem (14), the local optimal solution satisfying the KKT conditions is given by

$$r^\star_\iota=\log\left(1+\frac{p'^\star_\iota}{\left(\mathrm{diag}(\boldsymbol{\beta})(\mathbf{F}^\top\mathbf{p}'^\star+\mathbf{a}_\theta)\right)_\iota}\right),\qquad \iota=1,\ldots,L, \tag{26}$$

with the conditions $\mathbf{p}'^\star=\left[\frac{(w_1+\mu_1^\star)\beta_1(e^{r_1^\star}-1)}{\lambda^\star\left(\mathrm{diag}(\boldsymbol{\beta})(\mathbf{F}\mathbf{p}^\star+\mathbf{n})\right)_1+\lambda p_1^\star},\ldots,\frac{(w_L+\mu_L^\star)\beta_L(e^{r_L^\star}-1)}{\lambda\left(\mathrm{diag}(\boldsymbol{\beta})(\mathbf{F}\mathbf{p}^\star+\mathbf{n})\right)_L+\lambda^\star p_L^\star}\right]^\top$, $\mathbf{n}^\top\mathbf{p}'^\star=\bar{p}_\theta$, and $\frac{1}{\lambda^\star}(\mathbf{w}+\boldsymbol{\mu}^\star)=\mathrm{diag}(\boldsymbol{\beta}\circ(e^{\mathbf{r}^\star}-2\cdot\mathbf{1}+e^{-\mathbf{r}^\star}))^{-1}\mathbf{p}^\star\circ\mathbf{p}'^\star$. The proof is given in Appendix I. As can be seen, if $\tilde{\mathbf{B}}_\theta \geq 0$, the local optimal rate (26) is represented by the SINR in the virtual dual SIMO network, indicating the equivalence between this local optimal solution of (14) in the primal network and the optimal solution of (23) in the virtual dual SIMO network. Together with the SIMO-MISO duality given in Theorem 2, the local optimal rates (26) of the relaxed sum rate maximization problem with single constraint of the primal MISO network is proven to be optimal.

Next, we establish the following relationship between (26) and the suboptimal solution of (5) that satisfies the KKT conditions.

Lemma 6: For the sum rate maximization (5), its local optimal solutions obtained by the KKT conditions are identical to (26) when $\boldsymbol{\theta}^\star=\frac{1}{\lambda^\star}\boldsymbol{\varrho}^\star$, where $\boldsymbol{\varrho}\in\mathbb{R}^{M+K}_{\geq 0}$ and $\lambda > 0$ are, respectively, the Lagrange dual variables of the constraint set in (5) and the constraint $\mathbf{a}_\theta^\top\mathbf{p}\leq\bar{p}_\theta$ in (14). The proof is given in Appendix J.

Since the optimal solution of the relaxation (14) and the suboptimal solution of (5) are respectively the upper and lower bounds of the optimal value of (5), and by using Lemma 6 and Theorem 3, we can prove that the optimal solution of the relaxation (14) and the suboptimal solution of (5) are both identical to (26). Therefore, the tightness of the relaxation in (14) is validated. The readers may refer to Fig. 1, where the connections of the problems studied in this subsection are marked by red dashed lines.

C. Optimal Transmit Beamforming

By exploiting the SIMO-MISO duality derived in Theorem 2, the optimal transmit beamforming in the primal MISO network is shown to be the same as the optimal receive beamforming in the virtual dual SIMO network. The SIMO receive beamforming can be readily designed, since each lth receive beamforming can be optimized in a decoupled manner. Specifically, while the SINR (4) in the primal MISO network is a function of multiple transmit beamforming vectors, the SINR (18) in the virtual dual SIMO network only depends on $\mathbf{u}_l$. Thus, the optimization of the receive beamforming is decoupled from one another. The lth receive beamforming is optimized simply by maximizing its rate r_l, which is equivalent to the maximization of SINR γ_l' and the minimization of the mean-squared error (MSE), leading to the LMMSE beamforming vector as the optimal solution, i.e.,

$$\mathbf{u}_l^{\dagger\star} = \frac{1}{\|\mathbf{h}_{ll}^{\dagger}\big(\mathbf{H}_l \mathrm{diag}(\mathbf{p}'^{\star})\mathbf{H}_l^{\dagger} + n_l'\mathbf{I}\big)^{-1}\|_2}\mathbf{h}_{ll}^{\dagger}\big(\mathbf{H}_l \mathrm{diag}(\mathbf{p}'^{\star})\mathbf{H}_l^{\dagger} + n_l'\mathbf{I}\big)^{-1}, \tag{27}$$

where $\mathbf{H}_l \in \mathbb{C}^{L\times N_{\mathrm{T}}}$ is constituted by the column vectors $\mathbf{h}_{il}$ for $i = 1, \ldots, L$.

V. Joint Transmit Beamforming and Power Control Algorithm

In this section, we consider the joint design of transmit beamforming and power control. We first establish the link between the max-min rate problem and the relaxed sum rate maximization problem with single constraint. Then, by using the optimality condition (25) and the subgradient projection method, an algorithm is developed to solve (5).

A. Weighted Max-Min Rate Optimization

Let us consider the weighted max-min rate problem given by

$$\begin{array}{ll} \underset{\mathbf{r}\geq\mathbf{0}}{\text{maximize}} & \underset{l=1,\ldots,L}{\min} \dfrac{r_l}{\tilde{w}_l} \\ \text{subject to} & \log\rho(\tilde{\boldsymbol{B}}_{\theta}\mathrm{diag}(e^{\mathbf{r}})) \leq 0. \end{array} \tag{28}$$

The positive weight vector $\tilde{\mathbf{w}} = [\tilde{w}_1, \ldots, \tilde{w}_L]^{\top}$ in (28) reflects the priority of users, i.e., a larger $\tilde{w}_l$ means a higher priority of the lth user. The following result gives the optimal value and optimal solution of the weighted max-min rate problem in (28).

Lemma 7: The optimal solution $\mathbf{r}^{\star}$ of the max-min rate optimization (28) is proportional to the weight vector $\tilde{\mathbf{w}}$ [24]. The proof is given in Appendix K.

The following iterative algorithm computes the optimal solution of (28).

Algorithm 1 (Weighted Max-Min Rate Power Control):

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

① Compute $\mathbf{r}(t)$ by $\mathbf{p}(t)$:

$$r_l(t) = \log\left(1 + \frac{p_l(t)}{(\mathrm{diag}(\boldsymbol{\beta})(\mathbf{F}\mathbf{p}(t)+\mathbf{n}))_l}\right),\ l = 1,\ldots,L.$$

② Compute $\mathbf{p}(t+1)$:

$$p_l(t+1) = \left(\frac{\tilde{w}_l}{r_l(t)}\right) p_l(t),\ l = 1,\ldots,L.$$

③ Normalize $\mathbf{p}(t+1) \leftarrow \frac{\mathbf{p}(t+1)}{\mathbf{a}_\theta^\top \mathbf{p}(t+1)/\bar{p}_\theta}$.

④ Go to Step 1 until convergence.

Lemma 8: Starting from any initial point $\mathbf{p}(0)$, $\mathbf{r}(t)$ converges geometrically fast to the optimal solution of the max-min rate optimization in (28) [24]. The proof is given in Appendix L.

Remark 2: Observe that Step 2 in Algorithm 1 is similar to the distributed power control (DPC) algorithm in [29] that is widely used for power control in wireless cellular networks. In particular, as each user in the cognitive radio network can estimate its SINR γ_l and then compute its achievable rate, the power can be updated distributively in every iteration. The normalization in Step 3 can be performed either centrally or distributively by each user using the gossip algorithm [30].

In the following lemma, the connection between the weighted sum rate maximization problem and the weighted max-min rate problem with single constraint (14) is established.

Lemma 9: Conditioned on $\tilde{\mathbf{w}} = \mathbf{1}$ and $\mathbf{w} = \mathbf{x}(\tilde{\boldsymbol{B}}_\theta) \circ \mathbf{y}(\tilde{\boldsymbol{B}}_\theta)$, the optimal values and solutions of the max-min rate (28) and the weighted sum rate maximization (10) are the same, i.e., $\mathbf{r}^\star = -\log(\rho(\tilde{\boldsymbol{B}}_\theta))\mathbf{1}$ [24]. This can be proved by using th contraction mapping [31], or the Krause's theorem [32], as shown in Appendix M.

By exploiting Lemma 9, a fast and distributed power control algorithm to solve the sum rate maximization optimally is developed in the next subsection.

B. Power Control Algorithm of Weighted Sum Rate Maximization

Using Algorithm 1 as a computation submodule (which connects with the sum rate maximization through Lemma 9) and Lemma 5, Algorithm 2 is proposed for the power control of the weighted sum rate maximization. (Note that in Algorithm 2 the iteration index t_k represents the t_kth iteration of the kth loop of the algorithm, starting from the innermost loop for the execution of Algorithm 1.)

Algorithm 2 (Weighted Sum Rate Power Control):

Initialize $\mathbf{r}(0) \geq \mathbf{0}$ and $\boldsymbol{\theta}(0) \in \mathbb{R}^{K+M}_{\geq 0}$, $\boldsymbol{\theta}^\top \mathbf{1} = 1$.

① Compute the relaxed single power constraint modulated by $\boldsymbol{\theta}$:

$$\mathbf{a}_\theta(t_3) = \sum_{k=1}^{K} \mathbf{a}_k^\top \theta_k(t_3) + \sum_{m=1}^{M} \left(\mathbf{F}^\top \mathrm{diag}(\boldsymbol{\beta})\mathbf{b}_m\right)^\top \theta_{m+K}(t_3),$$

$$\bar{p}_\theta(t_3) = \sum_{k=1}^{K} \bar{p}_k \theta_k(t_3) + \sum_{m=1}^{M} \left(\bar{q}_m - \mathbf{b}_m^\top \mathrm{diag}(\boldsymbol{\beta})\mathbf{n}\right) \theta_{m+K}(t_3).$$

② Use Algorithm 1 to, respectively, find the primal power $\mathbf{p}(t_1+1)$ and the dual power $\mathbf{p}'(t_1+1)$ with $\mathbf{a}_\theta(t_3)$, $\bar{p}_\theta(t_3)$, and $\tilde{\mathbf{w}} = \mathbf{r}(t_3)$. When $\mathbf{p}(t_1)$, $\mathbf{p}'(t_1)$, and $\mathbf{r}(t_1)$ converge, set $\mathbf{p}(t_2+1) = \mathbf{p}(t_1)$, $\mathbf{p}'(t_2+1) = \mathbf{p}'(t_1)$, and $\mathbf{r}(t_2+1) = \mathbf{r}(t_1)$.

③ Update $\mathbf{r}(t_2+1)$ as follows:

$$\hat{w}_l = \frac{p_l(t_2)p_l'(t_2)}{\beta_l(e^{r_l(t_2)} + e^{-r_l(t_2)} - 2)}, \qquad l = 1, \dots, L,$$

$$\hat{\mathbf{w}} \leftarrow \frac{\hat{\mathbf{w}}}{\hat{\mathbf{w}}^\top \mathbf{1}},$$

$$r_l(t_2+1) \leftarrow \max\left\{r_l(t_2) + (w_l - \hat{w}_l), \epsilon\right\}, \qquad l = 1, \dots, L,$$

where ϵ is a given small positive scalar.

④ When $\mathbf{r}(t_2)$ converges, if $w_l < \hat{w}_l$, force $r_l = 0$, i.e., remove the lth user from the optimization, and go back to Step 2; otherwise, go to Step 5 and set $\mathbf{r}(t_3+1) = \mathbf{r}(t_2)$.

⑤ Update $\boldsymbol{\theta}(t_3+1)$ by

$$\boldsymbol{\theta}(t_3+1) = \max\left\{\boldsymbol{\theta}(t_3) - \alpha \nabla \mathbf{g}(\boldsymbol{\theta}(t_3)), 0\right\},$$

where $\nabla \mathbf{g}(\boldsymbol{\theta}(t_3)) = \left[\bar{p}_1 - \mathbf{a}_1^\top \mathbf{p}(t_3), \dots, \bar{p}_K - \mathbf{a}_1^\top \mathbf{p}(t_3), \bar{q}_1 - \mathbf{b}_1^\top \mathbf{q}(t_3), \dots, \bar{q}_M - \mathbf{b}_M^\top \mathbf{q}(t_3)\right]$, and the normalization $\boldsymbol{\theta}(t_3+1) \leftarrow \frac{\boldsymbol{\theta}(t_3+1)}{\mathbf{1}^\top \boldsymbol{\theta}(t_3+1)}$.

⑥ Go to Step 1 until convergence.

Remark 3: Algorithm 2 comprises three loops: The innermost loop computes the powers in the primal and the powers in virtual dual networks by using Algorithm 1. According to Lemma 9, the relaxed single power constraint tightens when Algorithm 1 converges. The second loop (Step 3 and 4) then updates the rates according to Lemma 5. Since $\mathbf{p}$ and $\mathbf{p}'$ are derived from the Perron right and left eigenvectors of the nonnegative matrix $\tilde{\mathbf{B}}_\theta \mathrm{diag}(e^{\mathbf{r}^\star})$ as shown in Appendix H, the optimality of the rate convergence can be proved by using Lemma 4 in [24]. For numerical stability, we set a small positive scalar $\epsilon > 0$; otherwise, a zero rate leads to the zero denominator in the second step of Algorithm 1. Then, since zero rates indicate inactive users, we design a mechanism to identify such inactive users and remove them from the optimization. This mechanism is realized by checking the Lagrange dual variable μ_l. Specifically, whenever the lth user is inactive at the optimum, we have $\mu_l > 0$, which is checked by the condition of $w_l < \hat{w}_l$ as stated in (25). Lastly, the outermost loop is used to tighten the relaxation by applying the subgradient projection method since the relaxation is proved to be quasiconvex by Lemma 3. Given the reciprocal channel in a time division duplex (TDD) system, the computation of the powers in either the primal MISO network or the dual SIMO network can be computed distributively. No channel information exchange is required, and each user only changes the computed power for the rate computation and the normalization steps, where the latter can be realized by the gossip algorithm [30]. In this case, Algorithm 2 for the power control with fixed transmit beamforming can be implemented in a distributed manner. However, when channels of the MISO and SIMO networks are not reciprocal or when we have a non-TDD system, a central controller can be used to compute the power in the dual SIMO network.

C. Joint Beamforming and Power Control Algorithm of Weighted Sum Rate Maximization

As shown in (27), the optimal LMMSE transmit beamforming is a function of the powers in the virtual dual SIMO network. The joint design algorithm is thus realized by simply adding an additional loop on top of Algorithm 2.

Algorithm 3 (Weighted Sum Rate Joint Beamforming and Power Control):

Initialize $\mathbf{r}(0) > 0$ and $\mathbf{U}(0)$ with $\|\mathbf{u}_l(0)\|_2 = 1,\ l = 1,\ \ldots,\ L$.

① Use Algorithm 2 to compute $\mathbf{p}(t_3)$, $\mathbf{p}'(t_3)$, and $\mathbf{r}(t_3)$. When $\mathbf{p}(t_3)$, $\mathbf{p}'(t_3)$, and $\mathbf{r}(t_3)$ converge, set $\mathbf{p}(t_4+1) = \mathbf{p}(t_3)$, $\mathbf{p}'(t_4+1) = \mathbf{p}'(t_3)$, and $\mathbf{r}(t_4+1) = \mathbf{r}(t_3)$.

② Compute the transmit beamforming vector

$$\mathbf{u}_l^\top(t_4+1) = \mathbf{h}_l^\dagger(\mathbf{H}^\dagger \mathrm{diag}(\mathbf{p}'(t_4))\mathbf{H} + n_l'\mathbf{I})^{-1},\ l = 1, \ldots,\ L,$$

$$\mathbf{u}_l(t+1) \leftarrow \frac{1}{\|\mathbf{u}_l(t_4+1)\|_2}\mathbf{u}_l(t_4+1),\ l = 1, \ldots, L.$$

③ Go to Step 1 until convergence.

Remark 4: Since the powers in the dual network are already computed by Algorithm 2, the computation for updating the beamformers has little overhead. The optimum transmit beamforming and the power control is iteratively approached by using Step 1 for the power control and Step 2 for the transmit beamforming [22]. Since the complete channel information is required for the computation of transmit beamforming, a central controller can be used to reduce the overhead of information exchange.

VI. Numerical Examples

In this section, we first numerically evaluate the performance of the proposed algorithm under the sufficient conditions of the quasi-invertibility of $\mathbf{B}_\theta$. Then, the percentage of the sum rate maximization which can be convexified by the proposed tight convex relaxation technique is numerically examined with various number of users, transmit antennas, and constraints. The normalized noise power is adopted

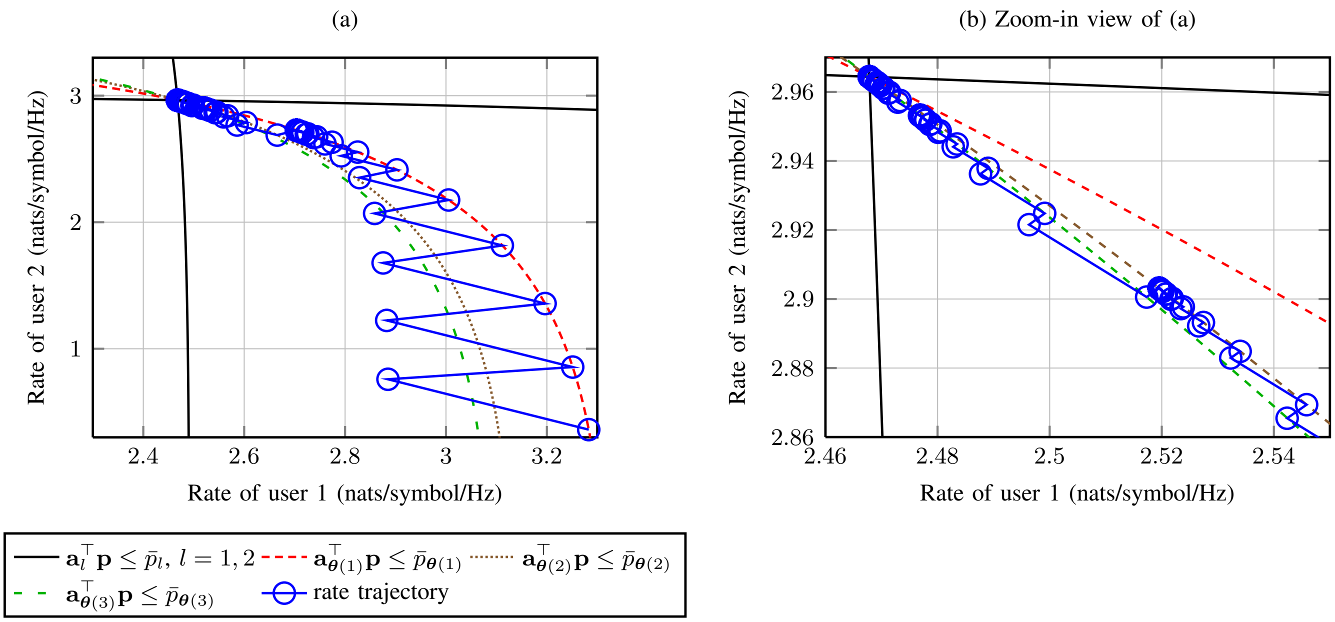


Fig. 4. Rate trajectory of Algorithm 2 for a two-user case with individual power constraints.

for simplicity, i.e., $\mathbf{n} = \mathbf{1}$. Note that for illustration purpose, we omit to show the evaluation in the innermost loop when Algorithm 1 is executed. In other words, each line depicted in the figures is the result of $\mathbf{r}(t_2)$. All the channel gains are randomly generated by a complex Gaussian random variable with zero mean and unit variance. The coefficients and the upper bounds of the constraints are randomly generated which follow the uniform distribution.

We first demonstrate that the proposed relaxation technique can convexify the sum rate maximization problem with multiple constraints and provide the optimal solutions. From the rate iteration illustrated in Fig. 4 that applies Algorithm 2 to a two-user case with individual power constraints as discussed in Sec. III, we observe that the relaxed single constraint becomes tighter to the individual power constraints with the update of $\boldsymbol{\theta}$. The rates eventually converge at the crosspoint of the two individual power constraints. This numerical experiment shows that the proposed relaxation technique can convexify the problem even when none of its constraints satisfy the quasi-invertibility.

Next, we illustrate the performance of Algorithm 2 and 3 for a five-user case with user 1 and user 2 being the primary users. We set $b_{11} = 1$ and $b_{22} = 1$, and generate the rest coefficients of the interference temperature constraints by using a uniform random variable with support $[0 0.1]$. Two power constraints and two interference temperature constraints are given as follows: $\mathbf{a}_1 = [0.2050, 0.6999, 0.1432, 0.4997, 0.4265]^\top$, $\mathbf{a}_2 = [0.6162, 0.6604, 0.7775, 0.5115, 0.1894]^\top$, $\mathbf{b}_1 = [1, 0.0140, 0.0793, 0.0081, 0.0368]^\top$, $\mathbf{b}_2 = [0.0864, 1, 0.0232, 0.0380, 0.0258]^\top$, $\bar{\mathbf{p}} = [4.3703, 3.6434]^\top$, and $\bar{\mathbf{q}} = [1.9239, 2.9822]^\top$. The transmit

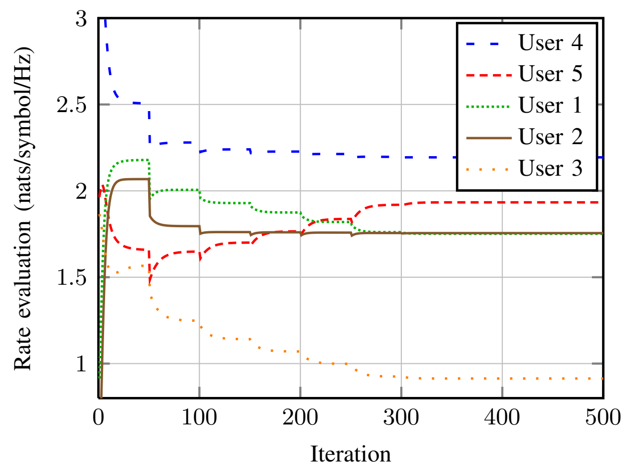


Fig. 5. Rate evaluation for a fixed transmit beamforming vectors. 50 rate iterations, 10 iterations for the subgradient projection method, five users, and three constraints are used.

beamforming is fixed and the effective channel gain is

$$\mathbf{G} = \begin{bmatrix} 5.0416 & 0.0970 & 0.0514 & 0.0958 & 0.0421 \\ 0.1851 & 3.5891 & 0.0232 & 0.0261 & 0.0501 \\ 0.0925 & 0.0165 & 3.9248 & 0.1137 & 0.0085 \\ 0.0386 & 0.0417 & 0.0924 & 7.8120 & 0.0970 \\ 0.2279 & 0.1147 & 0.3580 & 0.0638 & 4.6734 \end{bmatrix}.$$

With this $\mathbf{G}$ and the constraints mentioned above, Fig. 5 illustrates the rate iteration of Algorithm 2. The iteration numbers of the three loops in Algorithm 2 are respectively 50, 50, and 10, starting from the innermost loop. The results depicted in Fig. 5 show that the convergence of each loop is fast. Due to the update of $\boldsymbol{\theta}$, a jitter occurs at every 50 iterations.

The performance of Algorithm 3 is examined by the evaluation of the optimal value depicted in Fig. 6, where each curve represents the rate evaluation at the tth transmit beamforming iteration, that is, the rate evaluation of Algorithm 2 with the fixed transmit beamforming $\mathbf{U}(t)$. Again, each jitter represents the update of the subgradient projection for $\boldsymbol{\theta}$. Since the relaxation tightens with each iteration, the optimal value eventually converges to the optimal value of the original sum rate maximization with multiple constraints. From the outermost iteration in Algorithm 3, the optimized transmit beamforming improves the optimal values.

Lastly, we numerically examine the percentage of the sum rate maximization problem that can be convexified, i.e., the percentage that $\tilde{\mathbf{B}}_\theta \geq 0$, with various system parameters $(N_{\mathrm{T}}, L, K+M)$. 1000 channel realizations for the channel vector $\mathbf{h}_{li}^\dagger$, $i = 1, \ldots, L$ and $l = 1, \ldots, L$ are generated. Fig. 7(a) and Fig. 7(b) respectively show the two-user and four-user cases. The percentage of the convexity is less than 10% for the SISO network, but quickly increases when multiple transmit antennas are applied. This is because transmit beamforming can effectively suppress the interference that leads to a weak interference

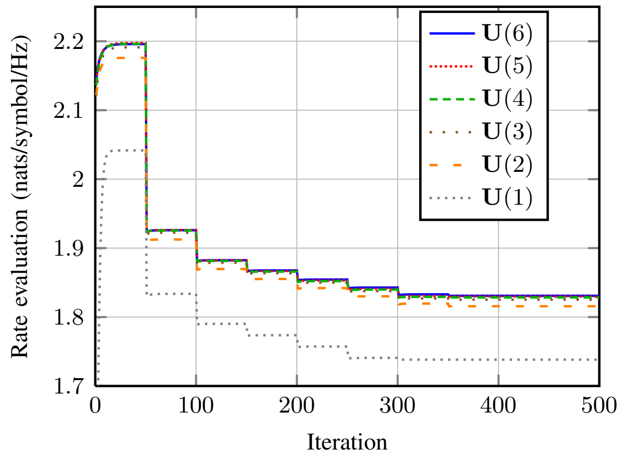


Fig. 6. Optimal value evaluation with various transmit beamforming. 50 rate iterations, 10 iterations for the subgradient projection method, five users, and three constraints are used.

setting favoring the quasi-invertibility condition to occur more frequently, as stated in Lemma 1. This demonstrates that our proposed algorithm is especially useful when the number of transmit antennas is large. For $N_{\mathrm{T}} = 2$, the percentage of the convexity is almost 70% when a single constraint is imposed. For the case with five constraints, however, the percentage of the convexity drops below 50%. After applying the proposed relaxation technique, the percentage of convexity becomes high, validating that the proposed relaxation technique effectively convexifies the nonconvex sum rate maximization problem with multiple constraints. Interestingly, in some circumstances, the likelihood of the relaxation with single constraint being convexified is greater than that of the optimization with a single constraint. Comparing the two-user case with the four-user case shown in Fig. 7(b), the percentage of the convexity decreases when more users operate in the cognitive radio network, since more nonnegative entries in $\tilde{\mathbf{B}}_\theta$ are required for the convexity. Nevertheless, by applying our relaxation technique, the likelihood of the convexity again increases significantly.

VII. Conclusions

In this paper, we have proposed a convex relaxation technique that transforms multiple power and interference temperature constraints into a single relaxed power constraint to bound the global optimal value of sum rate maximization. The virtual dual SIMO network is constructed with the aid of this single-constrained relaxation. The sum rate maximization of the virtual dual SIMO network is then shown to be equivalent to the relaxed sum rate maximization of the primal MISO network by using the Perron-Frobenius theorem. The derived SIMO-MISO duality is exploited to prove the tightness of the proposed relaxation technique. The optimal LMMSE transmit beamforming is also proved by the SIMO-MISO duality. Using the relaxation technique and the optimality condition represented by the Schur product of

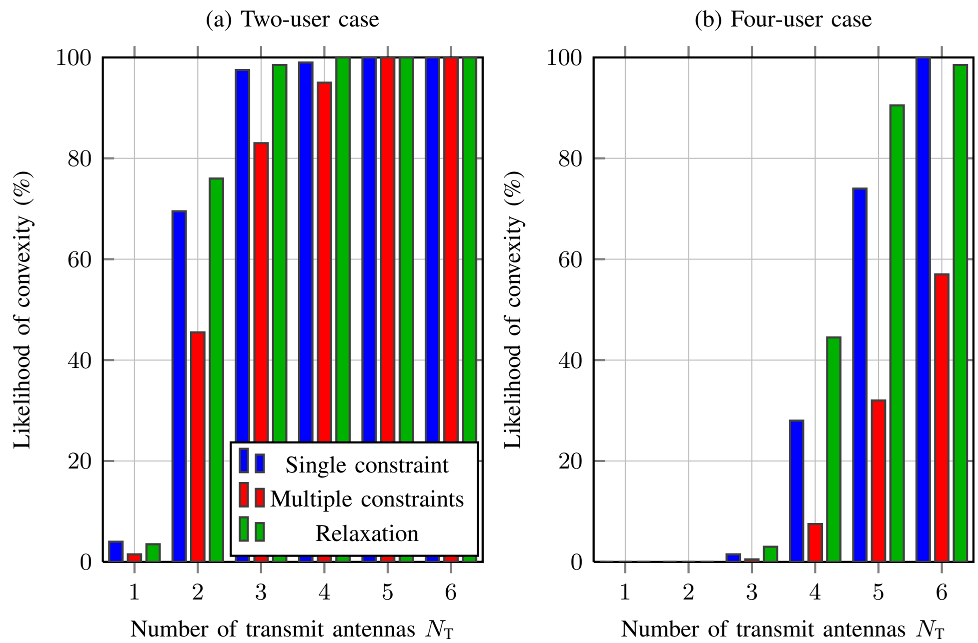


Fig. 7. The likelihood of convexity with various number of transmit antennas, constraints, and users.

the optimal powers in the primal MISO and the optimal powers in the virtual dual SIMO networks, an algorithm that leverages max-min rate control as a submodule is designed to jointly optimize the power control and transmit beamforming. As demonstrated by the numerical examples, the proposed algorithm converges fast and is computationally efficient in MISO cognitive radio networks, especially when the number of transmit antennas is large.

Appendix

A. Proof of Theorem 1

Proof: The kth power constraint can be represented by $\mathbf{B}_k$ as in the following reformulation:

$$\mathrm{diag}(\boldsymbol{\beta})\left(\mathbf{F}+\frac{1}{\bar{p}_k}\mathbf{n}\mathbf{a}_k^\top\right)\mathbf{p}=\mathrm{diag}(\boldsymbol{\beta})\left(\mathbf{F}\mathbf{p}+\frac{1}{\bar{p}_k}\mathbf{n}\mathbf{a}_k^\top\mathbf{p}\right),$$

$$\Rightarrow \mathrm{diag}(\boldsymbol{\beta})\left(\mathbf{F}+\frac{1}{\bar{p}_k}\mathbf{n}\mathbf{a}_k^\top\right)(\mathrm{diag}(e^{\mathbf{r}}-\mathbf{1}))\mathbf{q}\overset{(a)}{\leq}\mathrm{diag}(\boldsymbol{\beta})(\mathbf{F}\mathbf{p}+\mathbf{n})=\mathbf{q},$$

$$\Rightarrow \mathbf{B}_k\mathrm{diag}(e^{\mathbf{r}})\mathbf{q}\leq(\mathbf{I}+\mathbf{B}_k)\mathbf{q},$$

where (a) is due to $\mathbf{a}_k^\top\mathbf{p}\leq\bar{p}_k$. Likewise, the mth interference temperature constraint is reformulated as

$$\mathbf{q}=\mathrm{diag}(\boldsymbol{\beta})\big(\mathbf{F}\mathrm{diag}(e^{\mathbf{r}}-\mathbf{1})\mathbf{q}+\mathbf{n}\big)\geq\mathrm{diag}(\boldsymbol{\beta})\left(\mathbf{F}\mathrm{diag}(e^{\mathbf{r}}-\mathbf{1})\mathbf{q}+\mathbf{n}\frac{\mathbf{b}_m^\top\mathbf{q}}{\bar{q}_m}\right),$$

$$\Rightarrow \mathrm{diag}(\boldsymbol{\beta})\mathbf{F}\mathrm{diag}(e^{\mathbf{r}}-\mathbf{1})\mathbf{q}\leq\left(\mathbf{I}-\frac{1}{\bar{q}_m}\mathrm{diag}(\boldsymbol{\beta})\mathbf{n}\mathbf{b}_m^\top\right)\mathbf{q},$$

$$\overset{(a)}{\Rightarrow}\left(\mathbf{I}+\frac{1}{\bar{q}_m-\mathbf{b}_m^\top\mathrm{diag}(\boldsymbol{\beta})\mathbf{n}}\mathrm{diag}(\boldsymbol{\beta})\mathbf{n}\mathbf{b}_m^\top\right)\mathrm{diag}(\boldsymbol{\beta})\mathbf{F}\mathrm{diag}(\mathbf{e}^{\mathbf{r}}-\mathbf{1})\mathbf{q}\leq\mathbf{q},$$

$$\Rightarrow \mathbf{D}_m\mathrm{diag}(e^{\mathbf{r}})\mathbf{q}\leq(\mathbf{I}+\mathbf{D}_m)\mathbf{q},$$

where (a) is due to $\big(\mathbf{I}-\frac{1}{\bar{q}_m}\mathrm{diag}(\boldsymbol{\beta})\mathbf{n}\mathbf{b}_m^\top\big)^{-1} = \big(\mathbf{I}+\frac{1}{\bar{q}_m-\mathbf{b}_m^\top\mathrm{diag}(\boldsymbol{\beta})\mathbf{n}}\mathrm{diag}(\boldsymbol{\beta})\mathbf{n}\mathbf{b}_m^\top\big)$ by the Sherman-Morrison formula [33]

$$(\mathbf{X}+\mathbf{u}\mathbf{v}^\top)^{-1} = \mathbf{X}^{-1} - \frac{1}{1+\mathbf{v}^\top\mathbf{X}^{-1}\mathbf{u}}\mathbf{X}^{-1}\mathbf{u}\mathbf{v}^\top \tag{29}$$

with $\mathbf{X}=\mathbf{I}$, $\mathbf{u}=-\frac{1}{\bar{q}_m}\mathrm{diag}(\boldsymbol{\beta})\mathbf{n}$, and $\mathbf{v}=\mathbf{b}_m$ in (29). ■

B. Proof of Lemma 1

Proof: Given $\mathbf{B}=\frac{1}{\bar{p}_k}\mathrm{diag}(\boldsymbol{\beta})\mathbf{n}\mathbf{a}_k^\top$, by assuming $\tilde{\mathbf{B}}=\frac{\alpha}{\bar{p}_k}\mathrm{diag}(\boldsymbol{\beta})\mathbf{n}\mathbf{a}_k^\top$ for some positive $0<\alpha<1$, and using $(\mathbf{I}+\frac{1}{\bar{p}_k}\mathrm{diag}(\boldsymbol{\beta})\mathbf{n}\mathbf{a}_k^\top) = (\mathbf{I}-\frac{\alpha}{\bar{p}_k}\mathrm{diag}(\boldsymbol{\beta})\mathbf{n}\mathbf{a}_k^\top)^{-1}$ in Definition 1, we can apply the Sherman-Morrison formula as in (29) to obtain $\mathbf{I}+\frac{1}{\bar{p}_k}\mathrm{diag}(\boldsymbol{\beta})\mathbf{n}\mathbf{a}_k^\top = \mathbf{I}+\frac{1}{\bar{p}_k/\alpha-\mathbf{a}_k^\top\mathrm{diag}(\boldsymbol{\beta})\mathbf{n}}\mathrm{diag}(\boldsymbol{\beta})\mathbf{n}\mathbf{a}_k^\top$ which leads to $\alpha=\frac{1}{1+\mathbf{a}_k^\top\mathrm{diag}(\boldsymbol{\beta})\mathbf{n}/\bar{p}_k}$. ■

C. Proof of Corollary 1

Proof: Due to the quasi-invertibility of $\mathbf{B}_k$, the power constraints in (7) are reformulated to

$$\begin{aligned} &\mathbf{B}_k\mathrm{diag}(e^{\mathbf{r}})\mathbf{q} \le (\mathbf{I}+\mathbf{B}_k)\mathbf{q}, \quad k=1,\dots,K,\\ \Rightarrow\ &\tilde{\mathbf{B}}_k\mathrm{diag}(e^{\mathbf{r}})\mathbf{q} \le \mathbf{q}, \quad k=1,\dots,K, \end{aligned} \tag{30}$$

where $\tilde{\mathbf{B}}_k$ is given in (8). Likewise, the interference temperature constraints in (7) can be rewritten as $\tilde{\mathbf{D}}_m\mathrm{diag}(e^{\mathbf{r}})\mathbf{q}\le\mathbf{q}$ when $\tilde{\mathbf{D}}_m\ge 0$, $m=1,\dots,M$. We now state the Subinvariance Theorem [34]:

Lemma 10 (The Subinvariance Theorem [34]): Define a nonnegative and irreducible matrix $\mathbf{A}>\mathbf{0}$, a positive scalar $\Lambda>0$, and a nonnegative vector $\boldsymbol{\upsilon}\ge\mathbf{0}$ that satisfies $\mathbf{A}\boldsymbol{\upsilon}\le\Lambda\boldsymbol{\upsilon}$. We have the inequality $\rho(\mathbf{A})\le\Lambda$, where the equality $\rho(\mathbf{A})=\Lambda$ holds true if and only if $\mathbf{A}\boldsymbol{\upsilon}=\Lambda\boldsymbol{\upsilon}$.

Now, from Lemma 10 and by letting $\Lambda=1$ and $\mathbf{A}=\tilde{\mathbf{B}}_k\mathrm{diag}(e^{\mathbf{r}})$, the inequalities associated with the power constraints in (30) can be rewritten as $\rho\big(\tilde{\mathbf{B}}_k\mathrm{diag}(e^{\mathbf{r}})\big)\le 1,\quad k=1,\dots,K$. Likewise, by denoting $\Lambda=1$ and $\mathbf{A}=\tilde{\mathbf{D}}_m\mathrm{diag}(e^{\mathbf{r}})$, the interference temperature constraints can also be reformulated as $\rho\big(\tilde{\mathbf{D}}_m\mathrm{diag}(e^{\mathbf{r}})\big)\le 1,\ m=1,\dots,M$. By taking the logarithm, this algebraic manipulation transforms the constraint set in (7) to a convex set of spectral radius constraints owing to the log-convexity property of the spectral radius function [26]. ■

D. Proof of Lemma 2

Proof: The first order derivatives of the logarithmic spectral radius constraints, $\log\rho\big(\tilde{\mathbf{B}}_k\mathrm{diag}(e^{\mathbf{r}})\big)$ and $\log\rho\big(\tilde{\mathbf{D}}_m\mathrm{diag}(e^{\mathbf{r}})\big)$, with respect to the rate $\mathbf{r}$ take the form of [21], [35]

$$\begin{aligned} \frac{\partial\log\rho\big(\tilde{\mathbf{B}}_k\mathrm{diag}(e^{\mathbf{r}})\big)}{\partial\mathbf{r}} &= \mathbf{x}\big(\tilde{\mathbf{B}}_k\mathrm{diag}(e^{\mathbf{r}})\big)\circ\mathbf{y}\big(\tilde{\mathbf{B}}_k\mathrm{diag}(e^{\mathbf{r}})\big),\\ \frac{\partial\log\rho\big(\tilde{\mathbf{D}}_m\mathrm{diag}(e^{\mathbf{r}})\big)}{\partial\mathbf{r}} &= \mathbf{x}\big(\tilde{\mathbf{D}}_m\mathrm{diag}(e^{\mathbf{r}})\big)\circ\mathbf{y}\big(\tilde{\mathbf{D}}_m\mathrm{diag}(e^{\mathbf{r}})\big). \end{aligned} \tag{31}$$

According to the Perron-Frobenius theorem, the Perron right and left eigenvectors of an irreducible nonnegative matrix are nonnegative [34], indicating that the spectral radius constraints in (10) are monotonically increasing functions. The linear objective function $\mathbf{w}^\top\mathbf{r}$ in (10) is monotonically increasing in $\mathbf{r}$ as $\mathbf{w} \in \mathbb{R}_{\geq 0}$. Thus, maximizing the weighted sum rate leads to at least one tight spectral radius constraint. Given the $k^\star$th power constraint being tight, we have $\rho\big(\tilde{\mathbf{B}}_{k^\star}\text{diag}(e^{\mathbf{r}^\star})\big) = 1$ in (10) and $\mathbf{q}^\star = \mathbf{x}\big(\tilde{\mathbf{B}}_{k^\star}\text{diag}(e^{\mathbf{r}^\star})\big)$ according to Lemma 10, and the optimal power is $\mathbf{p}^\star = \text{diag}(e^{\mathbf{r}^\star} - \mathbf{1})\mathbf{x}\big(\tilde{\mathbf{B}}_{k^\star}\text{diag}(e^{\mathbf{r}^\star})\big)$. The proof of the tightness of the $m^\star$th interference temperature constraint is exactly the same. ∎

E. Proof of Lemma 3

Proof: Given $\boldsymbol{\theta}_1,\ \boldsymbol{\theta}_2 \in \mathbb{R}_{\geq 0}^{K+M}$ and $\lambda \in \mathbb{R}_{\geq 0},\ \lambda \leq 1$ and by defining $\mathbf{p}^\star_{\theta_\lambda}$ as the optimal power of the relaxed sum rate maximization with $\boldsymbol{\theta}_\lambda = \lambda\boldsymbol{\theta}_1 + (1-\lambda)\boldsymbol{\theta}_2$, we have $\big(\lambda\mathbf{a}_{\theta_1}^\top + (1-\lambda)\mathbf{a}_{\theta_2}^\top\big)\,\mathbf{p}^\star_{\theta_\lambda} \leq \lambda\bar{p}_{\theta_1} + (1-\lambda)\bar{p}_{\theta_2}$. This inequality implies that at least one of the inequalities, $\mathbf{a}_{\theta_i}^\top \leq \bar{p}_{\theta_i}$ for $i = 1,\ 2$, is true, meaning that $\mathbf{p}^\star_{\theta_\lambda}$ is a feasible power vector for one of the relaxed sum rate maximization with either θ_1 or θ_2. Assume that $\mathbf{p}^\star_{\theta_\lambda}$ is a feasible power vector for the relaxed sum rate maximization with $\boldsymbol{\theta}_1$. Then we obtain $\tilde{\mathbf{p}} = \frac{\bar{p}_{\theta_1}}{\mathbf{a}_{\theta_1}^\top\mathbf{p}^\star_{\theta_\lambda}}\mathbf{p}^\star_{\theta_\lambda}$, which is also a feasible point in the relaxed sum rate maximization with $\boldsymbol{\theta}_1$. Define $\mathbf{r}^\star_{\theta_\lambda}$ and $\tilde{\mathbf{r}}$ as the rates respectively associated with $\mathbf{p}^\star_{\theta_\lambda}$ and $\tilde{\mathbf{p}}$. Since $\mathbf{a}_{\theta_1}^\top\mathbf{p}^\star_{\theta_\lambda} \leq \bar{p}_{\theta_1}$, we have $\mathbf{p}^\star_{\theta_\lambda} \leq \tilde{\mathbf{p}}$ and thus $g(\lambda\boldsymbol{\theta}_1 + (1-\lambda)\boldsymbol{\theta}_2) = \mathbf{w}^\top\mathbf{r}^\star_{\theta_\lambda} \leq \mathbf{w}^\top\tilde{\mathbf{r}}$, where $g(\boldsymbol{\theta}) = \mathbf{w}^\top\mathbf{r}(\boldsymbol{\theta})$. Then, with the definition of $\mathbf{r}^\star_{\theta_1}$ as the optimal rate solution of the relaxed sum rate maximization with $\boldsymbol{\theta}_1$, we reach the conclusion that $g(\lambda\boldsymbol{\theta}_1 + (1-\lambda)\boldsymbol{\theta}_2) \leq \mathbf{w}^\top\tilde{\mathbf{r}} \leq \mathbf{w}^\top\mathbf{r}^\star_{\theta_1} = g(\lambda\boldsymbol{\theta}_1)$. ∎

F. Proof of Theorem 2

Proof: The relationship between $\mathbf{C}$ in (22) and $\mathbf{B}_\theta^\top$ is given by

$$\mathbf{B}_\theta^\top = \left(\mathbf{F}^\top + \frac{1}{\bar{p}_\theta}\mathbf{a}_\theta\mathbf{n}^\top\right)\text{diag}(\boldsymbol{\beta}) = \text{diag}(\boldsymbol{\beta})^{-1}\mathbf{C}\text{diag}(\boldsymbol{\beta}), \tag{32}$$

and the transpose of the quasi-inverse matrix $\tilde{\mathbf{B}}_\theta^\top$ can be reformulated as

$$\tilde{\mathbf{B}}_\theta^\top = \mathbf{B}_\theta^\top\big((\mathbf{I}+\mathbf{B}_\theta)^{-1}\big)^\top = (\mathbf{I}+\mathbf{B}_\theta^\top)^{-1}\mathbf{B}_\theta^\top. \tag{33}$$

Inserting (32) into (33), $\tilde{\mathbf{B}}_\theta^\top$ is represented as

$$\begin{aligned}\tilde{\mathbf{B}}_\theta^\top &= (\mathbf{I} + \text{diag}(\boldsymbol{\beta})^{-1}\mathbf{C}\text{diag}(\boldsymbol{\beta}))^{-1}\text{diag}(\boldsymbol{\beta})^{-1}\mathbf{C}\text{diag}(\boldsymbol{\beta}) = \text{diag}(\boldsymbol{\beta})^{-1}(\mathbf{I}+\mathbf{C})^{-1}\mathbf{C}\text{diag}(\boldsymbol{\beta})\\ &= \text{diag}(\boldsymbol{\beta})^{-1}\tilde{\mathbf{C}}\text{diag}(\boldsymbol{\beta}).\end{aligned} \tag{34}$$

The nonnegativity of $\boldsymbol{\beta}$ indicates that $\tilde{\mathbf{C}} \geq 0$ once $\tilde{\mathbf{B}}_\theta \geq 0$.

Using the relationship between $\tilde{\mathbf{C}}$ and $\tilde{\mathbf{B}}_\theta$ in (34), the optimality condition represented by the spectral radius constraint can be reformulated as follows:

$$\begin{aligned} 1 &= \rho(\tilde{\mathbf{B}}_\theta \mathrm{diag}(e^{\mathbf{r}^\star})) \\ &\overset{(a)}{=} \rho(\tilde{\mathbf{B}}_\theta^\top \mathrm{diag}(e^{\mathbf{r}^\star})) = \rho(\mathrm{diag}(\boldsymbol{\beta})^{-1}\tilde{\mathbf{C}}\mathrm{diag}(\boldsymbol{\beta}\circ e^{\mathbf{r}^\star})) \\ &\overset{(b)}{=} \rho(\mathrm{diag}(\boldsymbol{\beta}\circ e^{\mathbf{r}^\star})\mathrm{diag}(\boldsymbol{\beta})^{-1}\tilde{\mathbf{C}}) \\ &\overset{(c)}{=} \rho(\tilde{\mathbf{C}}\mathrm{diag}(e^{\mathbf{r}^\star})), \end{aligned}$$

where (a) is due to $\rho(\mathbf{XY}) = \rho(\mathbf{YX})$ and $\rho(\mathbf{X}) = \rho(\mathbf{X}^\top)$, (b) is due to $\rho(\mathbf{XY}) = \rho(\mathbf{YX})$, and (c) is due to $\rho(\mathbf{XY}) = \rho(\mathbf{YX})$ and the cancellation of $\mathrm{diag}(\boldsymbol{\beta})$ and $\mathrm{diag}(\boldsymbol{\beta})^{-1}$. ■

G. Proof of Lemma 4

Proof: Following the Subinvariance Theorem in Lemma 10, we have

$$\tilde{\mathbf{C}}\mathrm{diag}(e^{\mathbf{r}^\star})\mathbf{q}' = \mathbf{q}'. \tag{35}$$

Inserting $\tilde{\mathbf{C}} = \mathrm{diag}(\boldsymbol{\beta})\tilde{\mathbf{B}}_\theta^\top\mathrm{diag}(\boldsymbol{\beta})^{-1}$ (from (34)) into (35) gives

$$\begin{aligned} \mathrm{diag}(\boldsymbol{\beta})\tilde{\mathbf{B}}_\theta^\top\mathrm{diag}(e^{\mathbf{r}^\star}/\boldsymbol{\beta})\mathbf{q}' = \mathbf{q}' &\Rightarrow \mathrm{diag}(e^{\mathbf{r}^\star})\tilde{\mathbf{B}}_\theta^\top\left(\mathrm{diag}(e^{\mathbf{r}^\star}/\boldsymbol{\beta})\mathbf{q}'\right) = \mathrm{diag}(e^{\mathbf{r}^\star}/\boldsymbol{\beta})\mathbf{q}', \\ &\Rightarrow \left(\mathrm{diag}(e^{\mathbf{r}^\star}/\boldsymbol{\beta})\mathbf{q}'\right)^\top\left(\tilde{\mathbf{B}}_\theta\mathrm{diag}(e^{\mathbf{r}^\star})\right) = \left(\mathrm{diag}(e^{\mathbf{r}^\star}/\boldsymbol{\beta})\mathbf{q}'\right)^\top, \end{aligned}$$

showing $\mathrm{diag}(e^{\mathbf{r}^\star}/\boldsymbol{\beta})\mathbf{q}'$ is the Perron left eigenvector of the matrix $\tilde{\mathbf{B}}_\theta\mathrm{diag}(e^{\mathbf{r}^\star})$. Then, we have

$$\begin{aligned} \mathrm{diag}(e^{\mathbf{r}^\star}/\boldsymbol{\beta})\mathbf{q}' = \mathbf{y}(\tilde{\mathbf{B}}_\theta\mathrm{diag}(e^{\mathbf{r}^\star})) &\Rightarrow \mathrm{diag}(e^{\mathbf{r}^\star}/(\boldsymbol{\beta}\circ(e^{\mathbf{r}^\star}-\mathbf{1})))\mathbf{p}' = \mathbf{y}(\tilde{\mathbf{B}}_\theta\mathrm{diag}(e^{\mathbf{r}^\star})), \\ &\Rightarrow \mathbf{p}' = \mathrm{diag}(\boldsymbol{\beta}\circ(\mathbf{1}-e^{-\mathbf{r}^\star}))\mathbf{y}(\tilde{\mathbf{B}}_\theta\mathrm{diag}(e^{\mathbf{r}^\star})). \end{aligned}$$

■

H. Proof of Lemma 5

Proof: The Lagrangian of (10) is given by $\mathcal{L}(\mathbf{r},\lambda,\boldsymbol{\mu}) = \mathbf{w}^\top\mathbf{r} - \lambda\log\rho(\tilde{\mathbf{B}}_\theta\mathrm{diag}(e^{\mathbf{r}})) + \boldsymbol{\mu}^\top\mathbf{r}$. Since (10) is convex, we use the KKT conditions to obtain $\frac{\partial\mathcal{L}}{\partial\mathbf{r}} = \mathbf{w} - \lambda\frac{\partial\log\rho(\tilde{\mathbf{B}}_\theta\mathrm{diag}(e^{\mathbf{r}}))}{\partial\mathbf{r}} + \boldsymbol{\mu} = 0$. Using (31), the optimality condition takes the form of $\mathbf{w} + \boldsymbol{\mu}^\star = \lambda^\star\mathbf{x}(\tilde{\mathbf{B}}_\theta\mathrm{diag}(e^{\mathbf{r}}))\circ\mathbf{y}(\tilde{\mathbf{B}}_\theta\mathrm{diag}(e^{\mathbf{r}}))$. The optimality condition in (25) is thus obtained by using $\mathbf{p}^\star$ in (11) and $\mathbf{p}'^\star$ in (24) to replace the Perron left and right eigenvectors. ■

I. Proof of Theorem 3

Proof: The relaxed sum rate maximization (14) is equivalent to

$$\begin{aligned} \underset{\mathbf{p}\geq\mathbf{0}}{\text{maximize}} \quad & \sum_l w_l \log\left(1 + \frac{p_l}{\left(\mathrm{diag}(\boldsymbol{\beta})(\mathbf{F}\mathbf{p}+\mathbf{n})\right)_l}\right) \\ \text{subject to} \quad & \mathbf{a}_\theta^\top\mathbf{p} \leq \bar{p}_\theta. \end{aligned}$$

By changing the nonnegative constraint from $\mathbf{p}\geq\mathbf{0}$ to $\log\left(1+\frac{p_l}{\left(\mathrm{diag}(\boldsymbol{\beta})(\mathbf{F}\mathbf{p}+\mathbf{n})\right)_l}\right)\geq 0$, $l = 1,\ldots,L$, the Lagrangian function with Lagrange multipliers λ and $\boldsymbol{\mu}$ is

$$\mathcal{L}(\mathbf{p},\lambda,\boldsymbol{\mu}) = \sum_l (w_l+\mu_l)\left(\log\left(\left(\mathrm{diag}(\boldsymbol{\beta})(\mathbf{F}\mathbf{p}+\mathbf{n})\right)_l + p_l\right) - \log\left(\left(\mathrm{diag}(\boldsymbol{\beta})(\mathbf{F}\mathbf{p}+\mathbf{n})\right)_l\right)\right) - \lambda(\mathbf{a}_\theta^\top\mathbf{p} - \bar{p}_\theta).$$

We then apply the KKT conditions to obtain

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial p_\iota} =& \sum_l (w_l+\mu_l)\frac{\beta_l F_{l\iota}}{\big(\mathrm{diag}(\boldsymbol{\beta})(\mathbf{F}\mathbf{p}+\mathbf{n})\big)_l + p_l} + (w_\iota+\mu_\iota)\frac{1}{\big(\mathrm{diag}(\boldsymbol{\beta})(\mathbf{F}\mathbf{p}+\mathbf{n})\big)_\iota + p_\iota}\\ &-\sum_l (w_l+\mu_l)\frac{\beta_l F_{l\iota}}{\big(\mathrm{diag}(\boldsymbol{\beta})(\mathbf{F}\mathbf{p}+\mathbf{n})\big)_l} - \lambda a_{\theta\iota}, \qquad \iota = 1,\dots L.\end{aligned} \tag{36}$$

Let $e^{r_l} = \frac{\big(\mathrm{diag}(\boldsymbol{\beta})(\mathbf{F}\mathbf{p}+\mathbf{n})\big)_l + p_l}{\big(\mathrm{diag}(\boldsymbol{\beta})(\mathbf{F}\mathbf{p}+\mathbf{n})\big)_l}$, i.e., $e^{r_l}-1 = \frac{p_l}{\big(\mathrm{diag}(\boldsymbol{\beta})(\mathbf{F}\mathbf{p}+\mathbf{n})\big)_l}$, then (36) becomes

$$\frac{\partial \mathcal{L}}{\partial p_\iota} = -\sum_l \frac{(w_l+\mu_l)(e^{r_l}-1)\beta_l F_{l\iota}}{\big(\mathrm{diag}(\boldsymbol{\beta})(\mathbf{F}\mathbf{p}+\mathbf{n})\big)_l + p_l} + \frac{(w_\iota+\mu_\iota)}{\big(\mathrm{diag}(\boldsymbol{\beta})(\mathbf{F}\mathbf{p}+\mathbf{n})\big)_\iota + p_\iota} - \lambda a_{\theta\iota}, \qquad \iota = 1,\dots,L.$$

By defining

$$\mathbf{p}' = \left[\frac{(w_1+\mu_1)\beta_1(e^{r_1}-1)}{\lambda\big(\mathrm{diag}(\boldsymbol{\beta})(\mathbf{F}\mathbf{p}+\mathbf{n})\big)_1 + \lambda p_1}, \dots, \frac{(w_L+\mu_L)\beta_L(e^{r_L}-1)}{\lambda\big(\mathrm{diag}(\boldsymbol{\beta})(\mathbf{F}\mathbf{p}+\mathbf{n})\big)_L + \lambda p_L}\right]^\top, \tag{37}$$

we can reformulate the KKT conditions as

$$\sum_{l=1}^{L} F_{l\iota} p_l'^\star + a_{\theta\iota} = \frac{p_\iota'^\star}{\beta_\iota(e^{r_\iota^\star}-1)}, \qquad \iota = 1,\dots,L. \tag{38}$$

After rearranging (38), we get

$$r_\iota^\star = \log\left(1 + \frac{p_\iota'^\star}{\big(\mathrm{diag}(\boldsymbol{\beta})(\mathbf{F}^\top\mathbf{p}'^\star + \mathbf{a}_\theta)\big)_\iota}\right), \qquad \iota = 1,\dots,L. \tag{39}$$

As can be seen, (39) is identical to the optimal solution in the virtual SIMO dual network constructed using the Perron-Frobenius theorem in Section IV-A.

To derive the bounding constraint condition, we first rewrite (39) as $\mathbf{F}^\top\mathbf{p}'^\star + \mathbf{a}_\theta = \mathrm{diag}(\boldsymbol{\beta}\circ(e^{\mathbf{r}^\star}-\mathbf{1}))^{-1}\mathbf{p}'^\star$. After multiplying $\mathbf{p}^{\star\top}$ on both sides and rearranging the expression of $\mathbf{p}$ in (6) to get $\mathbf{F}\mathbf{p}^\star+\mathbf{n} = \mathrm{diag}(\boldsymbol{\beta}\circ(e^{\mathbf{r}^\star}-1))\mathbf{p}^\star$, the bounding constraint is given by $\mathbf{n}^\top\mathbf{p}'^\star = \mathbf{a}_\theta^\top\mathbf{p}^\star = \bar{p}$.

From (37), the optimal dual power can be rewritten as

$$p_l'^\star = \frac{w_l+\mu_l^\star}{\lambda^\star}\frac{\beta_l(e^{r_l^\star}-1)^2}{p_l^\star e^{r_l^\star}}, \qquad l = 1,\dots,L,$$

leading to $\frac{\mathbf{w}+\boldsymbol{\mu}^\star}{\lambda^\star} = \mathrm{diag}(\boldsymbol{\beta}\circ(e^{\mathbf{r}^\star}-2\cdot\mathbf{1}+e^{-\mathbf{r}^\star}))^{-1}\mathbf{p}^\star\circ\mathbf{p}'^\star$, which is equivalent to the optimality condition derived using the Perron-Frobenius theorem in (25). ■

J. Proof of Lemma 6

Proof: The multiple-constrained sum rate maximization (5) can be represented as

$$\begin{aligned}\underset{\mathbf{p}\geq\mathbf{0},\ \mathbf{q}\geq\mathbf{0}}{\text{maximize}} \quad & \sum_l w_l \log\left(1+\frac{p_l}{\big(\mathrm{diag}(\boldsymbol{\beta})(\mathbf{F}\mathbf{p}+\mathbf{n})\big)_l}\right)\\ \text{subject to} \quad & \mathbf{a}_k^\top\mathbf{p} \leq \bar{p}_k,\ k=1,\dots,K,\\ & \mathbf{b}_m^\top\mathbf{q} \leq \bar{q}_m\ m=1,\dots,M.\end{aligned}$$

By using (12), the interference temperature constraints can be rewritten as the power constraints, and thus the variables are reduced to only powers. By changing the nonnegative constraint from $\mathbf{p}\geq\mathbf{0}$ to $\log\left(1+\frac{p_l}{\big(\mathrm{diag}(\boldsymbol{\beta})(\mathbf{F}\mathbf{p}+\mathbf{n})\big)_l}\right) \geq 0$, $l = 1,\dots,L$, we write the Lagrangian function with Lagrange multipliers $\boldsymbol{\varrho}$ and $\boldsymbol{\mu}$ as

$$\mathcal{L}(\mathbf{p},\boldsymbol{\varrho},\boldsymbol{\mu}) = \sum_l w_l \log\left(1+\frac{p_l}{\big(\mathrm{diag}(\boldsymbol{\beta})(\mathbf{F}\mathbf{p}+\mathbf{n})\big)_l}\right) + \sum_l \mu_l \log\left(1+\frac{p_l}{\big(\mathrm{diag}(\boldsymbol{\beta})(\mathbf{F}\mathbf{p}+\mathbf{n})\big)_l}\right)$$
$$-\sum_{k=1}^{K}\varrho_k(\mathbf{a}_k^\top\mathbf{p}-\bar{p}_k) - \sum_{m=1}^{M}\varrho_{K+m}\big((\mathbf{F}^\top\mathrm{diag}(\boldsymbol{\beta})\mathbf{b}_m)^\top\mathbf{p}-\bar{q}_m+\mathbf{b}_m^\top\mathrm{diag}(\boldsymbol{\beta})\mathbf{n}\big).$$

We then apply the KKT conditions to obtain

$$\frac{\partial\mathcal{L}}{\partial p_\iota} = \sum_l (w_l+\mu_l)\frac{\beta_l F_{l\iota}}{\big(\mathrm{diag}(\boldsymbol{\beta})(\mathbf{F}\mathbf{p}+\mathbf{n})\big)_l+p_l} + (w_\iota+\mu_\iota)\frac{1}{\big(\mathrm{diag}(\boldsymbol{\beta})(\mathbf{F}\mathbf{p}+\mathbf{n})\big)_l+p_l}$$
$$-\sum_l (w_l+\mu_l)\frac{\beta_l F_{l\iota}}{\big(\mathrm{diag}(\boldsymbol{\beta})(\mathbf{F}\mathbf{p}+\mathbf{n})\big)_l} - \sum_{k=1}^{K}\varrho_k a_{k\iota} - \sum_{m=1}^{M}\varrho_{K+m}(\mathbf{F}^\top\mathrm{diag}(\boldsymbol{\beta})\mathbf{b}_m)_\iota, \qquad \iota=1,\ldots L.$$

Using the optimal dual power defined in (37), we have the optimality condition as

$$\sum_{l=1}^{L}F_{l\iota}p_l'^\star + \frac{1}{\lambda^\star}\left(\sum_{k=1}^{K}\varrho_k^\star a_{k\iota} + \sum_{m=1}^{M}\varrho_{K+m}^\star(\mathbf{F}^\top\mathrm{diag}(\boldsymbol{\beta})\mathbf{b}_m)_\iota\right)$$
$$\overset{(a)}{=} \sum_{l=1}^{L}F_{l\iota}p_l'^\star + \underbrace{\sum_{k=1}^{K}\theta_k^\star a_{k\iota} + \sum_{m=1}^{M}\theta_{m+K}^\star(\mathbf{F}^\top\mathrm{diag}(\boldsymbol{\beta})\mathbf{b}_m)_\iota}_{a_{\theta_\iota^\star}} = \frac{p_\iota'^\star}{\beta_\iota(e^{r_\iota^\star}-1)},$$

where (a) is true by letting $\boldsymbol{\theta}^\star = \frac{1}{\lambda^\star}\boldsymbol{\varrho}^\star$. Thus, we have proved that the above equality is identical to the optimality condition (38) for the relaxed sum rate maximization with the optimal value $\boldsymbol{\theta}^\star$. ■

K. Proof of Lemma 7

Proof: The epigraph form of the max-min rate maximization in (28) is given by

$$\begin{array}{ll}\underset{\delta,\,\mathbf{r}}{\text{maximize}} & \delta\\ \text{subject to} & \delta \le r_l/\tilde{w}_l, \quad l=1,\ldots,L,\\ & \log\rho(\tilde{\mathbf{B}}_\theta\mathrm{diag}(e^{\mathbf{r}}))\le 0.\end{array} \tag{40}$$

With the dual variables $\boldsymbol{\lambda}\in\mathbb{R}^L_{\ge 0}$ and $\mu\ge 0$ being, respectively, the Lagrange multipliers for the L inequality constraints $\delta\le r_l/\tilde{w}_l,\ l=1,\ldots,L$, and the spectral radius constraint $\log\rho\big(\tilde{\mathbf{B}}_\theta\mathrm{diag}(e^{\mathbf{r}})\big)\le 0$, the Lagrangian function of (40) is derived as $\mathcal{L}(\delta,\mathbf{r},\boldsymbol{\lambda},\mu)=\delta-\boldsymbol{\lambda}^\top(\delta\mathbf{1}-\mathrm{diag}(\tilde{\mathbf{w}})^{-1}\mathbf{r})-\mu\log\rho\big(\tilde{\mathbf{B}}_\theta\mathrm{diag}(e^{\mathbf{r}})\big)$. Taking the first order derivative of the Lagrangian function with respect to $\mathbf{r}$, we have $\frac{\partial\mathcal{L}}{\partial\mathbf{r}}=\mathrm{diag}(\tilde{\mathbf{w}})^{-1}\boldsymbol{\lambda}-\mu\mathbf{x}\big(\tilde{\mathbf{B}}_\theta\mathrm{diag}(e^{\mathbf{r}})\big)\circ\mathbf{y}\big(\tilde{\mathbf{B}}_\theta\mathrm{diag}(e^{\mathbf{r}})\big)$. By letting $\partial\mathcal{L}/\partial\mathbf{r}=0$, we obtain $\boldsymbol{\lambda}^\star=\mu^\star\mathrm{diag}(\tilde{\mathbf{w}})\mathbf{x}\big(\tilde{\mathbf{B}}_\theta\mathrm{diag}(e^{\mathbf{r}})\big)\circ\mathbf{y}\big(\tilde{\mathbf{B}}_\theta\mathrm{diag}(e^{\mathbf{r}})\big)$. The nonnegative dual variables and the nonnegative Perron eigenvectors of the nonnegative matrix $\tilde{\mathbf{B}}_\theta\mathrm{diag}(e^{\mathbf{r}})$ lead to the conclusion that $\boldsymbol{\lambda}^\star>0$. Combining this fact with complementary slackness, we thus have $\delta^\star\mathbf{1}-\mathrm{diag}(\tilde{\mathbf{w}})^{-1}\mathbf{r}^\star=\mathbf{0}\Rightarrow\mathbf{r}^\star=\delta^\star\tilde{\mathbf{w}}\ge\mathbf{0}$. ■

L. Proof of Lemma 8

Proof: By using the result $1/\delta^{\star} = \tilde{w}_l/r_l$ in Lemma 7, we can define the function $f : \mathbb{R}^L \to \mathbb{R}^L$ $\frac{1}{\delta^{\star}}\mathbf{p}^{\star} = f(\mathbf{p}^{\star})$ with

$$f_l(\mathbf{p}) = \frac{\tilde{w}_l}{r_l}p_l = \frac{\tilde{w}_l}{\log(1+p_l/(\mathrm{diag}(\boldsymbol{\beta})(\mathbf{F}\mathbf{p}+\mathbf{n}))_l)}p_l. \tag{41}$$

When the constraint in (28) is tight, we equivalently have a tight power constraint $\mathbf{a}_{\theta}\mathbf{p}^{\star}/\bar{p}_{\theta} = 1$ by applying the technique of change of variables used in the relaxed sum rate maximization in (14) and (15). Obviously, this power constraint is a weighted linear (monotone) norm in $\mathbf{p}^{\star}$. To prove the convergence of Algorithm 1, we use the following nonlinear Perron-Frobenius theory [32].

Theorem 4 (Krause's theorem [32]): Defining $\|\cdot\|$ as a monotone norm on $\mathbb{R}^L$, the concave mapping $f : \mathbb{R}^L \to \mathbb{R}^L$ has the following features: The conditional eigenvalue problem $f(\mathbf{z}) = \lambda\mathbf{z}$, $\lambda \in \mathbb{R}$, $\mathbf{z} \geq 0$ and $\|\mathbf{z}\| = 1$, has a unique positive solution $(\lambda^{\star}, \mathbf{z}^{\star})$. Moreover, $\lim_{k\to\infty}\frac{1}{\|\mathbf{z}(k)\|}f(\mathbf{z}(k))$ converges geometrically fast to $\mathbf{z}^{\star}$, where k is the number of fixed point iterations.

The proof of $f(\mathbf{p})$ being concave can be found in [36]. Let $f(\mathbf{p})$ be given in (41), $\lambda = 1/\delta$, and $\mathbf{p}$ is constrained by the monotone norm $\|\cdot\|$ of the tight constraint $\mathbf{a}_{\theta}^{\top}\mathbf{p}^{\star} = \bar{p}_{\theta}$. According to Theorem 4, $\mathbf{p}(k)$ in Algorithm 1 converges to the fixed point $\mathbf{p}^{\star} = \frac{1}{\|\mathbf{p}^{\star}\|}f(\mathbf{p}^{\star})$ geometrically fast. ■

M. Proof of Lemma 9

Proof: Before the proof, we state the Friedland-Karlin inequality that is fundamental in connecting the max-min rate optimization in (28) to the sum rate maximization in (5).

Theorem 5 (The Friedland-Karlin inequality [35]): Let $\mathbf{A} \in \mathbb{R}^{L\times L}_{\geq 0}$ be an irreducible nonnegative matrix. Assume that $\mathbf{x}(\mathbf{A})$ and $\mathbf{y}(\mathbf{A})$ are nonnegative Perron right and left eigenvectors of $\mathbf{A}$, normalized such that $\mathbf{x}(\mathbf{A})^{\top}\mathbf{y}(\mathbf{A}) = 1$. Suppose $\boldsymbol{\eta}$ is a nonnegative vector. Then,

$$\rho(\mathbf{A})\prod_{l=1}^{L}\eta_l^{(\mathbf{x}(\mathbf{A})\circ\mathbf{y}(\mathbf{A}))_l} \leq \rho\big(\mathbf{A}\mathrm{diag}(\boldsymbol{\eta})\big). \tag{42}$$

The equality holds if and only if all η_l are equal.

By inserting $\boldsymbol{\eta} = e^{\mathbf{r}}$ and $\mathbf{A} = \tilde{\mathbf{B}}_{\theta}$ into (42) and taking the logarithm on both sides, we obtain

$$\log\rho(\tilde{\mathbf{B}}_{\theta}) + \sum_{l=1}^{L}(\mathbf{x}(\tilde{\mathbf{B}}_{\theta})\circ\mathbf{y}(\tilde{\mathbf{B}}_{\theta}))_l r_l \leq \log\rho(\tilde{\mathbf{B}}_{\theta}\mathrm{diag}(e^{\mathbf{r}})),$$

$$\Rightarrow \sum_{l=1}^{L}(\mathbf{x}(\tilde{\mathbf{B}}_{\theta})\circ\mathbf{y}(\tilde{\mathbf{B}}_{\theta}))_l r_l \leq -\log\rho(\tilde{\mathbf{B}}_{\theta}) + \log\rho(\tilde{\mathbf{B}}_{\theta}\mathrm{diag}(e^{\mathbf{r}})). \tag{43}$$

When $\mathbf{r} = \mathbf{r}^{\star}$, we have $\log\rho(\tilde{\mathbf{B}}_{\theta}\mathrm{diag}(e^{\mathbf{r}^{\star}})) = 0$. Also, the equality in (43) is achieved if and only if all r_l are equal. Due to these two conditions and by defining $\mathbf{w} = \mathbf{x}(\tilde{\mathbf{B}}_{\theta})\circ\mathbf{y}(\tilde{\mathbf{B}}_{\theta})$, we have the optimal value $\sum_{l=1}^{L} w_l r_l^{\star} = -\log\rho(\tilde{\mathbf{B}}_{\theta})$ and the optimal solution $r_l^{\star} = -\log\rho(\tilde{\mathbf{B}}_{\theta})$ of the sum rate maximization.

For the weighted max-min rate maximization (28), the optimal rate with weight vector $\tilde{\mathbf{w}} = \mathbf{1}$ is $\mathbf{r}^\star = \delta^\star \mathbf{1}$ according to Lemma 7. Moreover, the spectral radius constraint is tight when the optimal rate is used, i.e., $\rho(\tilde{\mathbf{B}}_\theta \mathrm{diag}(e^{\tilde{\mathbf{w}}\delta^\star})) = 1$. Transforming this condition to the logarithmic domain, the optimal value is equal to the optimal value of considering the weighted sum rate maximization optimization with $\mathbf{w} = \mathbf{x}(\tilde{\mathbf{B}}_\theta) \circ \mathbf{y}(\tilde{\mathbf{B}}_\theta)$, i.e., $-\log(\rho(\tilde{\mathbf{B}}_\theta))$. ■

REFERENCES

[1] J. Mitola III, "Cognitive radio for flexible mobile multimedia communications," in *Proc. 6th IEEE Int. Workshop on Mobile Multimedia Communications (MOMUS'99)*, Nov. 1999.

[2] J. Mitola III and J. G. Q. Maguire, "Cognitive radio: Making software radios more personal," *IEEE Personal Communications*, vol. 6, no. 4, pp. 13–18, Aug. 1999.

[3] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201–220, Feb. 2005.

[4] K.-C. Chen and R. Prasad, *Cognitive Radio Networks*. Chichester, UK: John Wiley & Sons, 2009.

[5] E. Biglieri, A. J. Goldsmith, L. J. Greenstein, N. B. Mandayam, and H. V. Poor, *Principles of Cognitive Radio*. Cambridge, UK: Cambridge University Press, 2013.

[6] 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, June 2010.

[7] K. R. Krishnan and H. Luss, "Power selection for maximizing SINR in femtocells for specified SINR in macrocell," in *Proc. IEEE Wireless Commun. and Networking Conf. (WCNC)*, Mar. 2011.

[8] L. Zheng and C. W. Tan, "Cognitive radio network duality and algorithms for utility maximization," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 3, pp. 500–513, Mar. 2013.

[9] 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, Feb. 2011.

[10] 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, Apr. 2012.

[11] M. Ebrahimi, M. A. Maddah-Ali, and A. K. Khandani, "Power allocation and asymptotic achievable sum-rates in single-hop wireless networks," in *Proc. Conf. Inform. Sciences and Systems (CISS)*, Mar. 2006.

[12] Z.-Q. Luo and S. Zhang, "Dynamic spectrum management: Complexity and duality," *IEEE J. Sel. Top. Sign. Proces.*, vol. 2, no. 1, pp. 57–73, Feb. 2008.

[13] W. Yu and R. Lui, "Dual methods for nonconvex spectrum optimization of multicarrier systems," *IEEE Trans. Commun.*, vol. 54, no. 7, pp. 1310–1322, July 2006.

[14] S. Hayashi and Z.-Q. Luo, "Spectrum management for interference-limited multiuser communication systems," *IEEE Trans. Inf. Theory*, vol. 55, no. 3, pp. 1153–1175, Mar. 2009.

[15] M. Chiang, C. W. Tan, P. Palomar, D. O'Neill, and D. Julian, "Power control by geometric programming," *IEEE Trans. Wireless Commun.*, vol. 6, no. 7, pp. 2640–2651, July 2007.

[16] E. Matskani, N. D. Sidiropoulos, and L. Tassiulas, "Convex approximation algorithms for back-pressure power control," *IEEE Trans. Signal Process.*, vol. 60, no. 4, pp. 1957–1970, Apr. 2012.

[17] T. Wang and L. Vandendorpe, "On the SCALE algorithm for multiuser multicarrier power spectrum management," *IEEE Trans. Signal Process.*, vol. 60, no. 9, pp. 4992–4998, Sept. 2012.

[18] K. Eriksson, S. Shi, N. Vucic, and M. Schubert, "Globally optimal resource allocation for achieving maximum weighted sum rate," in *Proc. IEEE Glob. Comm. Conf. (GLOBECOM)*, Dec. 2010.

[19] H. Al-Shatri and T. Weber, "Achieving the maximum sum rate using D.C. programming in cellular networks," *IEEE Trans. Signal Process.*, vol. 60, no. 3, pp. 1331–1341, Mar. 2012.

[20] 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, Feb. 2011.

[21] ——, "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, Sept. 2011.

[22] 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, Dec. 2011.

[23] ——, "Fast algorithms and performance bounds for sum rate maximization in wireless networks," *IEEE/ACM Trans. Networking*, vol. 21, no. 3, pp. 706–719, June 2013.

[24] L. Zheng and C. W. Tan, "Maximizing sum rates in cognitive radio networks: Convex relaxation and global optimization algorithms," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 4, Oct. 2014.

[25] Y. K. Wong, "Some mathematical concepts for linear economic models," *Economic Activity Analysis O. Morgenstern, Ed. New York: Wiley*, pp. 283–339, 1954.

[26] S. Boyd and L. Vanderberghe, *Convex Optimization*. Cambridge University Press, 2004.

[27] H. Huh, H. C. Papadopoulos, and G. Caire, "Multiuser MIMO transmitter optimization for inter-cell interference mitigation," *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 4272–4285, Aug. 2010.

[28] L. Zhang, R. Zhang, Y.-C. Liang, Y. Xin, and H. V. Poor, "On gaussian MIMO BC-MAC duality with multiple transmit covariance constraints," *IEEE Trans. Inf. Theory*, vol. 58, no. 4, pp. 2064–2078, Apr. 2012.

[29] 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, Nov. 1993.

[30] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Randomized gossip algorithms," *IEEE Trans. Inf. Theory*, vol. 52, no. 6, pp. 2508–2530, June 2006.

[31] D. P. Bertsekas and J. N. . Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*. Athena Scientific, 1997.

[32] U. Krause, "Concave Perron-Frobenius theory and applications," *Nonlinear analysis*, vol. 47, pp. 1457–1466, 2001.

[33] W. W. Hager, "Updating the inverse of a matrix," *SIAM Rev.*, vol. 31, no. 2, pp. 221–239, June 1989.

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

[35] 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.

[36] D. W. H. Cai, C. W. Tan, and S. H. Low, "Optimal max-min fairness rate control in wireless networks: Perron-Frobenious characterization and algorithms," in *Proc. IEEE INFOCOM*, Mar. 2012.