Spectrum Management in Multiuser Cognitive Wireless Networks: Optimality and Algorithm

Chee Wei Tan, Shmuel Friedland, and Steven H. Low

Abstract—Spectrum management is used to improve performance in multiuser communication system, e.g., cognitive radio or femtocell networks, where multiuser interference can lead to data rate degradation. We study the nonconvex NP-hard problem of maximizing a weighted sum rate in a multiuser Gaussian interference channel by power control subject to affine power constraints. By exploiting the fact that this problem can be restated as an optimization problem with constraints that are spectral radii of specially crafted nonnegative matrices, we derive necessary and sufficient optimality conditions and propose a global optimization algorithm based on the outer approximation method. Central to our techniques is the use of nonnegative matrix theory, e.g., nonnegative matrix inequalities and the Perron-Frobenius theorem. We also study an inner approximation method and a relaxation method that give insights to special cases. Our techniques and algorithm can be extended to a multiple carrier system model, e.g., OFDM system or receivers with interference suppression capability.

Index Terms—Optimization, nonnegative matrix theory, dynamic spectrum access, power control, cognitive wireless networks.

I. INTRODUCTION

WHEN multiple users transmit simultaneously over the same frequency band of the wireless medium, the data rates are affected by multiuser interference. For example, interference is a major source of performance impairment in the Code Division Multiple Access (CDMA) cellular network (see [1], [2]) and the cognitive wireless networks (see [3]–[6]). How to optimize performance in the presence of interference is thus an important issue. Power control is often used to improve the spectral utilization and the system performance, e.g., maximizing the total data rates of all users [2], [3], [7]–[9]. We refer to power control techniques that adapt the spectrum allocation to maximize performance in a multiuser system as Dynamic Spectrum Management (DSM). There are many DSM algorithms proposed to maximize the total data rates of all users in the literature [1]–[4], [10], [11].[1]

Cognitive radio has recently emerged as a new technology that can substantially increase spectrum utilization efficiency by allowing unlicensed (secondary) users to share a common spectrum with licensed (primary) users so long as the interference caused is maintained below a certain prescribed level known as the *interference temperature* [4], [12], [13]. In other words, regulatory constraints can be translated into power constraints to influence the performance of all users. A practical approach to obtain achievable rates is to assume and treat multiuser interference as additive Gaussian noise. This assumption is relevant to cognitive wireless networks, where low-complexity coordination is desired for a decentralized implementation, e.g., when joint decoding of primary and secondary users is not possible.

Maximizing the weighted sum rate in a multiuser Gaussian channel where interference is treated as noise is a nonconvex problem [1], [4], [10], [11]. It is shown to be NP-hard in [10]. The algorithms for this problem in the literature are mostly based on game theory, e.g., [4], [13], or the Lagrange dual decomposition [11], [14], [15]. Due to the duality gap, dual algorithms cannot guarantee finding a global optimal solution. Moreover, finding a feasible primal solution for a given feasible dual solution, or vice versa, is difficult. Prohibitive as it seems, several recent advances have been made by exploiting the problem structure. The authors in [10] show that the duality gap is zero when the number of frequency tones is asymptotically large. Computational algorithms based on the difference of convex functions are proposed in [9]. Fast distributed algorithms can be found in [1], [16]–[18] that solve the problem suboptimally but with performance guarantees (and optimally under special cases). The authors in [19] use nonnegative matrix theory to analyze the problem and give a branch-and-bound algorithm.

This paper is organized as follows. We introduce the system model in Section II. In Section III, we state the weighted sum rate maximization problem and present an equivalent reformulation problem. In Section IV, a further preprocessing step converts the optimization variable from power to Signal-to-Interference Ratio (SIR) in logarithmic scale.[2] We then characterize the global optimality conditions, and propose a global optimization algorithm in a single-carrier model. In Section V, approximation and relaxation techniques as well as special cases are presented. In Section VI, numerical simulations show that the global optimization algorithm is computationally fast for small-to-medium problem size. In Section VII, we illustrate how our techniques can be extended to a multiple carrier model and to include multiuser interference suppression. We conclude the paper in Section VIII. All proofs are found in the appendix.

The following notation is used. Boldface uppercase letters denote matrices, boldface lowercase letters denote column vec-

[1]Related work in the wireline setting can be found in a Digital Subscriber Line (DSL) system, where users (twisted-pairs) in a cable binder share a common spectrum and interfere with one another [10], [11].

[2]Measurements of power and SIR variable in wireless networks are typically expressed in decibels (dB), i.e., base-10 logarithmic scale.

tors, italics denote scalars, and $\mathbf{u} \geq \mathbf{v}$ denotes componentwise inequality between vectors $\mathbf{u}$ and $\mathbf{v}$. We also let $(\mathbf{By})_l$ denote the lth element of $\mathbf{By}$. Let $\mathbf{x} \circ \mathbf{y}$ denote the Schur product of the vectors $\mathbf{x}$ and $\mathbf{y}$, i.e., $\mathbf{x} \circ \mathbf{y} = [x_1 y_1, \ldots, x_L y_L]^\top$. We write $\mathbf{B} \geq \mathbf{F}$ if $B_{ij} \geq F_{ij}$ for all i, j. The Perron-Frobenius eigenvalue of a nonnegative matrix $\mathbf{F}$ is denoted as $\rho(\mathbf{F})$, and the Perron right and left eigenvector of $\mathbf{F}$ associated with $\rho(\mathbf{F})$ are denoted by $\mathbf{x}(\mathbf{F}) \geq \mathbf{0}$ and $\mathbf{y}(\mathbf{F}) \geq \mathbf{0}$ (or, simply $\mathbf{x}$ and $\mathbf{y}$, when the context is clear) respectively. Recall that the Perron-Frobenius eigenvalue of $\mathbf{F}$ is the eigenvalue with the largest absolute value. Assume that $\mathbf{F}$ is a nonnegative irreducible matrix. Then $\rho(\mathbf{F})$ is simple and positive, and $\mathbf{x}(\mathbf{F}), \mathbf{y}(\mathbf{F}) > \mathbf{0}$ [20]. We will assume the normalization: $\mathbf{x}(\mathbf{F}) \circ \mathbf{y}(\mathbf{F})$ is a probability vector. The super-script $(\cdot)^\top$ denotes transpose. We denote $\mathbf{e}_l$ as the lth unit coordinate vector and $\mathbf{I}$ as the identity matrix. For any vector $\tilde{\boldsymbol{\gamma}} = (\tilde{\gamma}_1, \ldots, \tilde{\gamma}_L)^\top \in \mathbb{R}^L$, let $e^{\tilde{\boldsymbol{\gamma}}} = (e^{\tilde{\gamma}_1}, \ldots, e^{\tilde{\gamma}_L})^\top$. Let $P : X \to Y$ be a mapping from the space X to the space Y. For a subset $Z \subset X$, we denote by $P(Z)$ the image of the set Z.

II. SYSTEM MODEL

Consider a multiuser communication system with L users (logical transmitter/receiver pairs) sharing a common frequency. Each user employs a single-user decoder, i.e., treating interference as additive Gaussian noise, and has perfect channel state information at the receiver. We assume that fading occurs sufficiently slowly in the channel, i.e., flat-fading, so that the channel can be considered essentially fixed during transmission. Our system with L users can be modeled by a Gaussian interference channel having the baseband signal model:

$$y_l = h_{ll} x_l + \sum_{j \neq l} h_{lj} x_j + z_l, \tag{1}$$

where $y_l \in \mathbb{C}^{1 \times 1}$ is the received signal of the lth user, $h_{lj} \in \mathbb{C}^{1 \times 1}$ is the channel coefficient between the transmitter of the jth user and the receiver of the lth user, $x \in \mathbb{C}^{N \times 1}$ is the transmitted (information carrying) signal vector, and z_l's are the i.i.d. additive complex Gaussian noise coefficient with variance $n_l/2$ on each of its real and imaginary components. The first term on the right-hand side of (1) represents the desired signal, whereas the second term represents the interfering signals from other users. At each transmitter, the signal is constrained by an average power constraint, i.e., $\mathbb{E}[|x_l|^2] = p_l$, which we assume to be upper bounded by $\bar{p}_l$ for all l.

The vector $(p_1, \ldots, p_L)^\top$ is the transmit power vector and is the optimization variable of interest in this paper. Let $G = [G_{lj}]_{l,j=1}^L > 0_{L \times L}$ represents the channel gain, where $G_{lj} = |h_{lj}|^2$ is the channel gain from the jth transmitter to the lth receiver, and $\mathbf{n} = (n_1, \ldots, n_L)^\top > \mathbf{0}$, where n_l is the noise power at the lth receiver. Figure 1(a) shows the system model with the problem parameters for the 2-user case.

Next, we define a nonnegative matrix $\mathbf{F}$ with entries:

$$F_{lj} = \begin{cases} 0, & \text{if } l = j \\ \frac{G_{lj}}{G_{ll}}, & \text{if } l \neq j \end{cases} \tag{2}$$

and the vector

$$\mathbf{v} = \left(\frac{n_1}{G_{11}}, \frac{n_2}{G_{22}}, \ldots, \frac{n_L}{G_{LL}} \right)^\top. \tag{3}$$

Moreover, we assume that $\mathbf{F}$ is irreducible, i.e., each link has at least an interferer.

Assuming a linear matched-filter receiver, the SIR for the lth receiver (as a function of powers) is given by:

$$\gamma_l(\mathbf{p}) := \frac{p_l}{\sum_{j \neq l} F_{lj} p_j + v_l}, \ l = 1, \ldots, L. \tag{4}$$

Note that (4) as a mapping $\mathbf{p} \mapsto \boldsymbol{\gamma}(\mathbf{p})$ for $\mathbf{p} = (p_1, \ldots, p_L)^\top \geq \mathbf{0}$ can be compactly written as $p_l/((\mathbf{Fp})_l + v_l)$. Let us denote the vector $\boldsymbol{\gamma}(\mathbf{p}) = (\gamma_1(\mathbf{p}), \ldots, \gamma_L(\mathbf{p}))^\top$.

Now, (4) is the transformation from power to SIR. We next give the transformation from SIR to power, which is first obtained in [21] (also given in [2], Chap. 2).

Lemma 1: Assume that $\boldsymbol{\gamma}(\mathbf{p})$ is defined by (4). Then $\rho(\text{diag}(\boldsymbol{\gamma}(\mathbf{p}))\mathbf{F}) < 1$. Hence, for $\boldsymbol{\gamma} = \boldsymbol{\gamma}(\mathbf{p})$,

$$\mathbf{p} = P(\boldsymbol{\gamma}) := (\mathbf{I} - \text{diag}(\boldsymbol{\gamma})\mathbf{F})^{-1} \text{diag}(\boldsymbol{\gamma})\mathbf{v}. \tag{5}$$

Vice versa, if $\boldsymbol{\gamma}$ is in the set

$$\boldsymbol{\Gamma} := \{ \boldsymbol{\gamma} \geq \mathbf{0}, \ \rho(\text{diag}(\boldsymbol{\gamma})\mathbf{F}) < 1 \}, \tag{6}$$

then $\mathbf{p}$ given in (5) is nonnegative. Furthermore, $\boldsymbol{\gamma}(P(\mathbf{p})) = \boldsymbol{\gamma}$. That is, $\boldsymbol{\gamma} : \mathbb{R}_+^L \to \boldsymbol{\Gamma}$, and $P : \boldsymbol{\Gamma} \to \mathbb{R}_+^L$ are inverse mappings.

Let $\bar{\boldsymbol{\gamma}}$ denote the vector $(\bar{p}_1/v_1, \ldots, \bar{p}_L/v_L)^\top$. Then, we have $\boldsymbol{\gamma} \leq \bar{\boldsymbol{\gamma}}$ as an implicit constraint.

In wireless networks, transmit power constraints model resource budget constraints or may be imposed by radio regulatory body. We assume that all users have individual power constraints (see (8) in the next section). In addition, unlike traditional wireless networks, a cognitive network allows secondary users to transmit provided that the performance degradation caused to the primary user is null or tolerable [4]. We assume that interference temperature constraints are modeled as additional affine constraints that are imposed on the transmit powers. These constraints are chosen by the primary users based on their quality of service requirements [4], [13]. We first consider the problem under individual power constraints, and then incorporate the interference temperature affine power constraints in Section IV-B.

III. THE SUM RATE MAXIMIZATION PROBLEM

Under the assumption of single-user decoding, we further assume that each user employs Gaussian code. In practice, Gaussian codes can be replaced by finite-order signal constellations such as quadrature-amplitude modulation (QAM) or other practical coding schemes. Assuming a fixed bit error rate (BER) at the receiver, the achievable data rate r_l of the lth user can be computed by the Shannon capacity formula [22]:

$$\log \left(1 + \frac{\gamma_l(\mathbf{p})}{\Gamma} \right) \quad \text{nats/symbol}, \tag{7}$$

where Γ is the SNR gap to capacity, which is always greater than 1. In this paper, we absorb $(1/\Gamma)$ into G_{ll} for all l, and write the achievable data rate as $r_l = \log(1 + \gamma_l(\mathbf{p}))$.

Let $\mathbf{w} = (w_1, \ldots, w_L)^\top \geq 0$ be a given probability vector, where w_l is a weight assigned to the lth link to reflect priority (a larger weight reflects a higher priority). Denote the set of feasible powers:

$$\Omega = \{ \mathbf{p} \mid \mathbf{0} \leq \mathbf{p} \leq \bar{\mathbf{p}} \}. \tag{8}$$
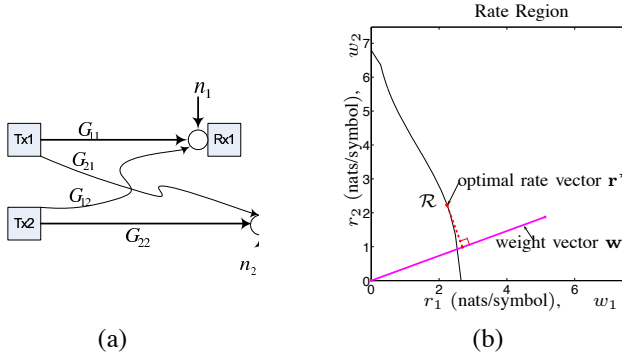
Fig. 1. (a) The system model for the 2-user case. (b) Illustrating an achievable rate region $\mathcal{R}$ for a 2-user Gaussian interference-limited channel. The positive weight vector $\mathbf{w}$ is superimposed on the rate region. Given a weight vector $\mathbf{w}$, the optimal rate vector $\mathbf{r}^\star = [r_1^\star, r_2^\star]$ is chosen on the boundary of the achievable rate region, where a perpendicular line from $\mathbf{w}$ (shown as the red dotted line) intersects with $\mathcal{R}$.

The sum rate maximization problem in a multiuser Gaussian channel can be stated as [1], [4], [10], [17]:

$$\text{maximize} \quad \sum_{l=1}^L w_l r_l = \sum_{l=1}^L w_l \log(1 + \gamma_l(\mathbf{p}))$$
$$\text{subject to} \quad \mathbf{p} \in \Omega. \tag{9}$$

We denote the optimal solution of (9) by $\mathbf{p}^\star = (p_1^\star, \dots, p_L^\star)^\top$. For any feasible $\mathbf{p} \in \Omega$, we call $\mathbf{r} = [r_1, \dots, r_L]$, evaluated at $\mathbf{p}$, a feasible rate vector that lies in the rate region $\mathcal{R}$, which is the set of all feasible rate vectors. The data rate evaluated at $\mathbf{p}^\star$ is given by $r_l^\star = \log(1 + \gamma_l(\mathbf{p}^\star))$. For the two user case, i.e., $L = 2$, Figure 1(b) gives the geometrical illustration of the weighted sum rate maximization problem in the rate region, i.e., finding $\mathbf{r}^\star \in \mathcal{R}$ for a given $\mathbf{w}$.

We now state an alternative formulation of (9) given in [19].

Theorem 1: The optimal value in (9) is equal to the optimal value of the problem:

$$\text{maximize} \quad \sum_l w_l \log(1 + \gamma_l)$$
$$\text{subject to} \quad \rho(\text{diag}(\boldsymbol{\gamma})(\mathbf{F} + (1/\bar{p}_l)\mathbf{ve}_l^\top)) \leq 1 \quad \forall l, \tag{10}$$
$$\text{variables:} \quad \gamma_l, \quad \forall l.$$

Now, $\boldsymbol{\gamma}^\star$ is an optimal solution to (10) if and only if $P(\boldsymbol{\gamma}^\star)$ is an optimal solution to (9). In particular, $\boldsymbol{\gamma}^\star$ satisfies

$$\rho\left(\text{diag}(\boldsymbol{\gamma}^\star)\left(\mathbf{F} + (1/\bar{p}_l)\mathbf{ve}_i^\top\right)\right) = 1 \tag{11}$$

for some integer i.

Note that (11) implies that the optimal solution of (9) is such that $p_i^\star = \bar{p}_i$ for some i. The transformation from SIR to power in (5) plays a central role in obtaining the constraint set of (10), which facilitates our algorithm design with nonnegative matrix theory in Section IV-A. It is also key to proving the optimality conditions in Section IV. Note that the formulation given in (10) shows that an optimal solution to (9), $\mathbf{p}^\star$, is in general not unique.

IV. GLOBAL OPTIMIZATION

In this section, we show that (10) is equivalent to an optimization problem that maximizes a convex function on a closed unbounded domain. For $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_L)^\top > 0$, let

$$\tilde{\boldsymbol{\gamma}} = \log \boldsymbol{\gamma}, \tag{12}$$

i.e., $\boldsymbol{\gamma} = e^{\tilde{\boldsymbol{\gamma}}}$. Then, (10) is equivalent to:

$$\text{maximize} \quad f(\tilde{\boldsymbol{\gamma}}) = \sum_l w_l \log(1 + e^{\tilde{\gamma}_l})$$
$$\text{subject to} \quad \log \rho(\text{diag}(e^{\tilde{\boldsymbol{\gamma}}})(\mathbf{F} + (1/\bar{p}_l)\mathbf{ve}_l^\top)) \leq 0 \quad \forall l,$$
$$\text{variables:} \quad \tilde{\boldsymbol{\gamma}} = (\tilde{\gamma}_1, \dots, \tilde{\gamma}_n)^\top \in \mathbb{R}^L. \tag{13}$$

Remark 1: Now, for a nonnegative matrix $B \in \mathbb{R}_+^{L \times L}$, $\log \rho(e^x B)$ is a convex function [23]. Therefore, the constraint

set in (13) is convex. Also, since $\log(1 + e^{\tilde{\gamma}_l})$ is strictly convex in $\tilde{\gamma}_l \in \mathbb{R}$ for all l, (13) is a convex maximization problem. The unboundedness of the convex set in (13) is due to the identity $0 = e^{-\infty}$.

We next denote the convex set in (13) by

$$D(\{\mathbf{F}\}) = \{\tilde{\boldsymbol{\gamma}} \in \mathbb{R}^L, \log \rho(\text{diag}(e^{\tilde{\boldsymbol{\gamma}}})(\mathbf{F} + (1/\bar{p}_l)\mathbf{ve}_l^\top)) \leq 0 \, \forall l\}, \tag{14}$$

and define the first order derivative function of the objective function in (13) by

$$f'(\tilde{\boldsymbol{\gamma}}) = \left(w_1 \frac{e^{\tilde{\gamma}_1}}{1 + e^{\tilde{\gamma}_1}}, \dots, w_L \frac{e^{\tilde{\gamma}_L}}{1 + e^{\tilde{\gamma}_L}}\right)^\top. \tag{15}$$

Theorem 2: The optimal value in (13) is achieved only on the boundary of $D(\{\mathbf{F}\})$. A point $\tilde{\boldsymbol{\gamma}}^\star$ is an optimal solution of (13) if and only if the following conditions hold:

$$f'(\boldsymbol{\beta})^\top (\tilde{\boldsymbol{\gamma}} - \boldsymbol{\beta}) \leq 0 \ \forall \boldsymbol{\beta} \in \{f(\boldsymbol{\beta}) = f(\tilde{\boldsymbol{\gamma}}^\star)\} \text{ and } \tilde{\boldsymbol{\gamma}} \in D(\{\mathbf{F}\}). \tag{16}$$

Now, since $D(\{\mathbf{F}\})$ is unbounded, we need to consider a small modification to (13) so that numerical methods that compute the maximum value of convex functions on bounded closed convex sets can be employed [24]. In particular, we consider (13) with additional constraints:

$$D(\{\mathbf{F}\}, K) = \{\tilde{\boldsymbol{\gamma}} \in D(\{\mathbf{F}\}), \quad \tilde{\boldsymbol{\gamma}} \geq -K\mathbf{1}\} \tag{17}$$

for an arbitrarily large $K \gg 1$. Note that $D(\{\mathbf{F}\}, K)$ is compact and convex. The following result in [19] gives the description of the set $D(\{\mathbf{F}\}, K)$.

Lemma 2: Let $\underline{\mathbf{p}} = P(e^{-K}\mathbf{1}) = (e^K \mathbf{I} - \mathbf{F})^{-1}\mathbf{v}$. Then $D(\{\mathbf{F}\}, K) \subseteq \{\tilde{\boldsymbol{\gamma}} \mid \log \boldsymbol{\gamma}(\underline{\mathbf{p}}) \leq \tilde{\boldsymbol{\gamma}} \leq \log \boldsymbol{\gamma}(\bar{\mathbf{p}})\}$.

For the purpose of algorithm design, we replace the set $D(\{\mathbf{F}\})$ in Theorem 2 by $D(\{\mathbf{F}\}, K)$. Since K can be made arbitrarily large, a solution in $D(\{\mathbf{F}\}, K)$ can be made arbitrarily close to $\tilde{\boldsymbol{\gamma}}^\star$ in Theorem 2. In the following, we propose a global optimization algorithm to find the optimal value of (13) on the closed bounded set $D(\{\mathbf{F}\}, K)$.

A. Outer Approximation Algorithm

We leverage nonnegative matrix theory and the outer approximation technique in [24], [25] to compute an extreme point of $D(\{\mathbf{F}\}, K)$ in (17) that yields $\tilde{\boldsymbol{\gamma}}^\star$ in Theorem 2. Our approach is as follows: The feasible region containing the optimal extreme point is first embedded inside a compact polyhedral convex set. Infeasible regions are then successively removed from this initial polyhedral set. This method generates a nested sequence of polyhedrons approximating $D(\{\mathbf{F}\}, K)$ from the *exterior*. It is noteworthy that the computational performance of this method depends on the choice of this initial polyhedron, i.e., the method is effective if this initial polyhedron is a tight fit [24], [25].

To make our problem amenable to outer approximation and to enable a fast computational algorithm (by finding the *tightest* initial polyhedron), we first approximate the convex set $D(\{\mathbf{F}\}, K)$ by a bigger polyhedral convex sets as follows.

- Choose a finite number of points $\boldsymbol{\zeta}_1, \dots, \boldsymbol{\zeta}_M$ on the boundary of $D(\{\mathbf{F}\}, K)$.
- Let $\text{H}_1(\boldsymbol{\xi}), \dots, \text{H}_N(\boldsymbol{\xi}), \boldsymbol{\xi} \in \mathbb{R}^L$ be the N supporting hyperplanes of $D(\{\mathbf{F}\})$. Note that we can have more than one supporting hyperplane at $\boldsymbol{\zeta}_i$, and at most L supporting hyperplanes. So each $\boldsymbol{\xi} \in D(\{\mathbf{F}\}, K)$ satisfies the inequality $H_j(\boldsymbol{\xi}) \leq 0$ for $j = 1, \dots, N$.

- Define the polytope

$$D(\zeta_1, \ldots, \zeta_M, K) = \{\xi \in \mathbb{R}^L, \; -K\mathbf{1} \le \xi \le \log \bar{\gamma},$$
$$\mathrm{H}_j(\xi) \le 0, j = 1, \ldots, N\},$$

(18)

and output this polytope that contains $D(\{\mathbf{F}\}, K)$.

Now, using the polytope $D(\zeta_1, \ldots, \zeta_M, K)$, we have

$$\max_{\tilde{\gamma} \in D(\zeta_1, \ldots, \zeta_M, K)} f(\tilde{\gamma}) \ge \max_{\tilde{\gamma} \in D(\{\mathbf{F}\}, K)} f(\tilde{\gamma}). \tag{19}$$

Furthermore, since $f(\tilde{\gamma})$ is strictly convex, the maximum on the lefthand side in (19) is achieved only at an extreme point of $D(\zeta_1, \ldots, \zeta_M, K)$. Though solving the lefthand side in (19), being a multiextremal problem, is still computationally challenging, it can lead to provably correct solutions under well-defined conditions.

We now show how to efficiently compute the supporting hyperplanes $\mathrm{H}_j, j = 1, \ldots, N$ for $N = L$ in $D(\zeta_1, \ldots, \zeta_M, K)$. To do that, we give a characterization of supporting hyperplanes of $D(\{\mathbf{F}\})$ at a boundary point $\zeta \in \partial D(\{\mathbf{F}\})$ based on the Friedland-Karlin inequalities in [26]. An extension of these inequalities are found in Theorem 6 in Appendix Section I.

Theorem 3: Let $\bar{\mathbf{p}} = (\bar{p}_1, \ldots \bar{p}_L)^\top > \mathbf{0}$ be given. Consider the convex set (14). Let ζ be a boundary point of $\partial D(\{\mathbf{F}\})$. Then $\zeta = \log \gamma(\mathbf{p})$, where $\mathbf{0} \le \mathbf{p} = (p_1, \ldots, p_L)^\top \le \bar{\mathbf{p}}$. The set $\mathcal{B} := \{l \in \{1, \ldots, L\}, \; p_l = \bar{p}_l\}$ is nonempty. For each matrix $\mathbf{B}_l = (\mathbf{F} + (1/\bar{p}_l)\mathbf{ve}_l^\top)$, let $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_L)^\top \in \mathbb{R}^L$ satisfy the condition $\rho(e^{\boldsymbol{\eta}}\mathbf{B}_l) = 1$. Define $\mathrm{H}_l(\zeta)$ as

$$\mathrm{H}_l(\zeta) = \sum_{j=1}^{L} x_j(\mathbf{B}_l) y_j(\mathbf{B}_l)((\zeta)_j - \eta_j). \tag{20}$$

Then $\mathrm{H}_l(\zeta) \le 0$, for $l \in \mathcal{B}$, are the supporting hyperplanes of $D(\{\mathbf{F}\})$ at ζ.

We now show how to choose the boundary points $\zeta_1, \ldots, \zeta_M \in \partial D(\{\mathbf{F}\})$ and to compute the supporting hyperplanes of $D(\{\mathbf{F}\})$ at each ζ_i. Let $\underline{\mathbf{p}} = P(e^{-K}\mathbf{1}) = (\underline{p}_1, \ldots, \underline{p}_L)^\top$ be defined as in Lemma 2. Choose $M_i \ge 2$ equidistant points in each interval $[\underline{p}_i, \bar{p}_i]$, and define

$$p_{j_i, i} = \frac{j_i \underline{p}_i + (M_i - j_i)\bar{p}_i}{M_i} \quad \text{for } j_i = 1, \ldots, M_i, \; i = 1, \ldots, L. \tag{21}$$

Let

$$\mathcal{P} = \left\{ \begin{array}{l} \mathbf{p}_{j_1, \ldots, j_L} = (p_{j_1, 1}, \ldots, p_{j_L, L})^\top, \\ \min(\bar{p}_1 - p_{j_1, 1}, \ldots, \bar{p}_L - p_{j_L, L}) = 0. \end{array} \right\}$$

That is, $\mathbf{p}_{j_1, \ldots, j_L} \in \mathcal{P}$ if and only $\mathbf{p}_{j_1, \ldots, j_L} \not< \bar{\mathbf{p}}$. Then

$$\{\zeta_1, \ldots, \zeta_M\} = \log \gamma(\mathcal{P}).$$

The supporting hyperplanes of $D(\{\mathbf{F}\})$ at each ζ_i are given by Theorem 3. Thus, we have established the L unique hyperplanes $H_l(\xi) \le 0$ for $l = 1, \ldots, L$ in $D(\zeta_1, \ldots, \zeta_M, K)$:

$$\mathrm{H}_l(\xi) = \sum_j (\mathbf{x}(\mathbf{F} + (1/\bar{p}_l)\mathbf{ve}_l^\top) \circ \mathbf{y}(\mathbf{F} + (1/\bar{p}_l)\mathbf{ve}_l^\top))_j(\xi)_j$$
$$+ \log \rho(\mathbf{F} + (1/\bar{p}_l)\mathbf{ve}_l^\top) \le 0, l = 1, \ldots, L. \tag{22}$$

Based on $D(\zeta_1, \ldots, \zeta_L, K)$ in (22), we propose the following algorithm that computes the optimal solution of (9).

Algorithm 1 (Sum Rate Outer Approximation Algorithm):

- **Input.** $D(\zeta_1, \ldots, \zeta_L, K)$ with $\mathrm{H}_l(\xi), \forall l$ given in (22).
- **Output.** Optimal solution $\mathbf{p}^\star$ of (9).

---

1) Compute the vertices of the enclosing linear polyhedron

$D^{(0)}$, described by the set of constraints:

$$\sum_j (\mathbf{x}(\mathbf{F} + (1/\bar{p}_l)\mathbf{ve}_l^\top) \circ \mathbf{y}(\mathbf{F} + (1/\bar{p}_l)\mathbf{ve}_l^\top))_j \tilde{\gamma}_j +$$
$$\log \rho(\mathbf{F} + (1/\bar{p}_l)\mathbf{ve}_l^\top) \le 0,$$

(23)

and $\tilde{\gamma}_l \ge -K$ for all l. Let $V^{(0)}$ be the set of vertices of $D^{(0)}$. Set $k = 1$ and go to Step 2.

2) Iteration k: Solve the problem:

$$\begin{array}{ll} \text{maximize} & \sum_l w_l \log(1 + e^{\tilde{\gamma}_l}) \\ \text{subject to} & \tilde{\gamma}_l \in D^{(k-1)} \end{array} \tag{24}$$

by selecting $\max\{\sum_l w_l \log(1 + e^{\tilde{\gamma}_l}) : v \in V^{(k-1)}\}$. Let $\tilde{\gamma}^k$ be the optimizer to (24).

3) Compute

$$\mathbf{p}^k = \left(\mathbf{I} - \mathrm{diag}(\exp(\tilde{\gamma}^k))\mathbf{F}\right)^{-1} \mathrm{diag}(\exp(\tilde{\gamma}^k))\mathbf{v}. \tag{25}$$

4) If $\mathbf{p}^k \le \bar{\mathbf{p}}$, stop: $\tilde{\gamma}^k$ is the solution to (13) and $\mathbf{p}^k$ is the solution to (9). Otherwise, let

$$J^k = \{l : \log \rho(\mathrm{diag}(\exp(\tilde{\gamma}^k))(\mathbf{F} + (1/\bar{p}_l)\mathbf{ve}_l^\top))$$
$$= \max_{1 \le j \le L} \log \rho(\mathrm{diag}(\exp(\tilde{\gamma}^k))(\mathbf{F} + (1/\bar{p}_j)\mathbf{ve}_j^\top))\}$$

and choose any $j^k \in J^k$.

5) Compute the left eigenvector $\mathbf{y}_{j^k}$ and right (Perron) eigenvector $\mathbf{x}_{j^k}$ of $\mathrm{diag}(\exp(\tilde{\gamma}^k))(\mathbf{F} + (1/\bar{p}_{j^k})\mathbf{ve}_{j^k}^\top)$. Set

$$G_{j^k}^k(\tilde{\gamma}) = \log \rho(\mathrm{diag}(\exp(\tilde{\gamma}^k))(\mathbf{F} + (1/\bar{p}_{j^k})\mathbf{ve}_{j^k}^\top)) +$$
$$\frac{[\exp(\tilde{\gamma}^k) \circ \mathbf{x}_{j^k} \circ \mathbf{y}_{j^k}]^\top (\tilde{\gamma} - \tilde{\gamma}^k)}{\rho(\mathrm{diag}(\exp(\tilde{\gamma}^k))(\mathbf{F} + (1/\bar{p}_{j^k})\mathbf{ve}_{j^k}^\top))}. \tag{26}$$

6) Set $D^{(k)} = D^{(k-1)} \cap \{\tilde{\gamma} : G_{j^k}^k(\tilde{\gamma}) \le 0\}$, $V^{(k)} = \{\text{extreme points of } D^{(k)}\}$.

7) Set $k \leftarrow k + 1$. Go to Step 2.

The following result establishes the convergence of $\mathbf{p}^k$ in Algorithm 1 to the global optimal solution of (9).

Theorem 4: Every limit point of the sequence $\mathbf{p}^k = \left(\mathbf{I} - \mathrm{diag}(\exp(\tilde{\gamma}^k))\mathbf{F}\right)^{-1} \mathrm{diag}(\exp(\tilde{\gamma}^k))\mathbf{v}$ solves (9).

We add the following remarks concerning the implementation and optimality of Algorithm 1.

Remark 2: At Step 5, the gradient of $\log \rho(\mathrm{diag}(\exp(\tilde{\gamma}))(\mathbf{F} + (1/\bar{p}_{j^k})\mathbf{ve}_{j^k}^\top))$ (given by $[\exp(\tilde{\gamma}) \circ \mathbf{x}_{j^k} \circ \mathbf{y}_{j^k}]/\rho(\mathrm{diag}(\exp(\tilde{\gamma}))(\mathbf{F} + (1/\bar{p}_{j^k})\mathbf{ve}_{j^k}^\top)))$ at $\tilde{\gamma}^k$ is used to construct the inequality cut given by (26) that separates $\tilde{\gamma}^k$ (infeasible with respect to (13)) from the feasible constraint set of (13).

Remark 3: A finite number of iterations can be obtained by replacing the stopping rule at Step 4 with stop if $\max_l p_l^k - \bar{p}_l \le \epsilon$ or stop if $\log \rho(\mathrm{diag}(\exp(\tilde{\gamma}^k))(\mathbf{F} + (1/\bar{p}_{j^k})\mathbf{ve}_{j^k}^\top)) \le \epsilon$, where ϵ is a positive error tolerance number.

Remark 4: At Step 6, an online vertex enumeration procedure (for example, see [24]) can be used to speed up the computation of the new vertex set $V^{(k)}$.

Remark 5: Note that every limit point of the sequence $\mathbf{x}_{j^k}$ for any $j^k \in J^k$ converges to the limit point of the sequence $\mathbf{p}^k$ in Algorithm 1.

B. Spectrum-sharing constraints

Cognitive radio power constraints or interference temperature can be readily incorporated in our optimization framework
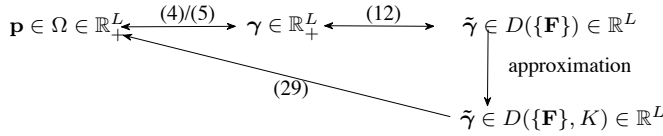
Fig. 2. A summary of the transformation and mapping between the power $\mathbf{p}$, the SIR $\boldsymbol{\gamma}$ and the SIR in the natural base logarithm $\tilde{\boldsymbol{\gamma}}$.

and algorithm by defining an appropriate

$$\Omega = \{\, \mathbf{p} \mid \mathbf{0} \leq \mathbf{p} \leq \bar{\mathbf{p}} \,\} \cap \{\, \mathbf{p} \mid \mathbf{A}\mathbf{p} \leq \bar{\mathbf{p}}_{\mathbf{A}} \,\}, \qquad (27)$$

where $\mathbf{A} \in \mathbb{R}_+^{m \times n}$. We also assume that Ω in (27) is such that the sets $\{\, \mathbf{p} \mid \mathbf{A}\mathbf{p} \leq \bar{\mathbf{p}}_{\mathbf{A}} \,\}$ and $\{\, \mathbf{p} \mid \mathbf{0} \leq \mathbf{p} \leq \bar{\mathbf{p}} \,\}$ are not strictly contained in each other. This implies that $p_l^\star = \bar{p}_l$ for some l. The set $\{\, \mathbf{p} \mid \mathbf{A}\mathbf{p} \leq \bar{\mathbf{p}}_{\mathbf{A}} \,\}$ is general enough to model constraints on the received power of a subset of links or constraints that limit the interference level.

First, we consider the individual power constraints. At Step 3, a feasible point to (9), $\hat{\mathbf{p}}^k$, is easily obtained by using a projection on to the box constraint $[\mathbf{0}, \bar{\mathbf{p}}]$:

$$\hat{p}_l^k = \min\{p_l^k, \bar{p}_l\} \ \forall \ l. \qquad (28)$$

Using Theorem 4, every limit point of the sequence $\hat{\mathbf{p}}^k$ solves (9). Further, it can be shown that the rate vector generated by $\hat{\mathbf{p}}^k$ always lies on the boundary of the rate region.

We next turn to the spectrum-sharing constraints given in (27) by considering the projection of $\mathbf{p}^k$ onto Ω [27]:

$$\begin{aligned} \text{minimize} \quad & \|\hat{\mathbf{p}}^k - \mathbf{p}^k\|_2 \\ \text{subject to} \quad & \hat{\mathbf{p}}^k \in \Omega. \end{aligned} \qquad (29)$$

Theorem 4 can be adapted to show that every limit point of the sequence $\hat{\mathbf{p}}^k$ given by (29) solves (9) with Ω in (27). We summarize the relationship between the sets $D(\{\mathbf{F}\})$, $D(\{\mathbf{F}\}, K)$, established by the transformation (4), (5), (12) and (29) used in Algorithm 1 in Figure 2.

V. APPROXIMATION AND RELAXATION TECHNIQUES

Although Algorithm 1 can find the optimal solution, it remains plausible only for small-to-medium problem size (before the vertex set of the outer approximating polytope reaches a prohibitive size). We now turn to other techniques that yield further insights to solving (9).

A. Inner Approximation

We first state the following result in [26, Theorem 3.1] that can be used to construct an inner approximation of $D(\{\mathbf{F}\})$.

Lemma 3: For any nonnegative matrix $\mathbf{F}$ and positive vector $\boldsymbol{\gamma}$,

$$\rho(\text{diag}(\boldsymbol{\gamma})\mathbf{F}) \leq \max_l \gamma_l \rho(\mathbf{F}). \qquad (30)$$

Equality is achieved in (30) if and only if γ_l are equal for all l.

Applying Lemma 3 to each constraint in (13), a smaller convex subset of $D(\{\mathbf{F}\})$ (replacing the constraints in (13) by restricted box constraints) can be easily obtained. Thus, a restricted problem of (13) and its solution (correspondingly, that of (9)) is given in the following.

Lemma 4: The optimization problem

$$\begin{aligned} \text{maximize} \quad & \sum_l w_l \log(1 + e^{\tilde{\gamma}_l}) \\ \text{subject to} \quad & \tilde{\boldsymbol{\gamma}} \leq \min_l - \log \rho(\mathbf{F} + (1/\bar{p}_l)\mathbf{v}\mathbf{e}_l^\top)\mathbf{1}. \end{aligned} \qquad (31)$$

has a solution that is an extreme point of $D(\{\mathbf{F}\})$ and is given by $\tilde{\boldsymbol{\gamma}} = -\log \rho(\mathbf{F} + (1/\bar{p}_i)\mathbf{v}\mathbf{e}_i^\top)\mathbf{1}$, where $i = \min_l - \log \rho(\mathbf{F} + (1/\bar{p}_l)\mathbf{v}\mathbf{e}_l^\top)$. Then, a feasible power vector to (9) is given by $\mathbf{x}(\mathbf{F} + (1/\bar{p}_i)\mathbf{v}\mathbf{e}_i^\top)$ (unique up to a constant).

Remark 6: Interestingly, the inner approximation in (31) yields the same solution as maximizing the minimum SIR subject to the individual power constraints (cf. Theorem 2 in [17]).

We briefly discuss how to enhance Algorithm 1 using the above result. Inner approximation techniques such as the Polyhedral Annexation method (cf. Chapter 6, [24]) can be applied using the constraint set in Lemma 31 as the initial inner approximation. In brief, the Polyhedral Annexation method generates a sequence of expanding polyhedrons contained in $D(\{\mathbf{F}\})$ that approximates $D(\{\mathbf{F}\})$ from the *interior*. This methodology is *dual* to and complements the outer approximation technique described in Section IV-A. We refer the readers to [24] on standard global optimization algorithms, e.g., branch-and-bound techniques, to combine and accelerate the inner and outer approximation techniques. Interestingly, the inner approximation given by (31) optimally solves (13) when the optimal solution of (13) is such that the SIR's are all equal (cf. conditions under which the equality in Lemma 3 above and Theorem 6 in the appendix holds).

B. Relaxation

We now turn to a relaxation technique that can solve special cases of (9). We define the matrix $\tilde{\mathbf{F}}$ as

$$\tilde{\mathbf{F}} = \mathbf{F} + \text{diag}(\bar{\boldsymbol{\gamma}})^{-1}. \qquad (32)$$

Lemma 5: Let $\mathbf{0} \leq \mathbf{p} \leq \bar{\mathbf{p}}$. Then,

$$\mathbf{p} \geq \text{diag}(\boldsymbol{\gamma}(\mathbf{p}))\tilde{\mathbf{F}}\mathbf{p}, \qquad (33)$$

and

$$\rho(\text{diag}(\boldsymbol{\gamma}(\mathbf{p}))\tilde{\mathbf{F}}) \leq 1. \qquad (34)$$

Next, using (34) and the logarithmic change of variable technique, we consider the following optimization problem:

$$\begin{aligned} \text{maximize} \quad & f(\tilde{\boldsymbol{\gamma}}) \\ \text{subject to} \quad & \log \rho(\text{diag}(e^{\tilde{\boldsymbol{\gamma}}})\tilde{\mathbf{F}} \leq 0, \\ & \tilde{\boldsymbol{\gamma}} \leq \log \bar{\boldsymbol{\gamma}}. \end{aligned} \qquad (35)$$

The following result shows that (35) is a relaxed problem of (13).

Lemma 6: The optimal value of (35) is not less than the optimal value of (9). Furthermore, the optimal solution of (35), $\tilde{\boldsymbol{\gamma}}'$, satisfies $\log \rho(\text{diag}(e^{\tilde{\boldsymbol{\gamma}}'})\tilde{\mathbf{F}}) = 0$.

Lemma 7: If $\mathbf{p}^\star = \bar{\mathbf{p}}$ or $\mathbf{p}^\star$ is such that $p_l^\star = 0$ for some l and $p_j^\star = \bar{p}_j$ for $j \neq l$, then

$$\rho(\text{diag}(\boldsymbol{\gamma}(\mathbf{p}^\star))\tilde{\mathbf{F}}) = 1. \qquad (36)$$

Remark 7: Since (35) is a relaxed problem as compared to (13), both Lemmas 6 and 7 imply that if the optimal solution of (35), $\boldsymbol{\gamma}'$, satisfies $P(\boldsymbol{\gamma}') \leq \bar{\mathbf{p}}$, then $P(\boldsymbol{\gamma}')$ is also the optimal solution of (9).

We obtain the following special case result by combining Lemma 7 and a result in [7].

Corollary 1 (2-user Sum Rate): When $L = 2$ and $\mathbf{w} = (\frac{1}{2}, \frac{1}{2})^\top$, $\mathbf{p}^\star$ is one of the following three vectors: $(\bar{P}_1, 0)^\top$,

$(0, \bar{P}_2)^\top$ or $(\bar{P}_1, \bar{P}_2)^\top$. Furthermore, $\mathbf{p}^\star = (\bar{P}_1, 0)^\top$ only if $\bar{p}_1 \bar{p}_2 F_{12} F_{21} - 0.5 v_1 v_2 - v_1 \bar{p}_2 \geq 0$, and $\mathbf{p}^\star = (0, \bar{P}_2)^\top$ only if $\bar{p}_1 \bar{p}_2 F_{12} F_{21} - 0.5 v_1 v_2 - v_2 \bar{p}_1 \gamma_2 \geq 0$.

Remark 8: From the above special case, the relaxed problem (35) is tight when $\mathbf{p}^\star$ is one of the on-off vectors: $(\bar{P}_1, 0)^\top$, $(0, \bar{P}_2)^\top$ and the corresponding necessary conditions in Corollary 1 are met.

VI. NUMERICAL EXAMPLES

In this section, we provide numerical examples to illustrate the performance of Algorithm 1 in Section IV-A and how the approximation technique in Section V can be applied in solving (9) for two users, i.e., $L = 2$. Consider the channel gain matrix given by

$$\mathbf{G} = \begin{bmatrix} 0.73 & 0.04 \\ 0.03 & 0.89 \end{bmatrix}. \qquad (37)$$

Let the maximum power constraint vector be $\bar{\mathbf{p}} = [1.8 \ 100.5]^\top$ mW and the noise power of each user be 0.1mW. The weight vector is given by $\mathbf{w} = \mathbf{x}(\mathbf{F} + (1/\bar{p}_i)\mathbf{ve}_i^\top) \circ \mathbf{y}(\mathbf{F} + (1/\bar{p}_i)\mathbf{ve}_i^\top)$, where $i = \arg\max_l \rho(\mathbf{F} + (1/\bar{p}_l)\mathbf{ve}_l^\top)$. We set $\epsilon = 1 \times 10^{-8}$ and $K = 100$ in Algorithm 1. Now, the optimal solution is achieved at the equal SIR allocation for the two users (equivalent to maximizing the minimum SIR problem), where $\mathbf{p}^\star = \mathbf{x}(\mathbf{F} + (1/\bar{p}_i)\mathbf{ve}_i^\top) = [1.8000 \ 1.442]^\top$ mW. Thus, the optimal sum rate is 2.2336 nats/symbol.

At the first iteration, the vertices of $V^{(0)}$ are $(-100.0, -100.0)$, $(-100.0, 103.6279)$, $(39.4757, -100.0)$, $(0.9959, 5.1941)$. The vertex of $V^{(0)}$ having the maximum objective function value is $\tilde{\boldsymbol{\gamma}}^1 = (39.4757 \ -100.0)$ and $\mathbf{p}^1 = [1.909 \times 10^{19} \ 0]^\top$. At Step 4, evaluating each constraint function at $\tilde{\boldsymbol{\gamma}}^1 = (39.4757 \ -100.0)$, we have $\log(\mathrm{diag}(\tilde{\boldsymbol{\gamma}}^1)(\mathbf{F} + (1/\bar{p}_1)\mathbf{ve}_1^\top)) = 36.9$ and $\log(\mathrm{diag}(\tilde{\boldsymbol{\gamma}}^1)(\mathbf{F} + (1/\bar{p}_2)\mathbf{ve}_2^\top)) = -33.4$. We thus choose $j^1 = 1$, and a new constraint is obtained at Step 5 as $\tilde{\gamma}_1 \leq 2.5757$.

Therefore, we have $D^1 = D^0 \cap \{\tilde{\boldsymbol{\gamma}} : \tilde{\gamma}_1 \leq 2.5757\}$ at Step 6. The vertices of D^1 are $(-100.0, -100.0)$, $(-100.0, 103.6279)$, $(2.5757, -100.0)$, $(2.5757, 0.8754)$, $(0.9959, 5.1941)$. We then proceed to Step 2 to find the optimal vertex of D^1.

After twenty nine more iterations, we arrive at the power vector $\mathbf{p}^{29} = [1.8000 \ 1.442]^\top$ mW, and $\max_l p_l^{29} - \bar{p}_l = 5.5485 \times 10^{-9}$, whereupon Algorithm 1 terminates. Figure 3 illustrates the evolution of the approximating polyhedron, where Figure 3(a) and (b) show the initial enclosing polyhedron $D^{(0)}$ and the polyhedron $D^{(14)}$ at the fourteen iteration, respectively. We observe that, by the fourteen iteration, $D^{(14)}$ provides a relatively good approximation to the feasible region $D(\{\mathbf{F}\}, K)$. Figure 4 (a) and (b) illustrates the convergence of the rate vectors generated by $[\log(1 + \gamma_1(\mathbf{p}^k)) \ \log(1 + \gamma_2(\mathbf{p}^k))]^\top$ and $[\log(1 + \gamma_1(\hat{\mathbf{p}}^k)) \ \log(1 + \gamma_2(\hat{\mathbf{p}}^k))]^\top$ respectively. As shown on Figure 4 (b), the rate vector converges close to the optimal rate vector by the tenth iteration. Figure 4 also illustrates the optimal rate vector.

Note that the inner approximation technique, i.e., solving (31) in Lemma 4 yields the optimal solution (solved by one optimization problem in contrast to a modest number of optimization problems in Algorithm 1).

TABLE I
A COMPARISON OF THE TYPICAL CONVERGENCE AND COMPLEXITY STATISTICS OF ALGORITHM 1 WITH THE PROBLEM SIZE. THE CPU TIME IS COMPUTED BASED ON AN IMPLEMENTATION ON A 64-BIT SUN/SOLARIS 10 (SUNOS 5.10) COMPUTER.

Problem size	Maximal number of generated vertices	Number of iterations	CPU time (minutes)
2	15	12	0.062
4	139	760	4.1
6	14022	1238	83
8	283681	1968	468

Next, we repeat our previous experiment but the parameters $\bar{\mathbf{p}} = [100.8 \ 100.5]^\top$ mW and $\bar{\mathbf{p}} = [300.8 \ 300.5]^\top$ mW are used instead. Figure 5 (a) and (b) illustrates the convergence for these two different maximum power constraints respectively. Unlike the previous experiment, the optimal solution in these two scenarios are such that one user is transmitting at maximum power and the other user does not transmit. As illustrated, the rate vector computed by Algorithm 1 already approaches close to the optimal rate vector by the seventh iteration, and Algorithm 1 terminates at the twenty-eigth iteration for both scenarios (with $\epsilon = 1 \times 10^{-8}$).

Lastly, we vary the problem size, i.e., the number of users, and evaluate the number of iterations, the maximal number of vertices upon exit and the CPU time required to run Algorithm 1 by setting $\epsilon = 1 \times 10^{-4}$. Algorithm 1 is implemented in Matlab and runs on a 64-bit Sun/Solaris 10 (SunOS 5.10) computer. Table I compares the average statistics on different problem size. As illustrated, Algorithm 1 is effective for a network of small-to-medium size. For a large number of users, the complexity of enumerating the vertices at Step 2 can grow prohibitively large as the number of iterations increases.

VII. EXTENSION TO MULTIPLE CARRIERS AND INTERFERENCE SUPPRESSION RECEIVERS

A. Multiple Carrier Model

In this section, we extend our previous results in Section III to a multiuser multiple carrier model, where a common spectrum is divided into K frequency tones. For illustrative purpose, we assume the standard synchronous orthogonal frequency-division modulation, where orthogonality among subchannels of the intended signal and the subchannels of the interference signal in different frequency tones is maintained.[3] Thus, transmissions can be modeled independently on each tone. The achievable rate at tone k can be modeled as [22]:

$$\log\left(1 + \frac{G_{ll,k} p_{l,k}}{\sum_{j \neq l} G_{lj,k} p_{j,k} + n_{l,k}}\right). \qquad (38)$$

The total data rate for each user is then obtained by adding its transmitted data rate over all the K tones. The total power budget of the lth user is constrained (across all K tones) by

$$\sum_{k=1}^{K} p_{l,k} \leq \bar{p}_l. \qquad (39)$$

It is easy to see that our previous model in Section II is a special case of this multiple carrier model assuming standard synchronous discrete multi-tone (DMT) modulation when $K = 1$.

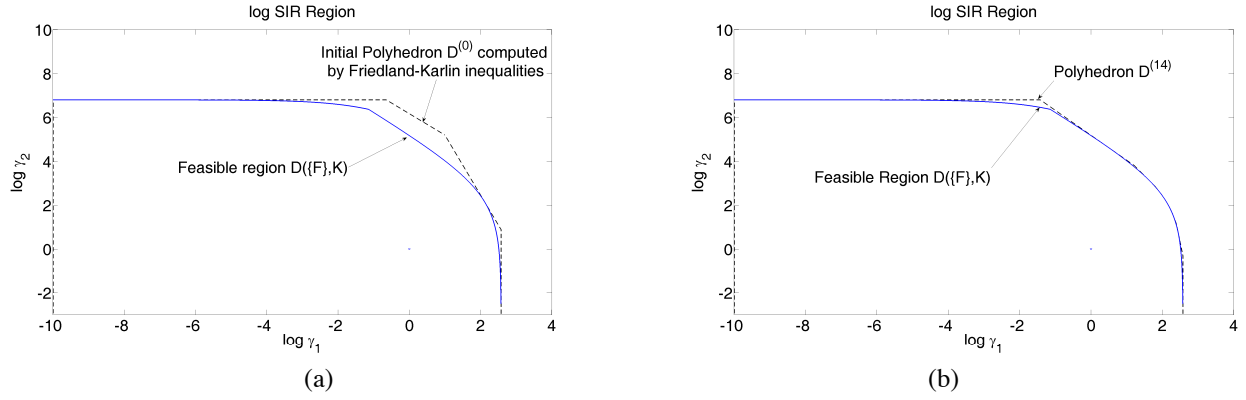[3]This is also known as discrete multitone modulation in the DSL context.

Fig. 3. Illustration of the convergence of Algorithm 1 in the $\tilde{\gamma}$ region with (a) the initial polyhedron given by $D^{(0)}$ and (b) the polyhedron $D^{(14)}$ at the fourteen iteration.
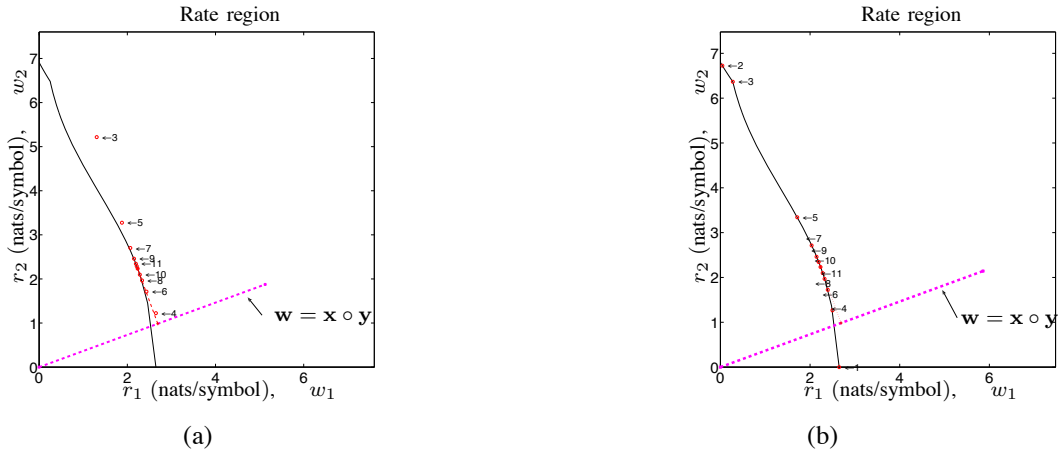

Fig. 4. Illustration of the convergence of Algorithm 1 in (a) the rate vector generated by $\mathbf{p}^k$ and in (b) the rate vector generated by $\hat{\mathbf{p}}^k$. Only the rate vectors obtained up to the eleventh iteration are shown.
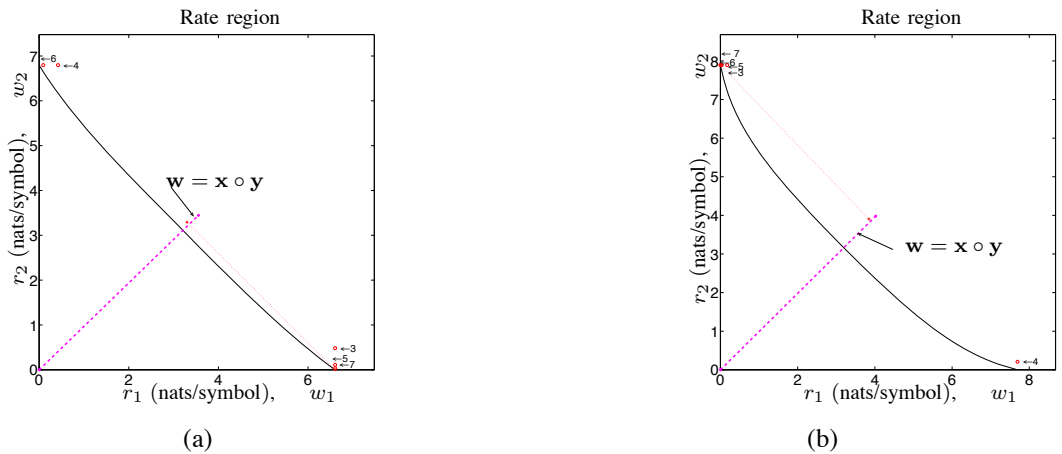

Fig. 5. Illustration of the convergence of Algorithm 1 on (a) a rate region with $\bar{\mathbf{p}} = [100.8 \ 100.5]^\top$ and on (b) a rate region with $\bar{\mathbf{p}} = [300.8 \ 300.5]^\top$. Only the rate vectors obtained up to the seventh iteration are indicated.

For brevity of notations, we define $\mathbf{p} \in \mathbb{R}_+^{(K \times L)}$ as a vector that stacks the $K \times L$ power allocation lined-up according to tones of all users. For example, if $L = 2$ and $K = 2$, then $\mathbf{p} = (p_{1,1}, p_{1,2}, p_{2,1}, p_{2,2})^\top$. Similarly, we define $\boldsymbol{\gamma} \in \mathbb{R}_+^{(K \times L)}$ as the SIR allocation, the matrix $\mathbf{F}$ having entries:

$$F_{ij,k} = \begin{cases} 0, & \text{if } i = j, \ i, j \in \{1, \ldots, L\}, \ k \in \{1, \ldots, K\} \\ \frac{G_{ij,k}}{G_{ii,k}}, & \text{if } i \neq j, \ i, j \in \{1, \ldots, L\}, \ k \in \{1, \ldots, K\} \end{cases} \tag{40}$$

and the vector

$$\mathbf{v} = \left(\frac{n_{1,1}}{G_{11,1}}, \frac{n_{1,2}}{G_{11,2}}, \ldots, \frac{n_{1,K}}{G_{11,K}}, \frac{n_{2,1}}{G_{22,1}}, \ldots, \frac{n_{L,K}}{G_{LL,K}} \right)^\top. \tag{41}$$

Note that asynchronous transmission can result in the lth user at tone k having interference from the power allocation at neighboring tones of tone k. Hence, in general, $\mathbf{F}$ is not a block diagonal matrix. However, assuming synchronous transmission of all users, $\mathbf{F}$ is block diagonal.

As in the previous, the lth user is given a positive weight parameter w_l to reflect its long-term priority. The problem of maximizing the weighted sum rate in a Gaussian channel with K frequency tones is given by:

$$\max_{\sum_{k=1}^K p_{l,k} \leq \bar{p}_l \ \forall l} \sum_{l=1}^L w_l \sum_{k=1}^K \log(1 + \gamma_{l,k}(\mathbf{p})), \tag{42}$$

where $\gamma_{l,k} = G_{ll,k} p_{l,k} / (\sum_{j \neq l} G_{lj,k} p_{j,k} + n_{l,k})$. In order to be consistent with our previous results, we shall consider the following equivalent problem:

$$\max_{\sum_{k=1}^K p_{l,k} \leq \bar{p}_l \ \forall l} \sum_{l=1}^L \sum_{k=1}^K \tilde{w}_{l,k} \log(1 + \gamma_{l,k}(\mathbf{p})). \tag{43}$$

where $\tilde{w}_{l,k} = (w_l / K)$ for all $l \in \{1, \ldots, L\}, k \in \{1, \ldots, K\}$, and $\tilde{\mathbf{w}} \in \mathbb{R}_+^{(K \times L)}$ is a probability vector.

As in Section III, it is instrumental to consider an equivalent reformulation of (43) in terms of the SIR variable.

Theorem 5: The optimal value of (43) is equal to the optimal value of the problem:

maximize $\sum_l \sum_k \tilde{w}_{lk} \log(1 + \gamma_{l,k})$

subject to $\rho(\text{diag}(\boldsymbol{\gamma})(\mathbf{F} + (1/\bar{p}_l) \sum_{k=1}^K \mathbf{v}\mathbf{e}_{((l-1)\times K+k)}^\top)) \leq 1$
$\forall l,$

variables: $\gamma_{l,k}, \quad \forall l, \ \forall k.$
$$\tag{44}$$

Now, $\boldsymbol{\gamma}^\star$ is an optimal solution of (44) if and only if $P(\boldsymbol{\gamma}^\star)$ is an optimal solution $\mathbf{p}^\star$ of the problem (43). In particular, any optimal solution $\boldsymbol{\gamma}^\star$ satisfies

$$\rho\left(\text{diag}(\boldsymbol{\gamma}^\star) \left(\mathbf{F} + (1/\bar{p}_l) \sum_{k=1}^K \mathbf{v}\mathbf{e}_{((l-1)\times K+k)}^\top \right) \right) = 1 \tag{45}$$

for some integer $l \in [1, L]$.

Using Theorem 5, it is straightforward to extend Algorithm 1 in Section IV-A to design optimal spectrum management schemes for multiple carrier channels.

We add the following brief discussion to get more intuition on the above theorem. Our techniques are general enough to deal with asynchronous transmission, where the powers in neighboring tones interfere with one another. This results in intercarrier interference (ICI) and tone coupling in (38); as mentioned earlier, ICI effects can be captured by $\mathbf{F}$ in (40), where the optimal solution to (43) is then characterized by the spectral radius of an appropriately constructed nonnegative matrix. This means, the appropriate choice of supporting

hyperplanes permits decoupling of both users and frequency tones in the SIR domain.

B. Receivers with Interference Suppression

Instead of treating multiuser interference as noise, more sophisticated receiver techniques can improve the data rate performance. For example, multiuser detection techniques that decode and cancel the signals of other users or using receive beamformers at the user increases the reliability of decoding the desired signal. The receiver techniques can also be integrated with the different cognitive spectrum sensing techniques [4], [5] that adapt the effective channel.

We briefly discuss how some of these different receiver techniques can be incorporated into our problem formulation and leave the details for a future work. In general, the various interference suppression techniques affect the spectra (Perron-Frobenius eigenvalues and eigenvectors) of the nonnegative matrices considered in this paper. For example, when multiuser detection is considered, i.e., user decoding is first ordered based on the received SNR, the resultant matrix $\mathbf{F}$ in (2) is then upper triangular. When transmit or receive beamformers are considered, the effective channel gain and cross interference, i.e., the entries of $\mathbf{F}$, become functions of beamformers. Let the set $\mathcal{F}$ be the set of all possible nonnegative matrices for a particular receive strategy with interference suppression. One possible way to optimize over $\mathcal{F}$ is to choose the matrix $\mathbf{F} \in \mathcal{F}$ such that it solves

$$\min_{\mathbf{F} \in \mathcal{F}} \max_l \ \rho(\text{diag}(\boldsymbol{\gamma})(\mathbf{F} + (1/\bar{p}_l)\mathbf{v}\mathbf{e}_l^\top)). \tag{46}$$

Algorithm 1 can then be used to optimize the power by using the optimal matrix $\mathbf{F}$ that solves (46).

VIII. CONCLUSION

We studied the nonconvex NP-hard problem of weighted sum rate maximization in a multiuser Gaussian channel that models a cognitive wireless network with affine power constraints, e.g., interference temperature constraints. Using tools from nonnegative matrix theory, in particular the Perron-Frobenius Theorem and the Friedland-Karlin inequalities, we obtained necessary and sufficient conditions to optimal power allocation and proposed a computationally fast global optimization algorithm. In particular, we designed an outer approximation technique that exploits the spectra of specially crafted nonnegative matrices to accelerate computation. We also studied an inner approximation method and a relaxed problem that provided insights to special cases. Our techniques and algorithm can be extended to the multiple carrier model, e.g., an OFDM cognitive wireless system, and systems with receivers having interference suppression capability.

ACKNOWLEDGEMENT

APPENDIX

A. Proof of Theorem 2

The objective function in (13) is a strictly convex function in $\tilde{\boldsymbol{\gamma}} \in \mathbb{R}^L$. Hence, the optimal value of (13) is achieved exactly

on an extreme point of the closed unbounded set specified in (13). It may happen that some coordinate of the extreme point (optimal solution) are $-\infty$. Since (13) is a strictly convex maximization problem, the optimality conditions given by (16) can be obtained (cf. [28]).

B. Proof of Theorem 3

Let $\mathbf{p} = P(e^{\boldsymbol{\zeta}})$. Since the set $\mathcal{B}$ is nonempty, we have $\rho(e^{\boldsymbol{\zeta}}\mathbf{B}_l) = 1$ if and only if $p_l = \bar{p}_l$. Hence, $\boldsymbol{\zeta}$ lies exactly at the intersection of the hypersurfaces $\log \rho(e^{\boldsymbol{\zeta}}\mathbf{B}_l) = 0, l \in \mathcal{B}$. Theorem 7 in Appendix Section I implies that the supporting hyperplanes of $D(\{\mathbf{F}\})$ at $\boldsymbol{\zeta}$ are $\mathrm{H}_l(\boldsymbol{\xi}) \leq 0$ for $l \in \mathcal{B}$.

C. Proof of Theorem 4

The proof of convergence of $\tilde{\boldsymbol{\gamma}}^k$ to the optimal solution of (13) by cutting plane methodology can be found in [24]. Since the map $\mathbf{p}^k = \left(\mathbf{I} - \mathrm{diag}(\exp(\tilde{\boldsymbol{\gamma}}^k))\mathbf{F}\right)^{-1} \mathrm{diag}(\exp(\tilde{\boldsymbol{\gamma}}^k))\mathbf{v}$ is bijective, the limit point of $\mathbf{p}^k$ solves (9). This proves the theorem.

D. Proof of Lemma 4

Lemma 4 is proved by using the fact that the objective function in (31) is monotonically increasing and the equality condition of Lemma 3. Since all users achieve equal SIR or more precisely, the maximum of the minimum SIR, the optimal power is given by $\mathbf{x}(\mathbf{F} + (1/\bar{p}_l)\mathbf{v}\mathbf{e}_i^{\top})$, unique up to a constant (cf. Theorem 2 in [17]).

E. Proof of Lemma 5

Since $0 \leq p_l \leq \bar{p}_l$, we have $\tilde{F}_{ll}p_l \leq v_l$ for all l. By the definition of $\boldsymbol{\gamma}(\mathbf{p})$ in (4), we have $p_l = \gamma_l(\mathbf{p}) \left(v_l + \sum_{j \neq l} F_{lj}p_j\right)$, which, together with the definition of $\tilde{\mathbf{F}}$, we have (33). To prove that the inequality (34) holds, we note that it is a consequence of the Wielandt's characterization of the spectral radius of an irreducible matrix [20]. Indeed, if $\mathbf{p} > \mathbf{0}$, i.e., all the coordinates of $\mathbf{p}$ are positive, then $\boldsymbol{\gamma}(\mathbf{p}) > 0$. Hence, $\mathrm{diag}(\boldsymbol{\gamma}(\mathbf{p}))\tilde{\mathbf{F}}$ is a positive matrix. Then, by the Wielandt's max-min characterization of the spectral radius, we have

$$\rho(\mathrm{diag}(\boldsymbol{\gamma}(\mathbf{p}))\tilde{\mathbf{F}}) \leq \max_{l=1,\ldots,L} \frac{(\mathrm{diag}(\boldsymbol{\gamma}(\mathbf{p}))\tilde{\mathbf{F}}\mathbf{p})_l}{p_l} \leq 1.$$

Observe next that if $p_l = 0$, then $\boldsymbol{\gamma}(\mathbf{p})_l = 0$. So if some of $p_l = 0$, then $\rho(\mathrm{diag}(\boldsymbol{\gamma}(\mathbf{p}))\tilde{\mathbf{F}})$ is the spectral radius of the maximal positive submatrix of $\mathrm{diag}(\boldsymbol{\gamma}(\mathbf{p}))\tilde{\mathbf{F}}$. Apply Wielandt's characterization to this positive submatrix to deduce (34).

F. Proof of Lemma 6

In view of (34), we see that the optimal value in (35) is achieved on a bigger set than that in (9). Since $f(\tilde{\boldsymbol{\gamma}})$ and $\log \rho(\mathrm{diag}(e^{\tilde{\boldsymbol{\gamma}}})\tilde{\mathbf{F}})$ increase with $\tilde{\boldsymbol{\gamma}}$, this proves the lemma.

G. Proof of Lemma 7

The definition of $\tilde{\mathbf{F}}$ implies (36) for $\mathbf{p}^\star = \bar{\mathbf{p}}$. Assume now that $p_l^\star = 0$ for some l. Then $\gamma_l(\mathbf{p}^\star) = 0$ for some l. Then, the lth row of $\mathrm{diag}(\boldsymbol{\gamma}(\mathbf{p}^\star))\tilde{\mathbf{F}}$ is zero. Let $F_{(l)}$ be the submatrix of $\mathbf{F}$ obtained by deleting the lth row and column. Let $\boldsymbol{\gamma}_{(l)}$ be the vector obtained from $\boldsymbol{\gamma}$ by deleting the lth coordinate. Hence, the characteristic polynomial of $\mathrm{diag}(\boldsymbol{\gamma})\mathbf{F}$, $\det(x\mathbf{I} - \mathrm{diag}(\boldsymbol{\gamma})\mathbf{F}$, is equal to $x \det(x\mathbf{I} - \mathrm{diag}(\boldsymbol{\gamma}_{(l)})F_{(l)})$. Therefore, $\rho(\mathrm{diag}(\boldsymbol{\gamma})\mathbf{F}) = \rho(\mathrm{diag}(\boldsymbol{\gamma}_{(l)})F_{(l)})$. Continuing in this manner, we deduce the lemma.

H. Proof of Corollary 1

The first part of Corollary 1 is a result of [7]. To prove the second part, we note that, using the mapping in (5), maximizing $w_1 \log(1 + \gamma_1) + w_2 \log(1 + \gamma_2)$ or equivalently, $\gamma_1 + \gamma_2 + \gamma_1\gamma_2$, subject to $\rho(\mathrm{diag}(\boldsymbol{\gamma})\tilde{\mathbf{F}}) = 1$ (from (36) in Lemma 7) is equivalent to the following linear program:

$$
\begin{aligned}
\text{maximize} \quad & (\bar{p}_1\bar{p}_2 F_{12}F_{21} - 0.5v_1v_2)(\gamma_1 + \gamma_2) - v_1\bar{p}_2\gamma_1 \\
& - v_2\bar{p}_1\gamma_2 \\
\text{subject to} \quad & \boldsymbol{\gamma} \leq \bar{\boldsymbol{\gamma}}.
\end{aligned}
$$
(47)

As a linear program with a simple box constraint, (47) has an optimal solution that is an extreme point. If $\mathbf{F}$ is a non-zero matrix, the solution is $(\bar{\gamma}_1, 0)^{\top}$ only if $\bar{p}_1\bar{p}_2 F_{12}F_{21} - 0.5v_1v_2 - v_1\bar{p}_2 \geq 0$, and $(0, \bar{\gamma}_2)^{\top}$ only if $\bar{p}_1\bar{p}_2 F_{12}F_{21} - 0.5v_1v_2 - v_2\bar{p}_1\gamma_2 \geq 0$. Using (5) and focusing on the feasible solution to (9), Corollary 1 is proved.

I. Results based on Friedland-Karlin inequalities

We state some nonnegative matrix theory results from [26] and [19] that are used in this paper. The following result extends [26, Theorem 3.1]:

Theorem 6: Let $\mathbf{A} \in \mathbb{R}_+^{L \times L}$ be an irreducible matrix. Assume that $\mathbf{x}(\mathbf{A}) = (x_1(\mathbf{A}), \ldots, x_L(\mathbf{A}))^{\top}$, $\mathbf{y}(\mathbf{A}) = (y_1(\mathbf{A}), \ldots, y_L(\mathbf{A}))^{\top} > \mathbf{0}$ are the Perron right and left eigenvectors of $\mathbf{A}$ respectively, normalized such that $\mathbf{x}(\mathbf{A}) \circ \mathbf{y}(\mathbf{A})$ is a probability vector. Suppose that $\boldsymbol{\gamma}$ is a nonnegative vector. Then

$$\rho(\mathbf{A}) \prod_l \gamma_l^{(\mathbf{x}(\mathbf{A}) \circ \mathbf{y}(\mathbf{A}))_l} \leq \rho(\mathrm{diag}(\boldsymbol{\gamma})\mathbf{A}).$$
(48)

If $\boldsymbol{\gamma}$ is a positive vector then equality holds if and only if all γ_l are equal. Furthermore, for any positive vector $\mathbf{z} = (z_1, \ldots, z_L)^{\top}$, the following inequality holds:

$$\rho(\mathbf{A}) \leq \prod_{l=1}^{L} \left(\frac{(\mathbf{A}\mathbf{z})_l}{z_l}\right)^{(\mathbf{x}(\mathbf{A}) \circ \mathbf{y}(\mathbf{A}))_l}.$$
(49)

If $\mathbf{A}$ is an irreducible matrix with positive diagonal elements, then equality holds in (49) if and only if $\mathbf{z} = t\mathbf{x}(\mathbf{A})$ for some positive t.

The following result gives an interpretation of the inequality (48) in terms of the supporting hyperplane of the convex function $\log \rho(e^{\boldsymbol{\xi}}\mathbf{B})$, where $\mathbf{B} \in \mathbb{R}_+^{L \times L}$ is irreducible and $\boldsymbol{\xi} \in \mathbb{R}^L$.

Theorem 7: Let $\mathbf{B} \in \mathbb{R}_+^{L \times L}$ be an irreducible matrix. Let $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_L)^{\top} \in \mathbb{R}^L$ satisfy the condition $\rho(e^{\boldsymbol{\eta}}\mathbf{B}) = 1$. Denote $\mathbf{A} = e^{\boldsymbol{\eta}}\mathbf{B}$ and assume that $\mathbf{x}(\mathbf{A}) = (x_1(\mathbf{A}), \ldots, x_L(\mathbf{A}))^{\top}, \mathbf{y}(\mathbf{A}) = (y_1(\mathbf{A}), \ldots, y_L(\mathbf{A}))^{\top} > \mathbf{0}$

are the Perron right and left eigenvectors of $\mathbf{A}$ respectively, normalized such that $\mathbf{x}(\mathbf{A}) \circ \mathbf{y}(\mathbf{A})$ is a probability vector. Let

$$H(\boldsymbol{\xi}) = \sum_{l=1}^{L} x_l(\mathbf{A}) y_l(\mathbf{A})(\xi_l - \eta_l). \qquad (50)$$

Then $H(\boldsymbol{\xi}) \leq 0$ is the unique supporting hyperplane to the convex set $\log \rho(e^{\boldsymbol{\xi}} \mathbf{B}) \leq 0$ at $\boldsymbol{\xi} = \boldsymbol{\eta}$.

REFERENCES

[1] M. Chiang, C. W. Tan, D. P. Palomar, D. O'Neill, and D. Julian. Power control by geometric programming. *IEEE Trans. on Wireless Communications*, 6(7):2640–2651, July 2007.

[2] M. Chiang, P. Hande, T. Lan, and C. W. Tan. Power control in wireless cellular networks. *Foundations and Trends in Networking*, 2(4):381–533, 2008.

[3] R. Etkin, A. P. Parekh, and D. Tse. Spectrum sharing in unlicensed bands. *IEEE Journal on Selected Areas in Communications*, 25(3):517–528, 2007.

[4] S. Haykin. Cognitive radio: brain-empowered wireless communications. *IEEE Journal on Selected Areas in Communications*, 23(2):201–220, 2005.

[5] Q. Zhao and B. M. Sadler. A survey of dynamic spectrum access. *IEEE Signal Processing Magazine*, 24(3):79–89, 2007.

[6] S. A. Jafar, S. Srinivasa, I. Maric, and A. Goldsmith. Breaking spectrum gridlock with cognitive radios: An information theoretic perspective. *Proc. of the IEEE*, 97(5):894–914, 2009.

[7] M. Ebrahimi, M. A. Maddah-Ali, and A. K. Khandani. Power allocation and asymptotic achievable sum-rates in single-hop wireless networks. *Proc. of IEEE 40th CISS*, 2006.

[8] P. Viswanath and D. N. C. Tse. Sum capacity of the vector gaussian broadcast channel and uplink-downlink duality. *IEEE Trans. on Information Theory*, 49(8):1912–1921, 2003.

[9] Y. H. Lin and R. L. Cruz. Power control and scheduling for interfering links. *Proc. of IEEE Information Theory Workshop*, 2004.

[10] Z.-Q. Luo and Z. Zhang. Dynamic spectrum management: Complexity and duality. *IEEE Journal on Selected Areas in Signal Processing*, 2(1):57–73, 2008.

[11] W. Yu, G. Ginis, and J. Cioffi. Distributed multiuser power control for digital subscriber lines. *IEEE Journal on Selected Areas in Communications*, 20(5):1105–1115, 2002.

[12] M. Gastpar. On capacity under receive and spatial spectrum-sharing constraints. *IEEE Trans. on Information Theory*, 53(2):471–487, 2007.

[13] 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. on Signal Processing*, 58(6):3251–3271, 2010.

[14] R. Cendrillon, W. Yu, M. Moonen, J. Verlinden, and T. Bostoen. Optimal multiuser spectrum balancing for digital subscriber lines. *IEEE Trans. on Communications*, 54(5):922–933, 2006.

[15] P. Tsiaflakis, M. Diehl, and M. Moonen. Distributed spectrum management algorithms for multiuser DSL networks. *IEEE Trans. on Signal Processing*, 56(10):4825–4843, 2008.

[16] M. Chiang. Balancing transport and physical layers in wireless multihop networks: Jointly optimal congestion control and power control. *IEEE Journal on Selected Areas in Communications*, 23(1):104–116, 2005.

[17] C. W. Tan, M. Chiang, and R. Srikant. Fast algorithms and performance bounds for sum rate maximization in wireless networks. *Proc. of IEEE Infocom*, 2009.

[18] 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 SIR. *Proc. of IEEE ISIT*, 2009.

[19] C. W. Tan, S. Friedland, and S. H. Low. Nonnegative matrix inequalities and their application to nonconvex power control optimization. *SIAM J. on Matrix Analysis and Applications*, in submission.

[20] A. Berman and R. J. Plemmons. *Nonnegative Matrices in the Mathematical Sciences*. Academic Press, USA, 1st edition, 1979.

[21] G. J. Foschini and Z. Miljanic. A simple distributed autonomous power control algorithm and its convergence. *IEEE Trans. on Vehicular Technology*, 42(4):641–646, 1993.

[22] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. John Wiley & Sons, USA, 1991.

[23] J. F. C. Kingman. A convexity property of positive matrices. *Proceedings of the American Mathematical Society*, 12(2):283–284, 1961.

[24] H. Tuy. *Convex analysis and global optimization*. Kluwer Academic Publishers, 1998.

[25] J. E. Kelley. The cutting plane method for solving convex programs. *Journal of the SIAM*, 8(4):703–712, 1960.

[26] S. Friedland and S. Karlin. Some inequalities for the spectral radius of non-negative matrices and applications. *Duke Mathematical Journal*, 42(3):459–490, 1975.

[27] D. P. Bertsekas. *Nonlinear Programming*. Athena Scientific, Belmont, MA, USA, 2nd edition, 2003.

[28] A. S. Strekalovskii. On the global extremum problem. *Soviet Math, Doklady*, 292(5):1062–1066, 1987.

Chee Wei Tan (M'08) received the M.A. and Ph.D. degree in Electrical Engineering from Princeton University, Princeton, NJ, in 2006 and 2008, respectively. He is an Assistant Professor at City University Hong Kong. Previously, he was a Postdoctoral Scholar at the California Institute of Technology (Caltech), Pasadena, CA. His research interests are in wireless and broadband communications, signal processing and nonlinear optimization.

Shmuel Friedland received all his degrees in Mathematics from Israel Institute of Technology,(IIT), Haifa, Israel: B.Sc in 1967, M.Sc. in 1969, D.Sc. in 1971.

He held Postdoc positions in Weizmann Institute of Science, Israel; Stanford University; IAS, Princeton. From 1975 to 1985, he was a member of Institute of Mathematics, Hebrew U., Jerusalem, and was promoted to the rank of Professor in 1982. Since 1985 he is a Professor at University of Illinois at Chicago. He was a visiting Professor in University of Wisconsin; Madison; IMA, Minneapolis; IHES, Bures-sur-Yvette; IIT, Haifa; Berlin Mathematical School, Berlin.

Friedland contributed to the following fields of mathematics: one complex variable, matrix and operator theory, numerical linear algebra, combinatorics, ergodic theory and dynamical systems, mathematical physics, mathematical biology, algebraic geometry. He authored about 170 papers, with many known coauthors, including one Fields Medal winner. He received the first Hans Schneider prize in Linear Algebra, jointly with M. Fiedler and I. Gohberg, in 1993. He was awarded recently a smoked salmon for solving the set-theoretic version of the salmon problem: http://www.dms.uaf.edu/~eallman. For more details on Friedland's vita and research, see http://www.math.uic.edu/~friedlan.
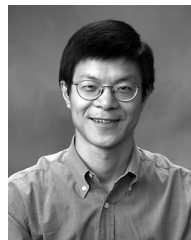
Steven H. Low (F'08) received the B.S. degree from Cornell University, Ithaca, NY, and the Ph.D. degree from the University of California, Berkeley, both in electrical engineering.

He is a Professor with the Computer Science and Electrical Engineering departments at California Institute of Technology, Pasadena, and an Adjunct Professor with Swinburne University,Melbourne, Australia. Prior to that, he was with AT&T Bell Laboratories, Murray Hill, NJ, and the University of Melbourne, Melbourne, Australia.

Prof. Low was a co-recipient of the IEEE Bennett Prize Paper Award in 1997 and the 1996 R&D 100 Award. He was a member of the Networking and Information Technology Technical Advisory Group for the US Presidents Council of Advisors on Science and Technology (PCAST) from 2006 to 2007. He was on the Editorial Board of the IEEE/ACM TRANSACTIONS ON NETWORKING from 1997 to 2006 and the Computer Networks Journal from 2003 to 2005. He is currently on the editorial boards of Computing Surveys and Foundations and Trends in Networking and is a Senior Editor of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATION.