Spectrum Management in Multiuser Cognitive Wireless Networks: Optimality and Algorithm

Chee Wei Tan\(^1\), Shmuel Friedland\(^2\) and Steven H. Low\(^3\)

\(^1\) City University of Hong Kong \quad \(^2\) University of Illinois at Chicago \quad \(^3\) California Institute of Technology

Abstract— Spectrum management is used to improve performance in multiuser communication systems, e.g., cognitive radio or femtocell networks, where multiuser interference can lead to throughput 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 throughput 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 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 and their 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 vectors, italics denote scalars, and \(\mathbf{u} \geq \mathbf{v}\) denotes componentwise

---

\(^1\) 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.
inequality between vectors \( \mathbf{u} \) and \( \mathbf{v} \). We also let \( (\mathbf{By})_l \) denote the \( l \)th element of \( \mathbf{By} \). Let \( \mathbf{x} \circ \mathbf{y} \) denote the Schur product of the vectors \( \mathbf{x} \) and \( \mathbf{y} \), i.e., \( \mathbf{x} \circ \mathbf{y} = [x_1y_1, \ldots, x_Ly_L]^\top \). 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 \( \mathbf{x}(\mathbf{F}) \geq 0 \) and \( \mathbf{y}(\mathbf{F}) \geq 0 \) (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 \( \mathbf{x}(\mathbf{F}), \mathbf{y}(\mathbf{F}) > 0 \) [20]. We will assume the normalization: \( \mathbf{x}(\mathbf{F}) \circ \mathbf{y}(\mathbf{F}) = 1 \) is a probability vector. The super-script \( (\cdot)^\top \) denotes transpose. We denote \( e_l \) as the \( l \)th unit coordinate vector and \( \mathbf{I} \) as the identity matrix. For any vector \( \mathbf{\tilde{\gamma}} = (\tilde{\gamma}_1, \ldots, \tilde{\gamma}_L)^\top \in \mathbb{R}^L \) let \( e^{\mathbf{\tilde{\gamma}}} = (e^{\tilde{\gamma}_1}, \ldots, e^{\tilde{\gamma}_L})^\top \). 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:

\[
y_l = h_l x_l + \sum_{j \neq l} h_{lj} x_j + z_l, \tag{1}
\]

where \( y_l \in \mathbb{C}^{1 \times 1} \) is the received signal of the \( l \)th user, \( h_{lj} \in \mathbb{C}^{1 \times 1} \) is the channel coefficient between the transmitter of the \( j \)th user and the receiver of the \( l \)th 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., \( \mathbb{E}[|x_l|^2] = p \), 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 \( G = [G_{lj}]_{L \times L} \) represents the channel gain, where \( G_{lj} = h_{lj}^2 \) is the channel gain from the \( j \)th transmitter to the \( l \)th receiver, and \( \mathbf{n} = (n_1, \ldots, n_L)^\top \geq 0 \), where \( n_l \) is the noise power at the \( l \)th 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:

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

and

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

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 \( l \)th receiver (as a function of powers) is given by:

\[
g_l(p) := \frac{P_l}{\sum_{j \neq l} F_{lj} p_j + v_l}, \quad l = 1, \ldots, L. \tag{4}
\]

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

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 \( g_l(\mathbf{p}) \) is defined by (4). Then \( \rho(\text{diag}(g_l(\mathbf{p})))< 1 \). Hence, for \( \gamma = g_l(\mathbf{p}) \),

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

Vice versa, if \( \gamma \) is in the set

\[
\Gamma := \{ \gamma \geq 0, \quad \rho(\text{diag}(\gamma(\mathbf{F}))) < 1 \},
\]

then \( \mathbf{p} \) given in (5) is nonnegative. Furthermore, \( \gamma(P(\gamma)) = \gamma \). That is, \( \gamma : \mathbb{R}^L_+ \to \Gamma \), and \( P : \Gamma \to \mathbb{R}^L_+ \) are inverse mappings. Let \( \gamma \) denote the vector \( (p_1/v_1, \ldots, p_L/v_L)^\top \). Then, we have \( \gamma \leq \gamma \) 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 additional affine constraints can be imposed on the transmit power of all users as interference temperature constraints. 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 interference temperature affine 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 \( l \)th user can be computed by the Shannon capacity formula [22]:

\[
\log \left( 1 + \frac{g_l(\mathbf{p})}{\Gamma} \right) \text{ nats/symbol,} \tag{7}
\]

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

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 \( l \)th link to reflect priority (a larger weight reflects a higher priority). Denote the set of feasible powers:

\[
\Omega = \{ \mathbf{p} \mid 0 \leq \mathbf{p} \leq \bar{\mathbf{p}} \}. \tag{8}
\]
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

$$D(\{F\}) = \{\tilde{c} \in \mathbb{R}^L, \log \rho(d_{1}^{\oplus} \rho(\frac{e^{c}}{\rho} F + (1/\rho)\omega e_{1}^{\top})) \leq 0 \forall \rho\}$$

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

$$f^1(\tilde{c}) = \left(\frac{e^{\tilde{c}_{1}}}{1 + e^{\tilde{c}_{1}}}, \ldots, \frac{e^{\tilde{c}_{L}}}{1 + e^{\tilde{c}_{L}}}\right)^{\top}.$$  

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

$$f^1(\beta) \leq 0 \quad \forall \beta \in \{f(\beta) = f(\tilde{c}^*)\} \text{ and } \tilde{c}^* \in D(\{F\}).$$

Now, since $D(\{F\})$ 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:

$$D(\{F\}, K) = \{\tilde{c} \in D(\{F\}), \tilde{c} \geq -K1\}$$

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

Lemma 2: Let $\mathbf{P} = \rho(e^{-K1}F - F)^{\top}$ Then $D(\{F\}, K) \subseteq \{\tilde{c} | \log \rho(\mathbf{P} \leq \gamma \leq \log \rho(\mathbf{P})\}.$

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

**A. Outer Approximation Algorithm**

We leverage nonnegative matrix theory and the outer approximation technique in [24], [25] to compute an extreme point of $D(\{F\}, K)$ in (17) that yields $\tilde{c}^*$ 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 $D(\{F\})$ 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 $D(\{F\}, K)$ by a bigger polyhedral convex sets as follows.

- Choose a finite number of points $\zeta_{1}, \ldots, \zeta_{M}$ on the boundary of $D(\{F\}, K)$.

- Let $H_{i}(\xi), \ldots, H_{N}(\xi) \xi \in \mathbb{R}^{L}$ be the $N$ supporting hyperplanes of $D(\{F\})$. Note that we can have more than one supporting hyperplane at $\zeta_{i}$, and at most $L$ supporting hyperplanes. So each $\xi \in D(\{F\}, K)$ satisfies the inequality $H_{j}(\xi) \leq 0 \text{ for } j = 1, \ldots, N.$
Define the polytope
\[ D(\zeta_1, \ldots, \zeta_M, K) = \{ \xi \in \mathbb{R}^L, -K1 \leq \xi \leq \log \gamma \}, \]
\[ H_j(\xi) \leq 0, j = 1, \ldots, N \}, \] (18)
and output this polytope that contains \( D(\{F\}, K) \).

Now, using the polytope \( D(\zeta_1, \ldots, \zeta_M, K) \), we have
\[ \max_{\gamma \in D(\zeta_1, \ldots, \zeta_M, K)} f(\gamma) \geq \max_{\gamma \in \partial D(\{F\}, K)} f(\gamma). \] (19)

Furthermore, since \( f(\gamma) \) 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 multietremeval 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 \( H_j, j = 1, \ldots, N \) for \( N = L \in D(\zeta_1, \ldots, \zeta_M, K) \). To do that, we give a characterization of supporting hyperplanes of \( D(\{F\}) \) at a boundary point \( \xi \in \partial D(\{F\}) \) based on the Friedland–Karlin inequalities in [26]. An extension of these inequalities is found in Theorem 6 in Appendix Section J.

**Theorem 3:** Let \( \tilde{p} = (\tilde{p}_1, \ldots, \tilde{p}_L)^\top > 0 \) be given. Consider the convex set (14). Let \( \zeta \) be a boundary point of \( \partial D(\{F\}) \). Then \( \zeta = \log \gamma(p) \), where \( 0 \leq p = (p_1, \ldots, p_L)^\top \leq \tilde{p} \). The set \( B = \{ l \in \mathbb{L}, p = \tilde{p}_l \} \) is nonempty. For each matrix \( B_l = (F + (1/\tilde{p}_l)ve_l^\top) \), let \( \eta = (\eta_1, \ldots, \eta_L)^\top \in \mathbb{R}^L \) satisfy the condition \( p(e^{\eta_l}B_l) = 1 \). Define \( \delta_l(\zeta) \) as
\[ \delta_l(\zeta) = \sum_{j=1}^L x_j(B_l)\eta_l((\zeta_j - \eta_j)). \] (20)

Then \( \delta_l(\zeta) \leq 0 \), for \( l \in B, \) are the supporting hyperplanes of \( D(\{F\}) \) at \( \zeta \).

We now show how to choose the boundary points \( \zeta_1, \ldots, \zeta_M \in \partial D(\{F\}) \) and to compute the supporting hyperplanes of \( D(\{F\}) \) at each \( \zeta_i \). Let \( \bar{p} = P(e^{K1}) = (\bar{p}_1, \ldots, \bar{p}_L)^\top \) be defined as in Lemma 2. Choose \( M_i \geq 2 \) equidistant points in each interval \( [p_i, \bar{p}_i] \), and define
\[ p_j, i = \frac{p_j + (M_i - j)\bar{p}_i}{M_i} \text{ for } j = 1, \ldots, M_i, i = 1, \ldots, L. \] (21)

Let
\[ \mathcal{P} = \{ p_{j_1}, \ldots, j_L = (p_{j_1}, \ldots, p_{j_L})^\top, \text{min}(p_j - p_{j-1}, \ldots, \bar{p}_L - p_{j_L}) = 0 \}. \]
That is, \( p_{j_1}, \ldots, j_L \in \mathcal{P} \) if and only if \( p_{j_1}, \ldots, j_L \in \mathcal{P} \). Then \( \{\zeta_1, \ldots, \zeta_M\} = \log \gamma(\mathcal{P}) \).

The supporting hyperplanes of \( D(\{F\}) \) at each \( \zeta_i \) are given by Theorem 3. Thus, we have established the \( L \) unique hyperplanes \( H_l(\xi) \leq 0 \) for \( l = 1, \ldots, L \) in \( D(\zeta_1, \ldots, \zeta_M, K) \):
\[ H_l(\xi) = \sum_j x_j(F + (1/\tilde{p}_l)ve_l^\top) \circ y_j(F + (1/\tilde{p}_l)ve_l^\top))(\xi_j) \]
\[ + \log \rho(F + (1/\tilde{p}_l)ve_l^\top)) \leq 0, l = 1, \ldots, L. \] (22)

Based on \( D(\zeta_1, \ldots, \zeta_L, K) \) in (22), we propose the following algorithm that computes the optimal solution of (13).

**Algorithm 1 (Sum Rate Outer Approximation Algorithm):**

- **Input.** \( D(\zeta_1, \ldots, \zeta_L, K) \) with \( H_l(\xi), \forall l \) given in (22).
- **Output.** Optimal solution \( p^* \) of (9).

1) Compute the vertices of the enclosing linear polyhedron \( D(\zeta), \) described by the set of constraints:
\[ \sum_j x_j(F + (1/\tilde{p}_l)ve_l^\top) \circ y_j(F + (1/\tilde{p}_l)ve_l^\top))(\xi_j) \]
\[ + \log \rho(F + (1/\tilde{p}_l)ve_l^\top)) \leq 0, \]
and \( \gamma_l \geq -K \) for all \( l \). Let \( V^{(k)} \) be the set of vertices of \( D(\zeta), \) Set \( k = 1 \) and go to Step 2.

2) **Iteration k:** Solve the problem:
\[ \text{maximize } \sum_i u_i \log(1 + e^\gamma_i) \]
subject to \( \gamma_i \in D^{(k-1)} \) by selecting \( \max \{ \sum_i u_i \log(1 + e^\gamma_i) : u \in V^{(k-1)} \} \), and let \( \gamma_k \) be the optimizer to (24).

3) Compute
\[ p_k = (I - \text{diag}(\exp(\gamma_k))F)^{-1} \text{diag}(\exp(\gamma_k))v. \] (25)

4) If \( p_k \leq \tilde{p}, \) stop: \( \gamma_k \) is the solution to (13) and \( p_k \) is the solution to (9). Otherwise, let \( j^* = \{ l : \log \rho(\text{diag}(\exp(\gamma_k)))F + (1/\tilde{p}_l)ve_l^\top) \}
\[ = \max_{1 \leq l \leq L} \log \rho(\text{diag}(\exp(\gamma_k)))F + (1/\tilde{p}_l)ve_l^\top) \}
and choose any \( j^* \in J^k \).

5) Compute the left eigenvector \( y_{j^*} \) and right (Perron) eigenvector \( x_{j^*} \) of \( \text{diag}(\exp(\gamma_k))F + (1/\tilde{p}_l)ve_l^\top) \).
Set
\[ G_{j^*}^k(\gamma) = \log \rho(\text{diag}(\exp(\gamma_k)))F + (1/\tilde{p}_l)ve_l^\top) = \text{max}_{\text{diag}(\exp(\gamma_k))F + (1/\tilde{p}_l)ve_l^\top)} \]
\[ \text{subject to } \gamma_i \in D(\zeta) \text{ for } i \neq j^* \}
and \( \gamma_k \) is used to construct the inequality cut given by (26) that separates \( \gamma_k \) (infeasible with respect to (13)) from the feasible constraint set of (13).

6) Set \( D^{(k)} = D^{(k-1)} \cap \{ \gamma : G_{j^*}^k(\gamma) \leq 0 \}. \)
7) Let \( k \leftarrow k + 1. \) Go to Step 2.

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

**Theorem 4:** Every limit point of the sequence \( p_k = (I - \text{diag}(\exp(\gamma_k))F)^{-1} \text{diag}(\exp(\gamma_k))v \) solves (9).

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

**Remark 2:** At Step 5, the gradient of \( \log \rho(\text{diag}(\exp(\gamma_k)))F + (1/\tilde{p}_l)ve_l^\top) \) given by \( \exp(\gamma) \circ x_{j^*} \circ y_{j^*} \)
\[ \rho(\text{diag}(\exp(\gamma_k)))F + (1/\tilde{p}_l)ve_l^\top) \]
\( \text{at } \gamma_k \) is used to construct the inequality cut given by (26) that separates \( \gamma_k \) (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 \( \max_j p_j - \tilde{p}_j \leq \varepsilon \) or stop if \( \log \rho(\text{diag}(\exp(\gamma_k)))F + (1/\tilde{p}_l)ve_l^\top) \leq \varepsilon \), where \( \varepsilon \) 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 \( x_{j^*} \)
\( \text{for any } j^* \in J^k \) converges to the limit point of the sequence \( p^k \) in Algorithm 1.
Equality is achieved in (30) if and only if (9). Further, it can be shown that the rate vector generated by restricted box constraints can be easily obtained. Thus, a projection on to the box constraint yields the same solution as maximizing the minimum inner approximation given by (31) optimally solves (13). Furthermore, the optimal solution of (35) is not less than the inner 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 and Theorem 6 holds).

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 $D(\{F\})$.

Lemma 3: For any nonnegative matrix $F$ and positive vector $\gamma$, 
$$\rho(\text{diag}(\gamma)F) \leq \max_l \gamma_l \rho(F).$$

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

Applying Lemma 3 to each constraint in (13), a smaller convex subset of $D(\{F\})$ (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
$$\begin{align*}
\text{maximize} & \quad \sum_l w_l \log(1 + e^{\gamma_l}) \\
\text{subject to} & \quad \tilde{\gamma} \leq \min_l - \log(\rho(F + (1/\tilde{p}_l)v e_i^T)). \quad (31)
\end{align*}$$

has a solution that is an extreme point of $D(\{F\})$ and is given by $\tilde{\gamma} = - \log\rho(F + (1/\tilde{p}_l)v e_i^T)$, where $i = \min_l - \log(\rho(F + (1/\tilde{p}_l)v e_i^T)$. Then, $x(F + (1/\tilde{p}_l)v e_i^T)$ (unique up to a constant) solves (31).

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 $D(\{F\})$ that approximates $D(\{F\})$ 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 and Theorem 6 holds).

B. Relaxation

We now turn to a relaxation technique that can solve special cases of (9). We define the matrix $\bar{F}$ as
$$\bar{F} = F + \text{diag}(\tilde{\gamma})^{-1}.$$

Lemma 5: Let $0 \leq p \leq \tilde{p}$. Then, 
$$p \geq \text{diag}(\gamma(p))\bar{F}p.$$ 

and
$$\rho(\text{diag}(\gamma(p))\bar{F}) \leq 1.$$ 

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

maximize $f(\tilde{\gamma})$
subject to 
$$\log(\rho(\text{diag}(e^{\tilde{\gamma}})F) \leq 0,$$
$$\tilde{\gamma} \leq \log(\tilde{\gamma}).$$

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), $\tilde{\gamma}$, satisfies $\log(\rho(\text{diag}(e^{\tilde{\gamma}})F) = 0$.

Lemma 7: If $p^* = \tilde{p}$ or $p^*$ is such that $p^*_l = 0$ for some $l$ and $p^*_j = \tilde{p}_j$ for $j \neq l$, then
$$\rho(\text{diag}(\gamma(p^*)\bar{F})) = 1.$$

Remark 7: Since (35) is a relaxed problem as compared to (9), both Lemmas 6 and 7 imply that if the optimal solution of (35), $\gamma$, satisfies $P(\gamma) \leq \tilde{p}$, then $P(\gamma)$ 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 $w = (\frac{1}{2}, \frac{1}{2})^T$, $p^*$ is one of the following three vectors: $(\tilde{P}_1, 0)^T$, $(0, \tilde{P}_2)^T$ or $(\tilde{P}_1, \tilde{P}_2)^T$. Furthermore, $p^* = (\tilde{P}_1, 0)^T$ only if $p_1 p_2 F_1 F_2 - 0.5v_1 v_2 - v_1 p_2 \geq 0$, and $p^* = (0, \tilde{P}_2)^T$ only if $p_1 p_2 F_1 F_2 - 0.5v_1 v_2 - v_2 p_1 \gamma_2 \geq 0$.

Remark 8: From the above special case, the relaxed problem (35) is tight when $p^*$ is one of the on-off vectors: $(\tilde{P}_1, 0)^T$, $(0, \tilde{P}_2)^T$ 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

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

Let the maximum power constraint vector be $\tilde{p} = [1.8 \ 100.5]^T$ mW and the noise power of each user be 0.1 mW. The weight vector is given by $w = x(F + (1/|\tilde{P}_i|)v_2|^T) y(F + (1/|\tilde{P}_i|)v_1|^T)$, where $i = \arg \max_0 \rho(F + (1/|\tilde{P}_i|)v_1|^T)$. 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 $p^* = x(F + (1/|\tilde{P}_i|)v_2|^T) = [1.8000 \ 1.442]^T$. 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 $\tilde{v}_1 = (39.4757 \ -100.0)$ and $p^1 = [1.909 \times 10^{10} \ 0]^T$. At Step 4, evaluating each constraint at $\tilde{v}_1 = (39.4757 \ -100.0)$, we have $\log(\text{diag}(\tilde{v}_1^T)(F + (1/|\tilde{P}_i|)v_1|^T)) = 36.8$ and $\log(\text{diag}(\tilde{v}_1^T)(F + (1/|\tilde{P}_i|)v_2|^T)) = -33.4$. We thus choose $j^1 = 1$, and a new constraint is obtained at Step 5 as $\tilde{v}_1 \leq 2.5757$.

Therefore, we have $D^1 = D^0 \cap \{\tilde{v} : \tilde{v}_1 \leq 2.5757\}$ 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 iterations, we arrive at the power vector $p^{29} = [1.8000 \ 1.442]^T$, and $\max_0 p^{29} - \tilde{p}_i = 5.5485 \times 10^{-9}$, 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 fourteenth iteration, respectively. We observe that, by the fourteenth iteration, $D^{(14)}$ provides a relatively good approximation to the feasible region $D((F), K)$. Figure 4 (a) and (b) illustrates the convergence of the rate vectors generated by $\log(1 + \gamma_i (p_k))$ and $\log(1 + \gamma_i (p_k^T))$ 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.

Table I

<table>
<thead>
<tr>
<th>Problem size</th>
<th>Maximal number of generated vertices</th>
<th>Number of iterations</th>
<th>CPU time (minutes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>15</td>
<td>12</td>
<td>0.062</td>
</tr>
<tr>
<td>4</td>
<td>130</td>
<td>760</td>
<td>4.1</td>
</tr>
<tr>
<td>6</td>
<td>14022</td>
<td>1238</td>
<td>83</td>
</tr>
<tr>
<td>8</td>
<td>283681</td>
<td>1968</td>
<td>468</td>
</tr>
</tbody>
</table>

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.

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 multuser multiple carrier model, where a common spectrum is divided into $K$ frequency tones denoted by $(K)$. For illustration 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. Thus, transmissions can be modeled independently on each tone. The achievable rate at tone $k$ can be modeled as [22]:

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

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

$$\sum_{k=1}^{K} p_{t,k} \leq \tilde{p}_t.$$ (39)

3This is also known as discrete multitone modulation in the DSL context.
Fig. 3. Illustration of the convergence of Algorithm 1 in the $\gamma$ 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 $p^h$ (b) the rate vector generated by $\hat{p}^h$. 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 $p = [100.8\ 100.5]^T$ (b) a rate region with $p = [300.8\ 300.5]^T$. Only the rate vectors obtained up to the seventh iteration are indicated.
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 $K = 1$.

For brevity of notations, we define $\mathbf{p} \in \mathbb{R}^L_+$ 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 $\mathbf{p} = (p_{1,1}, p_{1,2}, p_{2,1}, p_{2,2})^T$. Similarly, we define $\gamma \in \mathbb{R}^L_+$ as the SIR allocation, and the matrix $\mathbf{F}$ having entries:

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

and

$$\mathbf{v} = \left( \begin{array}{c} n_{1,1} \frac{G_{11,1}}{G_{11,1} + \gamma_{1,1}}, \\ n_{1,2} \frac{G_{11,2}}{G_{11,2} + \gamma_{1,2}}, \\ \vdots \\ n_{1,K} \frac{G_{11,K}}{G_{11,K} + \gamma_{1,K}}, \\ n_{2,1} \frac{G_{22,1}}{G_{22,1} + \gamma_{2,1}}, \\ \vdots \\ n_{L,K} \frac{G_{L,L,K}}{G_{L,L,K} + \gamma_{L,K}} \end{array} \right)^T$$

(40)

(41)

Note that asynchronous transmission can result in the $l$th 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 $l$th 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:

$$\max \sum_{i=1}^L w_l \sum_{k=1}^K \log(1 + \gamma_{i,k}(\mathbf{p})),$$

where $\gamma_{i,k} = G_{i,l,k} p_{l,k} / \sum_{j=1}^L G_{i,j,k} p_{j,k} + \gamma_{i,k}$. In order to be consistent with our previous results, we shall consider the following equivalent problem:

$$\max \sum_{i=1}^L w_l \sum_{k=1}^K \hat{w}_{i,k} \log(1 + \gamma_{i,k}(\mathbf{p})),$$

where $\hat{w}_{i,k} = (w_l / K)$ for all $l \in \{1, \ldots, L\}, k \in \{1, \ldots, K\}$, and $\mathbf{\hat{w}} \in \mathbb{R}^L_+$. A probability vector.

As in Section III, it is instrumental to consider an alternative formulation of (43).

**Theorem 5:** The optimization problem (43) is equivalent to the following optimization problem:

$$\max \sum_{i=1}^L \sum_{k=1}^K \hat{w}_{i,k} \log(1 + \gamma_{i,k}^\ast)$$

subject to $\rho(\text{diag}(\gamma^\ast)(\mathbf{F} + (1/p) \sum_{k=1}^K \mathbf{v}_{(l)1 \times (l+k)}) \leq 1 \forall l$, $\forall k$.

(44)

Now, $\gamma^\ast$ is an optimal solution of (44) if and only if $P(\gamma^\ast)$ is an optimal solution of the problem (43). In particular, any optimal solution $\gamma^\ast$ satisfies

$$\rho \left( \text{diag}(\gamma^\ast) \left( \mathbf{F} + (1/p) \sum_{k=1}^K \mathbf{v}_{(l)1 \times (l+k)} \right) \right) = 1$$

(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 multiplex 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 power allocations in neighboring tones interfere with one another. This results in intercarrier interference (ICI) and tone coupling in (38); as mentioned earlier, ICI effects are captured by $\mathbf{F}$ in (40), where the optimal solution to (43) is then characterized by the spectral radius of an appropriately constructed interference 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 change 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

$$\max_{\mathbf{F} \in \mathcal{F}} \min_{\mathbf{v} \in \mathbb{R}^{K \times L}_+} \rho(\text{diag}(\gamma)(\mathbf{F} + (1/p) \mathbf{v} \mathbf{v}^T)).$$

(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., an interference temperature constraint. 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**

This research has been supported in part by ARO MURI Award W911NF-08-1-0233.
APPENDIX

A. Proof of Theorem 2
Since \( w > 0 \), the objective function in (13) is a strictly convex function in \( \tilde{\gamma} \in \mathbb{R}^L \). 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 Lemma 2
From the definition of \( K \), we have \( \rho(e^K(F + (1/\tilde{p})ve_\tilde{p}^T)) < 1 \) for \( l = 1, \ldots, L \). Thus, \( -K \mathbf{1} \in D((F)) \). Let \( \gamma = e^K \mathbf{1} \). Assume that \( \tilde{\gamma} \in D((F), K) \). Then \( \gamma \geq -K \mathbf{1} \). Hence, \( \gamma = e^\tilde{\gamma} \geq \gamma \). Since \( \rho(dia_g(\gamma)(F)) < 1 \), we have \( p = P(\gamma) \geq P(\gamma) \). Hence, \( \gamma \) is nonempty, we have \( P(\gamma) \leq \tilde{p} \). Continuing in this manner, we deduce the lemma.

C. Proof of Theorem 3
Let \( p = P(e^C) \). Since the set \( B \) is nonempty, we have \( \rho(e^{CB}) = 1 \) if and only if \( p = \tilde{p} \). Hence, \( \zeta \) lies exactly at the intersection of the hypersurfaces \( \log \rho(e^{CB}) = 0, l \in B \). Theorem 7 in Appendix Section J implies that the supporting hyperplanes of \( D((F)) \) at \( \zeta \) are \( H_l(\xi) \leq 0 \) for \( l \in B \).

D. Proof of Theorem 4
The proof of convergence of \( \gamma^k \) to the optimal solution of (13) by cutting plane methodology can be found in [24]. Since the map \( p^k = (I - d(\exp(\gamma^k))F)^{-1}d(\exp(\gamma^k))v \) is bijective, the limit point of \( p^k \) solves (9). This proves the theorem.

E. Proof of Lemma 4
Lemma 4 is easily proved as the objective function in (31) is monotonicly increasing and by using 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 \( x(F + (1/\tilde{p})ve_\tilde{p}^T) \), unique up to a constant (cf. Theorem 2 in [17]).

F. Proof of Lemma 5
Since \( 0 \leq p_l \leq \bar{p}_l \), we have \( \tilde{p}_l \leq p_l \leq v_l \) for all \( l \). By the definition of \( \gamma(p) \) in (4), we have \( p_l = \gamma(p) \left( \begin{array}{c} \gamma_1 + \sum_{j \neq l} F_{lj} \end{array} \right) \), which together with the definition of \( \tilde{F} \), 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 \( p > 0 \), i.e., all the coordinates of \( p \) are positive, then \( \gamma(p) > 0 \). Hence, dia_g(\gamma(p))F is a positive matrix. Then, by the Wielandt’s max-min characterization of the spectral radius, we have
\[
\rho(dia_g(\gamma(p))F) \leq \max_{l=1}^{L} \frac{dia_g(\gamma(p))F_{ll}}{p_l} \leq 1.
\]
Observe next that if \( p_l = 0 \), then \( \gamma(p)_l = 0 \). So if some of \( p_l = 0 \), then \( \rho(dia_g(\gamma(p))F) \) is the spectral radius of the maximal positive submatrix of \( dia_g(\gamma(p))F \). Apply to this positive submatix Wielandt’s characterization to deduce (34).

G. 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 \( f(\gamma) \) and \( \log \rho(dia_g(e^{\gamma})F) \) increase with \( \gamma \), this proves the lemma.

H. Proof of Lemma 7
The definition of \( \tilde{F} \) implies (36) for \( p^* = \tilde{p} \). Assume now that \( p^*_l = 0 \) for some \( l \). Then \( \gamma_l(\tilde{p})^* = 0 \) for some \( l \). Then, the \( l \)th row of \( dia_g(\gamma_l(p^*))F \) is zero. Let \( F_{(l)} \) be the submatrix of \( F \) obtained by deleting the \( l \)th row and column. Let \( \gamma_{(l)} \) be the vector obtained from \( \gamma \) by deleting the \( l \)th coordinate. Hence, the characteristic polynomial of \( dia_g(\gamma_l(p^*))F \), \( det(xI - dia_g(\gamma_l(p^*))F) \), is equal to \( x(1 - dia_g(\gamma_{(l)}F_{(l)})) \). Therefore, \( \rho(dia_g(\gamma_l(p^*))F) = \rho(dia_g(\gamma_{(l)}F_{(l)})) \). Continuing in this manner, we deduce the lemma.

I. 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 \( \rho(dia_g(\gamma_l(p^*))F) = 1 \) (from (36) in Lemma 7) is equivalent to the following linear program:
maximize \( \tilde{p}_1 \tilde{p}_2 F_1 F_2 - 0.5 \tilde{v}_1 \tilde{v}_2 (\gamma_1 + \gamma_2) - v_1 \tilde{p}_2 \gamma_1 \)
subject to \( \gamma \leq \tilde{\gamma} \).

As a linear program with two variables over a simple box constraint, (47) has an optimal solution that is an extreme point. If \( F \) is a non-zero matrix, the solution is \( (\gamma_1, 0)^T \) only if \( \tilde{p}_1 \tilde{p}_2 F_1 F_2 - 0.5 \tilde{v}_1 \tilde{v}_2 - v_1 \tilde{p}_2 \gamma_1 \geq 0 \), and \( (0, \gamma_2)^T \) only if \( \tilde{p}_1 \tilde{p}_2 F_1 F_2 - 0.5 \tilde{v}_1 \tilde{v}_2 - v_2 \gamma_2 \geq 0 \). Using (5), the corresponding optimal power vector is deduced as given by Corollary 1.

J. 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 \( A \in \mathbb{R}^{L \times L} \) be an irreducible matrix. Assume that \( x(A) = (x_1(A), \ldots, x_L(A))^T \), \( y(A) = (y_1(A), \ldots, y_L(A))^T \) are the Perron right and left eigenvectors of \( A \) respectively, normalized such that \( x(A)^T y(A) \) is a probability vector. Suppose that \( \gamma \) is a nonnegative vector. Then
\[
\rho(A) \prod_{l=1}^{L} \gamma_l (x(A) \cdot y(A))_l \leq \rho(dia_g(\gamma(A))A).
\]

If \( \gamma \) is a positive vector then equality holds if and only if all \( \gamma_l \) are equal. Furthermore, for any positive vector \( z = (z_1, \ldots, z_L)^T \), the following inequality holds:
\[
\rho(A) \leq \prod_{l=1}^{L} \left( \frac{\text{trace}(A z_l)}{z_l} \right) (x(A) \cdot y(A))_l.
\]

If \( A \) is an irreducible matrix with positive diagonal elements, then equality holds in (49) if and only if \( z = tx(A) \) for some positive \( t \). The following result gives an interpretation of the inequality (48) in terms of the supporting hyperplane of the convex
Theorem 7: Let $B \in \mathbb{R}^{L\times L}$ be an irreducible matrix. Let $\eta = (\eta_1, \ldots, \eta_L)^T \in \mathbb{R}^L$ satisfy the condition $\rho(e^{\eta}B) = 1$. Denote $A = e^{\eta}B$ and assume that $x(A) = (x_1(A), \ldots, x_L(A))^T$, $y(A) = (y_1(A), \ldots, y_L(A))^T > 0$ are the Perron right and left eigenvectors of $A$ respectively, normalized such that $x(A) \circ y(A)$ is a probability vector. Let

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

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

REFERENCES


