Maximizing Sum Rates in Gaussian Interference-limited Channels*

Shmuel Friedland† and Chee Wei Tan‡

June 17, 2008

Abstract

We study the problem of maximizing sum rates in Gaussian interference-limited channels. We show that this maximum problem can be restated as a maximization problem of a convex function on a closed convex set. We suggest three algorithms to find the exact and approximate values of the optimal rates.
Keywords and phrases: Sum rates, maximization of convex functions, spectral radii of irreducible matrices, linear programming, wireless networks.
2000 Mathematics Subject Classification. 15A42, 15A48, 47N10, 49K35, 65K05, 94A40.

1 Introduction

Power control is used in cellular and ad-hoc networks to provide a high signal-to-noise (SNR) ratio for a reliable connection. A higher SNR also allows a wireless system that uses link adaptation to transmit at a higher data rate, thus leading to a greater spectral efficiency. Transmission rate adaptation by power control is an active research area in communication networks that can be used for both interference management and utility maximization [17].

The motivation of this work comes from maximizing sum rate, (data throughput), in wireless communications. Due to the broadcast nature of radio transmission, data rates in a wireless network are affected by interference. This is particularly true in Code Division Multiple Access (CDMA) systems, where users transmit at the same time over the same frequency bands and their spreading codes are not perfectly orthogonal. Transmit power control is often used to control signal interference to maximize the total transmission rates of all users.

Obtaining the optimal transmit power control scheme, that maximizes total throughput, requires solving a nonconvex problem. Hence, it is hard to obtain optimal solutions [18, 5, 7, 13, 22, 4]. In this paper we show that we can convert the problem of maximizing sum rate to a problem of maximizing a convex function

*Dedicated to the memory of Samuel Karlin

†Department of Mathematics, Statistics and Computer Science, University of Illinois at Chicago, Chicago, Illinois 60607-7045, USA, and Berlin Mathematical School, Berlin, Germany, e-mail:friedlan@uic.edu

‡Electrical Engineering Department, Princeton University, NJ 08544, USA, e-mail:cheetan@princeton.edu

on a closed unbounded convex set. This approach leads to some new algorithms for finding the exact and approximate solutions to the maximum problem.

We now state briefly the maximal problem. (See §2 for all definitions, notations and motivations.) Let $F = [f_{ij}]_{i=j=1}^{L}, \mathbf{v} = (v_1, \ldots, v_L)^\top$ be an $L \times L$ matrix, with zero diagonal and positive off diagonal elements, and a positive vector, respectively. Define the following transformation $\boldsymbol{\gamma}(\mathbf{p}) = (\gamma_1(\mathbf{p}), \ldots, \gamma_L(\mathbf{p}))^\top$ from the set of nonnegative vectors $\mathbf{p} \in \mathbb{R}_+^L$ to itself:

$$\gamma_l(\mathbf{p}) = \frac{p_l}{\sum_{m=1}^{L} f_{lm} p_m + v_l}, \text{ for } l = 1, \ldots, L, \text{ where } \mathbf{p} = (p_1, \ldots, p_l) \geq \mathbf{0}.$$

Let $\bar{\mathbf{p}} = (\bar{p}_1, \ldots, \bar{p}_L)^\top$ be a given positive vector. For a given probability vector $\mathbf{w} = (w_1, \ldots, w_L)^\top$, let

$$\Phi_{\mathbf{w}}(\boldsymbol{\gamma}) := \sum_{i=1}^{L} w_i \log(1 + \gamma_i), \text{ where } \mathbf{w}, \boldsymbol{\gamma} \in \mathbb{R}_+^L. \tag{1.1}$$

Then the sum rate maximization problem in interference-limited channels is $\max_{\mathbf{0} \leq \mathbf{p} \leq \bar{\mathbf{p}}} \Phi(\boldsymbol{\gamma}(\mathbf{p}))$. The exact solution to this problem is known to be NP-complete [14].

For any vector $\tilde{\boldsymbol{\gamma}} = (\tilde{\gamma}_1, \ldots, \tilde{\gamma}_L)^\top \in \mathbb{R}^L$ let $e^{\tilde{\boldsymbol{\gamma}}} = (e^{\tilde{\gamma}_1}, \ldots, e^{\tilde{\gamma}_L})^\top$. The purpose of this paper to show that the sum rate maximization problem equivalent to the maximization problem of the convex function $\Phi_{\mathbf{w}}(e^{\tilde{\boldsymbol{\gamma}}})$, where $\tilde{\boldsymbol{\gamma}}$ varies on a closed unbounded convex domain $D(\{F\} \subset \mathbb{R}^L)$. (See (4.6).) Note that this formulations shows that we can not expect, in general, that a maximal solution $\mathbf{p}^\star$ is unique. We show how to approximate this domain by a convex polytope $D(\boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_M, K)$, formed mainly by the supporting hyperplanes of $D(\{F\})$. This approximation is obtained by using the Friedland-Karlin inequality stated in [8]. Hence, it is easier to find the maximum $\Phi(e^{\tilde{\boldsymbol{\gamma}}})$ on $D(\boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_M, K)$, than on $D(\{F\})$. Finally we approximate the function $\Phi(e^{\tilde{\boldsymbol{\gamma}}})$ by the linear function $\mathbf{w}^\top \tilde{\boldsymbol{\gamma}} = \sum_{l=1}^{L} w_l \tilde{\gamma}_l$. Then the maximum of $\mathbf{w}^\top \tilde{\boldsymbol{\gamma}}$ on $D(\boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_M, K)$ is a linear programming problem, which can be solved in polynomial time using the ellipsoid algorithm [10].

It is also possible to generalize our work to find application to optimal spectrum management in DSL broadband access systems with interference-limited performance across the tones when users share a cable binder [14, 22].

We survey briefly the contents of the paper. In §2 we state definitions, notations and a short motivation. We give a characterization of the image of the multidimensional box $[\mathbf{0}, \bar{\mathbf{p}}] \subset \mathbb{R}_+^L$ by the map $\boldsymbol{\gamma}$, in terms of certain inequalities of the spectral radii of corresponding nonnegative matrices. In §3 we study the sum rate maximization problem. We give necessary and sufficient conditions for an extremal point $\mathbf{p} \in [\mathbf{0}, \bar{\mathbf{p}}]$ of $\Phi_{\mathbf{w}}(\boldsymbol{\gamma}(\mathbf{p}))$ to be a local maximum. In §4 we study several relaxation versions of the sum rate maximization problem. These relaxation version are simpler than the original maximal problem. When replacing the maximal function $\Phi_{\mathbf{w}}$ by the linear function $\mathbf{w}^\top \tilde{\boldsymbol{\gamma}}$, we have a closed-form solution under certain conditions on $\mathbf{w}$. This solution is obtained by using the Friedland-Karlin inequality stated in [8]. In §5 we give three algorithms to solve the sum rate maximization problem and its two other approximation versions described above. In §6, viewed as an appendix, we restate some useful results for [8] and give several applications and extensions, which are needed in this paper.

2 Notations and preliminary results

As usual, let $\mathbb{R}^{m\times n} \supset \mathbb{R}_+^{m\times n}$ denote the set of $m \times n$ matrices and its subset of nonnegative matrices. For $A, B \in \mathbb{R}^{m\times n}$, we denote $A \le B$ if $B - A \in \mathbb{R}_+^{m\times n}$. We denote $A \lneq B, A < B$ if $B - A$ is a nonzero nonnegative and positive matrix, respectively. We denote the entries of a matrix $A \in \mathbb{R}^{m\times n}$ by the small letters, i.e $A = [a_{ij}]_{i,j=1}^{m,n}$. Identify $\mathbb{R}^m = \mathbb{R}^{m\times 1}, \mathbb{R}_+^m = \mathbb{R}_+^{m\times 1}$.

A column vector is denoted by the **bold** letter $\mathbf{x} = (x_1, \dots, x_m)^\top \in \mathbb{R}^m$. Then $e^{\mathbf{x}} := (e^{x_1}, \dots, e^{x_m})^\top$. For $\mathbf{x} > \mathbf{0}$, we let $\mathbf{x}^{-1} := (\frac{1}{x_1}, \dots, \frac{1}{x_m})^\top$ and $\log \mathbf{x} = (\log x_1, \dots, \log x_L)^\top$. For any $\mathbf{x} = (x_1, \dots, x_L)^\top \in \mathbb{R}^L$, we let $e^{\mathbf{x}} = (e^{x_1}, \dots, e^{x_L})^\top$. Let $\mathbf{x}\circ\mathbf{y}$ denote the Schur product of the vectors $\mathbf{x}$ and $\mathbf{y}$, i.e., $\mathbf{x}\circ\mathbf{y} = [x_1y_1, \dots, x_Ly_L]^T$. Let $\mathbf{1} = (1, \dots, 1)^\top \in \mathbb{R}^L$. For $\underline{\mathbf{p}} \le \bar{\mathbf{p}} \in \mathbb{R}^L$, denote by $[\underline{\mathbf{p}}, \bar{\mathbf{p}}]$ the set of all $\mathbf{x} \in \mathbb{R}^L$ satisfying $\underline{\mathbf{p}} \le \mathbf{x} \le \bar{\mathbf{p}}$. For a vector $\mathbf{y} = (y_1, \dots, y_L)^\top$, denote by $\mathrm{diag}(\mathbf{y})$ the diagonal matrix $\mathrm{diag}(y_1, \dots, y_L)$. We also let $(B\mathbf{y})_l$ denote the lth element of $B\mathbf{y}$. The Perron-Frobenius eigenvalue of a nonnegative matrix F is denoted as $\rho(F)$, and the Perron (right) and left eigenvector of F associated with $\rho(F)$ are denoted by $\mathbf{x}(F)$ and $\mathbf{y}(F)$ (or simply $\mathbf{x}$ and $\mathbf{y}$ when the context is clear) respectively. Assume that F is a nonnegative irreducible matrix. Then $\mathbf{x}(F), \mathbf{y}(F) > \mathbf{0}$. We will assume the normalization: $\mathbf{x}(F) \circ \mathbf{y}(F)$ is a probability vector. The super-script $(\cdot)^\top$ denotes transpose. For a positive integer n, denote by $\langle n\rangle$ the set $\{1, \dots, n\}$. 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.

Consider a wireless network, e.g., cellular network, with L logical transmitter/receiver pairs. Transmit powers are denoted as $p_1, \dots, p_L$. In the cellular uplink case, all logical receivers may reside in the same physical receiver, i.e., the base station. Let $G = [g_{ij}]_{i,j=1}^L > 0_{L\times L}$ representing the channel gain, where g_{ij} is the channel gain from the jth transmitter to the ith receiver, and $\mathbf{n} = (n_1, \dots, n_L)^\top > \mathbf{0}$, where n_l is the noise power for the lth receiver be given. The Signal-to-Interference Ratio (SIR) for the lth receiver is denoted by γ_l.

For $\mathbf{p} = (p_1, \dots, p_L)^\top \ge \mathbf{0}$, we define the following transformation: $\mathbf{p} \mapsto \boldsymbol{\gamma}(\mathbf{p})$, where

$$\gamma_l(\mathbf{p}) := \frac{g_{ll}p_l}{\sum_{j\ne l} g_{lj}p_j + n_l},\ l = 1, \dots, L,\ \boldsymbol{\gamma}(\mathbf{p}) = (\gamma_1(\mathbf{p}), \dots, \gamma_L(\mathbf{p}))^\top. \tag{2.1}$$

Define

$$F = [f_{ij}]_{i,j=1}^L,\ \text{where } f_{ij} = \begin{cases} 0, & \text{if } i = j \\ \frac{g_{ij}}{g_{ii}}, & \text{if } i \ne j \end{cases} \tag{2.2}$$

and

$$\mathbf{g} = (g_{11}, \dots, g_{LL})^\top,\ \mathbf{n} = (n_1, \dots, n_L)^\top, \mathbf{v} = \left(\frac{n_1}{g_{11}}, \frac{n_2}{g_{22}}, \dots, \frac{n_L}{g_{LL}}\right)^\top. \tag{2.3}$$

Then

$$\boldsymbol{\gamma}(\mathbf{p}) = \mathbf{p} \circ (F\mathbf{p} + \mathbf{v})^{-1}. \tag{2.4}$$

Claim 2.1 *Let $\mathbf{p} \ge \mathbf{0}$ be a nonnegative vector. Assume that $\boldsymbol{\gamma}(\mathbf{p})$ is defined by (2.1). Then $\rho(\mathrm{diag}(\boldsymbol{\gamma}(\mathbf{p}))F) < 1$, where F is defined by (2.2). Hence, for $\boldsymbol{\gamma} = \boldsymbol{\gamma}(\mathbf{p})$,*

$$\mathbf{p} = P(\boldsymbol{\gamma}) := (I - \mathrm{diag}(\boldsymbol{\gamma})F)^{-1}\,\mathrm{diag}(\boldsymbol{\gamma})\mathbf{v}. \tag{2.5}$$

Vice versa, if γ *is in the set*

$$\mathbf{\Gamma} := \{\gamma \geq \mathbf{0},\ \rho(\mathrm{diag}(\gamma)F) < 1\}, \tag{2.6}$$

then the vector $\mathbf{p}$ *defined by (2.5) is nonnegative. Furthermore,* $\gamma(P(\mathbf{p})) = \gamma$*. That is,* $\gamma : \mathbb{R}_+^L \to \mathbf{\Gamma}$*, and* $P : \mathbf{\Gamma} \to \mathbb{R}_+^L$ *are inverse mappings.*

Proof. Observe that (2.1) is equivalent to the equality

$$\mathbf{p} = \mathrm{diag}(\gamma)F\mathbf{p} + \mathrm{diag}(\gamma)\mathbf{v}. \tag{2.7}$$

Assume first that $\mathbf{p}$ is a positive vector, i.e., $\mathbf{p} > \mathbf{0}$. Hence, $\gamma(\mathbf{p}) > \mathbf{0}$. Since all off-diagonal entries of $\mathbf{F}$ are positive it follows that the matrix $\mathrm{diag}(\gamma)F$ is irreducible. As $\mathbf{v} > \mathbf{0}$, we deduce that $\max_{i\in[1,n]} \frac{(\mathrm{diag}(\gamma)F\mathbf{p})_i}{p_i} < 1$. The min max characterization of Wielandt of $\rho(\mathrm{diag}(\gamma)F)$, [21] and [9, (38), pp.64], implies $\rho(\mathrm{diag}(\gamma)F) < 1$. Hence, $\gamma(\mathbf{p}) \in \mathbf{\Gamma}$. Assume now that $\mathbf{p} \geq \mathbf{0}$. Note that $p_i > 0 \iff \gamma_i(\mathbf{p}) > 0$. So $\mathbf{p} = \mathbf{0} \iff \gamma(\mathbf{p}) = 0$. Clearly, $\rho(\gamma(\mathbf{0})F) = \rho(0_{L\times L}) = 0 < 1$. Assume now that $\mathbf{p} \gneq \mathbf{0}$. Let $\mathcal{A} = \{i :\ p_i > 0\}$. Denote $\gamma(\mathbf{p})(\mathcal{A})$ the vector composed of positive entries of $\gamma(\mathbf{p})$. Let $F(\mathcal{A})$ be the principal submatrix of $\mathbf{F}$ with rows and columns in $\mathcal{A}$. It is straightforward to see that $\rho(\mathrm{diag}(\gamma(\mathbf{p}))F) = \rho(\mathrm{diag}(\gamma(p)(\mathcal{A})F(\mathcal{A}))$. The arguments above imply that

$$\rho(\mathrm{diag}(\gamma(\mathbf{p}))F) = \rho(\mathrm{diag}(\gamma(\mathbf{p})(\mathcal{A})\mathbf{F}(\mathcal{A})) < 1.$$

Assume now that $\gamma \in \mathbf{\Gamma}$. Then

$$(I - \mathrm{diag}(\gamma)F)^{-1} = \sum_{k=0}^{\infty} (\mathrm{diag}(\gamma)F)^k \geq 0_{L\times L}. \tag{2.8}$$

Hence, $P(\gamma) \geq \mathbf{0}$. The definition of $P(\gamma)$ implies that $\gamma(P(\gamma)) = \gamma$. □

Claim 2.2 *The set* $\mathbf{\Gamma} \subset \mathbb{R}_+^L$ *is monotonic with respect to the order* $\geq$*. That is if* $\gamma \in \mathbf{\Gamma}$ *and* $\gamma \geq \beta \geq \mathbf{0}$ *then* $\beta \in \mathbf{\Gamma}$*. Furthermore, the function* $P(\gamma)$ *is monotone on* $\mathbf{\Gamma}$*.*

$$P(\gamma) \geq P(\beta) \text{ if } \gamma \in \mathbf{\Gamma} \text{ and } \gamma \geq \beta \geq \mathbf{0}. \tag{2.9}$$

Equality holds if and only if $\gamma = \beta$*.*

Proof. Clearly, if $\gamma \geq \beta \geq \mathbf{0}$ then $\mathrm{diag}(\gamma)F \geq \mathrm{diag}(\beta)F$ which implies $\rho(\mathrm{diag}(\gamma)F) \geq \rho(\mathrm{diag}(\beta)F)$. Hence, $\mathbf{\Gamma}$ is monotonic. Use the Neumann expansion (2.8) to deduce the monotonicity of P. The equality case is straightforward. □

Note that $\gamma(\mathbf{p})$ is not monotonic in $\mathbf{p}$. Indeed, if one increases only the ith coordinate of $\mathbf{p}$, then one increases the ith coordinate of $\gamma(\mathbf{p})$ and decreases all other coordinates of $\gamma(\mathbf{p})$.

As usual, let $\mathbf{e}_i = (\delta_{i1}, \ldots, \delta_{iL})^\top,\ i = 1, \ldots, L$ be the standard basis in $\mathbb{R}^L$. In what follows, we need the following result.

Theorem 2.3 *Let $l \in [1, L]$ be an integer and $a > 0$. Denote $[0, a]_l \times \mathbb{R}_+^{L-1}$ the set of all $\mathbf{p} = (p_1, \dots, p_L)^\top \in \mathbb{R}_+^L$ satisfying $p_l \le a$. Then the image of the set $[0, a]_l \times \mathbb{R}_+^{L-1}$ by the map $\boldsymbol{\gamma}$ (2.1), is given by*

$$\rho(\mathrm{diag}(\boldsymbol{\gamma})(F + \frac{1}{a}\mathbf{v}\mathbf{e}_l^\top)) \le 1, \ \mathbf{0} \le \boldsymbol{\gamma}. \tag{2.10}$$

Furthermore, $\mathbf{p} = (p_1, \dots, p_L) \in \mathbb{R}_+^L$ satisfies the condition $p_l = a$ if and only if $\boldsymbol{\gamma} = \boldsymbol{\gamma}(\mathbf{p})$ satisfies

$$\rho(\mathrm{diag}(\boldsymbol{\gamma})(F + \frac{1}{a}\mathbf{v}\mathbf{e}_l^\top)) = 1. \tag{2.11}$$

Proof. Suppose that $\boldsymbol{\gamma}$ satisfies (2.10). We claim that $\boldsymbol{\gamma} \in \boldsymbol{\Gamma}$. Suppose first that $\boldsymbol{\gamma} > \mathbf{0}$. Then $\mathrm{diag}(\boldsymbol{\gamma})(F + t_1\mathbf{v}\mathbf{e}_l^\top) \lneq \mathrm{diag}(\boldsymbol{\gamma})(F + t_2\mathbf{v}\mathbf{e}_l^\top)$ for any $t_1 < t_2$. [9, Lemma 2, §2, Ch. XIII] yields

$$\begin{aligned} \rho(\mathrm{diag}(\boldsymbol{\gamma})F) < \rho(\mathrm{diag}(\boldsymbol{\gamma})(F + t_1\mathbf{v}\mathbf{e}_l^\top)) < \rho(\mathrm{diag}(\boldsymbol{\gamma})(F + t_2\mathbf{v}\mathbf{e}_l^\top)) < \\ \rho(\mathrm{diag}(\boldsymbol{\gamma})(F + \frac{1}{a}\mathbf{v}\mathbf{e}_l^\top)) \le 1 \text{ for } 0 < t_1 < t_2 < \frac{1}{a}. \end{aligned} \tag{2.12}$$

Thus $\boldsymbol{\gamma} \in \boldsymbol{\Gamma}$. Combine the above argument with the arguments of the proof of Claim 2.1 to deduce that $\boldsymbol{\gamma} \in \boldsymbol{\Gamma}$ for $\boldsymbol{\gamma} \ge \mathbf{0}$.

We now show that $P(\boldsymbol{\gamma})_l \le a$. The continuity of P implies that it is enough to consider the case $\boldsymbol{\gamma} > \mathbf{0}$. Combine the Perron-Frobenius theorem with (2.12) to deduce

$$0 < \det(I - \mathrm{diag}(\boldsymbol{\gamma})(F + t\mathbf{v}\mathbf{e}_l^\top)) \text{ for } t \in [0, a^{-1}). \tag{2.13}$$

We now expand the right-hand side of the above inequality. Let $B = \mathbf{x}\mathbf{y}^\top \in \mathbb{R}^{L \times L}$ be a rank one matrix. Then B has $L-1$ zero eigenvalues and one eigenvalue equal to $\mathbf{y}^\top\mathbf{x}$. Hence, $I - \mathbf{x}\mathbf{y}^\top$ has $L-1$ eigenvalues equal to 1 and one eigenvalue is $(1 - \mathbf{y}^\top\mathbf{x})$. Therefore, $\det(I - \mathbf{x}\mathbf{y}^\top) = 1 - \mathbf{y}^\top\mathbf{x}$. Since $\boldsymbol{\gamma} \in \boldsymbol{\Gamma}$ we get that $(I - \mathrm{diag}(\boldsymbol{\gamma})F)$ is invertible. Thus, for any $t \in \mathbb{R}$

$$\begin{aligned} \det(I - \mathrm{diag}(\boldsymbol{\gamma})(F + t\mathbf{v}\mathbf{e}_l^\top)) = \\ \det(I - \mathrm{diag}(\boldsymbol{\gamma})F)\det(I - t((I - \mathrm{diag}(\boldsymbol{\gamma})F)^{-1}\mathrm{diag}(\boldsymbol{\gamma})\mathbf{v})\mathbf{e}_l^\top) \\ \det(I - \mathrm{diag}(\boldsymbol{\gamma})F)(1 - t\mathbf{e}_l^\top(I - \mathrm{diag}(\boldsymbol{\gamma})F)^{-1}\mathrm{diag}(\boldsymbol{\gamma})\mathbf{v}). \end{aligned} \tag{2.14}$$

Combine (2.13) with the above identity to deduce

$$1 > t\mathbf{e}_l^\top(I - \mathrm{diag}(\boldsymbol{\gamma})F)^{-1}\mathrm{diag}(\boldsymbol{\gamma})\mathbf{v} = tP(\boldsymbol{\gamma})_l \text{ for } t \in [0, a^{-1}). \tag{2.15}$$

Letting $t \nearrow a^{-1}$, we deduce that $P(\boldsymbol{\gamma})_l \le a$. Hence, the set of $\boldsymbol{\gamma}$ defined by (2.10) is a subset of $\boldsymbol{\gamma}([0, a]_l \times \mathbb{R}_+^{L-1})$.

Let $\mathbf{p} \in [0, a]_l \times \mathbb{R}_+^{L-1}$ and denote $\boldsymbol{\gamma} = \boldsymbol{\gamma}(\mathbf{p})$. We show that $\boldsymbol{\gamma}$ satisfies (2.10). Claim 2.1 implies that $\rho(\mathrm{diag}(\boldsymbol{\gamma})F) < 1$. Since $\mathbf{p} = P(\boldsymbol{\gamma})$ and $p_l \le a$ we deduce (2.15). Use (2.14) to deduce (2.13). As $\rho(\mathrm{diag}(\boldsymbol{\gamma})F) < 1$, the inequality (2.13) implies that $\rho(\mathrm{diag}(\boldsymbol{\gamma})F + t\mathbf{v}^\top\mathbf{e}_l) < 1$ for $t \in (0, a^{-1})$. Hence, (2.10) holds.

It is left to show the condition (2.11) holds if and only if $P(\boldsymbol{\gamma})_l = a$. Assume that $\mathbf{p} = (p_1, \dots, p_L)^\top \in \mathbb{R}_+^L$, $p_l = a$ and let $\boldsymbol{\gamma} = \boldsymbol{\gamma}(\mathbf{p})$. We claim that equality holds in (2.10). Assume to the contrary that $\rho(\mathrm{diag}(\boldsymbol{\gamma})(F + \frac{1}{a}\mathbf{v}\mathbf{e}_l^\top)) < 1$. Then, there exists $\boldsymbol{\beta} > \boldsymbol{\gamma}$ such that $\rho(\mathrm{diag}(\boldsymbol{\beta})(F + \frac{1}{a}\mathbf{v}\mathbf{e}_l^\top)) < 1$. Since P is monotonic $P(\boldsymbol{\beta})_l > p_l = a$.

On the other hand, since $\boldsymbol{\beta}$ satisfies (2.10), we deduce that $P(\boldsymbol{\beta})_l \le a$. This contradiction yields (2.11). Similarly, if $\boldsymbol{\gamma} \ge \mathbf{0}$ and (2.11) then $P(\boldsymbol{\gamma})_l = a$. □

Corollary 2.4 *Let $\bar{\mathbf{p}} = (\bar{p}_1, \ldots, \bar{p}_L)^\top > \mathbf{0}$ be a given positive vector. Then $\boldsymbol{\gamma}([\mathbf{0}, \bar{\mathbf{p}}])$, the image of the set $[\mathbf{0}, \bar{\mathbf{p}}]$ by the map $\boldsymbol{\gamma}$ (2.1), is given by*

$$\rho\left(\operatorname{diag}(\boldsymbol{\gamma})\left(F + \frac{1}{\bar{\mathbf{p}}_l}\mathbf{v}\mathbf{e}_l^\top\right)\right) \le 1, \text{ for } l = 1, \ldots, L, \text{ and } \gamma \in \mathbb{R}_+^L. \tag{2.16}$$

In particular, any $\boldsymbol{\gamma} \in \mathbb{R}_+^L$ satisfying the conditions (2.16) satisfies the inequalities

$$\boldsymbol{\gamma} \le \bar{\boldsymbol{\gamma}} = (\bar{\gamma}_1, \ldots, \bar{\gamma}_L)^\top, \text{ where } \bar{\gamma}_l = \frac{\bar{p}_l}{v_l}, \ i = 1, \ldots, L. \tag{2.17}$$

Proof. Theorem 2.3 yields that $\boldsymbol{\gamma}([\mathbf{0}, \bar{\mathbf{p}}])$ is given by (2.16). (2.4) yields

$$\gamma_l(\mathbf{p}) = \frac{p_l}{((F\mathbf{p})_l + v_l)} \le \frac{p_l}{v_l} \le \frac{\bar{p}_l}{v_l} \text{ for } \mathbf{p} \in [\mathbf{0}, \bar{\mathbf{p}}].$$

Note that equality holds for $\mathbf{p} = \bar{p}_l \mathbf{e}_l$. □

3 The sum rate maximization problem

Let $\mathbf{w} = (w_1, \ldots, w_L)^\top \ge 0$ be a given probability vector. Assuming a singe-user decoder at each user, i.e., treating interference as noise, we can use the Shannon capacity formula for the attainable data rate. Then the problem of maximizing sum rate in wireless communication can be stated as the following maximum problem

$$\max_{\mathbf{0} \le \mathbf{p} \le \bar{\mathbf{p}}} \sum_{l=1}^{L} w_l \log(1 + \gamma_l(\mathbf{p})). \tag{3.1}$$

Lemma 3.1 *Let $\mathbf{w}$ be a probability vector, and assume that $\mathbf{p}^\star = (p_1^\star, \ldots, p_L^\star)^\top$ is a maximal solution to (3.1). Then $p_i^\star = \bar{p}_i$ for some i. Furthermore if $w_j = 0$ then $p_j^\star = 0$.*

Proof. Assume to the contrary that $\mathbf{p}^\star < \bar{\mathbf{p}}$. Let $\boldsymbol{\gamma}^\star = \boldsymbol{\gamma}(\mathbf{p}^\star)$. Since P is continuous on $\boldsymbol{\Gamma}$, there exists $\boldsymbol{\gamma} \in \boldsymbol{\Gamma}$ such that $\boldsymbol{\gamma} > \boldsymbol{\gamma}^\star$ such that $P(\boldsymbol{\gamma}) < \bar{\mathbf{p}}$. Clearly, $\Phi_{\mathbf{w}}(\boldsymbol{\gamma}(\mathbf{p}^\star)) < \Phi_{\mathbf{w}}(\boldsymbol{\gamma})$. As $\boldsymbol{\gamma} = \boldsymbol{\gamma}(P(\boldsymbol{\gamma}))$, we deduce that $\mathbf{p}^\star$ is not a maximal solution to (3.1), contrary to our assumptions.

Suppose that $w_j = 0$. For $\mathbf{p} = (p_1, \ldots, p_L)^\top$, let $\mathbf{p}_j$ be obtained from $\mathbf{p}$ by replacing the *jth* coordinate in $\mathbf{p}$ by 0. Assume that $p_j > 0$. Then $\gamma_i(\mathbf{p}) < \gamma_i(\mathbf{p}_j)$ for $i \ne j$. Since $w_j = 0$, it follows that $\Phi_{\mathbf{w}}(\boldsymbol{\gamma}(\mathbf{p})) < \Phi_{\mathbf{w}}(\boldsymbol{\gamma}(\mathbf{p}_j))$. □

Combine the above lemma with Theorem 2.3 and Corollary 2.4 to deduce an alternative formulation of (3.1).

Theorem 3.2 *The maximum problem (3.1) is equivalent to the maximum problem.*

$$\begin{array}{ll} \textit{maximize} & \sum_l w_l \log(1+\gamma_l) \\ \textit{subject to} & \rho(\mathrm{diag}(\boldsymbol{\gamma})(F + (1/\bar{p}_l)\mathbf{v}\mathbf{e}_l^\top)) \le 1 \quad \forall\, l \in \langle L\rangle, \\ \textit{variables:} & \gamma_l, \quad \forall\, l. \end{array} \tag{3.2}$$

$\boldsymbol{\gamma}^\star$ is a maximal solution of the above problem if and only if $P(\boldsymbol{\gamma}^\star)$ is a maximal solution $p^\star$ of the problem (3.1). In particular, any maximal solution $\boldsymbol{\gamma}^\star$ satisfies the equality (2.16) for some integer $l \in [1, L]$.

We now give the following simple necessary conditions for a maximal solution $\mathbf{p}^\star$ of (3.1). We first need the following result, which is obtained by straightforward differentiation.

Lemma 3.3 *Denote by*

$$\nabla \Phi_{\mathbf{w}}(\boldsymbol{\gamma}) = \left(\frac{w_1}{1+\gamma_1}, \ldots, \frac{w_L}{1+\gamma_L}\right)^\top = \mathbf{w} \circ (\mathbf{1} + \boldsymbol{\gamma})^{-1}$$

the gradient of $\Phi_{\mathbf{w}}$. Let $\boldsymbol{\gamma}(\mathbf{p})$ be defined as in (2.1). Then $\mathrm{H}(\mathbf{p}) = [\frac{\partial \gamma_i}{\partial p_j}]_{i=j=1}^L$, the Hessian matrix of $\boldsymbol{\gamma}(\mathbf{p})$, is given by

$$H(\mathbf{p}) = \mathrm{diag}((F\mathbf{p} + \mathbf{v})^{-1})(-\,\mathrm{diag}(\boldsymbol{\gamma}(\mathbf{p}))F + I).$$

In particular,

$$\nabla_{\mathbf{p}} \Phi_{\mathbf{w}}(\boldsymbol{\gamma}(\mathbf{p})) = \mathrm{H}(\mathbf{p})^\top \nabla \Phi_{\mathbf{w}}(\boldsymbol{\gamma}(\mathbf{p})).$$

Corollary 3.4 *Let $\mathbf{p}^\star = (p_1^\star, \ldots, p_L^\star)^\top$ be a maximal solution to the problem (3.1). Divide the set $\langle L\rangle = \{1, \ldots, L\}$ to the following three disjoint sets $\mathrm{S}_{\max}, \mathrm{S}_{\mathrm{in}}, \mathrm{S}_0$:*

$$\mathrm{S}_{\max} = \{i \in \langle L\rangle,\ p_i^\star = \bar{p}_i\},\ \mathrm{S}_{\mathrm{in}} = \{i \in \langle L\rangle,\ p_i^\star \in (0, \bar{p}_i)\},\ \mathrm{S}_0 = \{i \in \langle L\rangle,\ p_i^\star = 0\}.$$

Then the following conditions hold.

$$\begin{array}{l} (\mathrm{H}(\mathbf{p}^\star)^\top \nabla \Phi_{\mathbf{w}}(\boldsymbol{\gamma}(\mathbf{p}^\star)))_i \ge 0 \ \textit{for}\ i \in \mathrm{S}_{\max}, \\ (\mathrm{H}(\mathbf{p}^\star)^\top \nabla \Phi_{\mathbf{w}}(\boldsymbol{\gamma}(\mathbf{p}^\star)))_i = 0 \ \textit{for}\ i \in \mathrm{S}_{\mathrm{in}}, \\ (\mathrm{H}(\mathbf{p}^\star)^\top \nabla \Phi_{\mathbf{w}}(\boldsymbol{\gamma}(\mathbf{p}^\star)))_i \le 0 \ \textit{for}\ i \in \mathrm{S}_0 \end{array} \tag{3.3}$$

Proof. Assume that $p_i^\star = \bar{p}_i$. Then $\frac{\partial}{\partial p_i}\Phi_{\mathbf{w}}(\boldsymbol{\gamma}(\mathbf{p}))(\mathbf{p}^\star) \ge 0$. Assume that $0 < p_i^\star < \bar{p}_i$. Then $\frac{\partial}{\partial p_i}\Phi_{\mathbf{w}}(\boldsymbol{\gamma}(\mathbf{p}))(\mathbf{p}^\star) = 0$. Assume that $p_i^\star = 0$. Then $\frac{\partial}{\partial p_i}\Phi_{\mathbf{w}}(\boldsymbol{\gamma}(\mathbf{p}))(\mathbf{p}^\star) \le 0$. □

We now show that the maximum problem (3.2) can be restated as the maximum problem of convex function on a closed unbounded domain. For $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_L)^\top > 0$ let $\tilde{\boldsymbol{\gamma}} = \log \boldsymbol{\gamma}$, i.e. $\boldsymbol{\gamma} = e^{\tilde{\boldsymbol{\gamma}}}$. Recall that for a nonnegative irreducible matrix $B \in \mathbb{R}_+^{L\times L}$ $\log \rho(e^{\mathbf{x}} B)$ is a convex function [11]. Furthermore, $\log(1 + e^t)$ is a strict convex function in $t \in \mathbb{R}$. Hence, the maximum problem (3.2) is equivalent to the problem

$$\begin{array}{rl} \text{maximize} & \sum_l w_l \log(1 + e^{\tilde{\gamma}_l}) \\ \text{subject to} & \log \rho(\mathrm{diag}(e^{\tilde{\boldsymbol{\gamma}}})(F + (1/\bar{p}_l)\mathbf{v}\mathbf{e}_l^\top)) \le 0 \quad \forall\, l \in \langle L\rangle, \\ \text{variables:} & \tilde{\boldsymbol{\gamma}} = (\tilde{\gamma}_1, \ldots, \tilde{\gamma}_n)^\top \in \mathbb{R}^L. \end{array} \tag{3.4}$$

The unboundedness of the convex set in (3.4) is due to the identity $0 = e^{-\infty}$. In view of Lemma 3.1, it is enough to consider the maximal problem (3.1) in the case where $\mathbf{w} > \mathbf{0}$.

Theorem 3.5 *Let $\mathbf{w} > \mathbf{0}$ be a probability vector. Consider the maximum problem (3.1). Then any point $\mathbf{0} \le \mathbf{p}^\star \le \bar{\mathbf{p}}$ satisfying the conditions (3.3) is a local maximum.*

Proof. Since $\mathbf{w} > \mathbf{0}$, $\Phi_{\mathbf{w}}(e^{\tilde{\boldsymbol{\gamma}}})$ is a strict convex function in $\tilde{\boldsymbol{\gamma}} \in \mathbb{R}^L$. Hence, the maximum of (3.4) is achieved exactly on the extreme points of the closed unbounded set specified in (3.4). (It may happen that some coordinate of the extreme point are $-\infty$.) Translating this observation to the maximal problem (3.1), we deduce the theorem. □

We now give simple lower and upper bounds on the value of (3.1).

Lemma 3.6 *Consider the maximal problem (3.1). Let $B_l = (F + (1/\bar{p}_l)\mathbf{v}\mathbf{e}_l^\top))$ for $l = 1, \ldots, L$. Denote $R = \max_{l\in\langle L\rangle} \rho(B_l)$. Let $\bar{\boldsymbol{\gamma}}$ be defined by (2.17). Then*

$$\Phi_{\mathbf{w}}((1/R)\mathbf{1}) \le \max_{\mathbf{p}\in[\mathbf{0},\bar{\mathbf{p}}]} \Phi_{\mathbf{w}}(\boldsymbol{\gamma}(\mathbf{p}) \le \Phi_{\mathbf{w}}(\bar{\boldsymbol{\gamma}}).$$

Proof. By Corollary 2.4, $\boldsymbol{\gamma}(\mathbf{p}) \le \bar{\boldsymbol{\gamma}}$ for $\mathbf{p} \in [\mathbf{0}, \bar{\mathbf{p}}]$. Hence, the upper bounds holds. Clearly, for $\boldsymbol{\gamma} = (1/R)\mathbf{1}$, we have that $\rho(\mathrm{diag}(\boldsymbol{\gamma})B_l) \le 1$ for $l \in \langle L\rangle$. Then, from Theorem 3.2, $\Phi_{\mathbf{w}}((1/R)\mathbf{1})$ yields the lower bound. Equality is achieved in the lower bound when $\mathbf{p}^\star = t\mathbf{x}(B_i)$, where $i = \arg\max_{l\in\langle L\rangle} \rho(B_l)$, for some $t > 0$. □

4 Relaxations of the maximal problem

In this section, we study several relaxed versions of (3.2), which will be used later to construct algorithms to solve (3.1).

From the definition of $\boldsymbol{\gamma}(p)$, we deduce that

$$p_l = \gamma_l(\mathbf{p}) \left(\frac{n_l}{g_{ll}} + \sum_{j\ne l} \frac{g_{lj}}{g_{ll}} p_j \right). \tag{4.1}$$

Define

$$\tilde{F} = [\tilde{f}_{lj}]_{l,j=1}^L, \ \tilde{f}_{ll} = \frac{n_l}{g_{ll}\bar{p}_l}, \ \tilde{f}_{lj} = \frac{g_{lj}}{g_{ll}} \text{ for } j \ne l, \ l = 1, \ldots, L. \tag{4.2}$$

Lemma 4.1 *Let $\mathbf{0} \le \mathbf{p} \le \bar{\mathbf{p}}$. Assume that $\boldsymbol{\gamma}(\mathbf{p})$ and $\tilde{F}$ are defined by (2.1) and (4.2), respectively. Then*

$$\mathbf{p} \ge \mathrm{diag}(\boldsymbol{\gamma}(\mathbf{p}))\tilde{F}\mathbf{p}, \tag{4.3}$$

and

$$\rho(\mathrm{diag}(\boldsymbol{\gamma}(\mathbf{p}))\tilde{F}) \le 1. \tag{4.4}$$

Proof. The assumption that $0 \le p_l \le \bar{p}_l$ implies that $\frac{n_l}{g_{ll}} \ge \tilde{f}_{ll} p_l$. Then using (4.1), the definition of $\tilde{F}$ and the above observation implies (4.3). The inequality (4.4) is a consequence of Wielandt's characterization of the spectral radius of an irreducible matrix [9]. Indeed, if $\mathbf{p} > \mathbf{0}$, i.e. all the coordinates of $\mathbf{p}$ are positive, then $\boldsymbol{\gamma}(\mathbf{p}) > 0$. Hence, $\mathrm{diag}(\boldsymbol{\gamma}(\mathbf{p})))\tilde{F}$ is a positive matrix. Then Wielandt's characterizations claims

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

Observe next that if $p_l = 0$, then $\boldsymbol{\gamma}(\mathbf{p})_l = 0$. So if some of $p_l = 0$, then $\rho(\mathrm{diag}(\boldsymbol{\gamma}(\mathbf{p}))\tilde{F})$ is the spectral radius of the maximal positive submatrix of $\mathrm{diag}(\boldsymbol{\gamma}(\mathbf{p}))\tilde{F}$. Apply to this positive submatrix Wielandt's characterization to deduce (4.4). □

Lemma 4.2 *The maximum*

$$\max_{\boldsymbol{\gamma}\in\mathbb{R}_+^L,\rho(\mathrm{diag}(\boldsymbol{\gamma})\tilde{F})\le 1} \Phi_{\mathbf{w}}(\boldsymbol{\gamma}) \tag{4.5}$$

is not less than the maximum of the problem (3.1).

Proof. In view of (4.4), we see that the maximum in (4.5) is on a bigger set than the maximum in (3.1). □

By considering a change of variable in the constraint set of (4.5), we define the following set:

$$D(\tilde{F}) = \{\tilde{\gamma} \in \mathbb{R}^L, \quad \log \rho(\mathrm{diag}(e^{\tilde{\gamma}})\tilde{F}) \le 0\}, \tag{4.6}$$

which is convex in $\mathbb{R}^L$.

Corollary 4.3 *We have $\rho(\mathrm{diag}(\boldsymbol{\gamma}')\tilde{F}) = 1$ in (4.5), where $\boldsymbol{\gamma}'$ solves (4.5) optimally.*

Lemma 4.4 *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 $j \ne l$, then*

$$\rho(\mathrm{diag}(\boldsymbol{\gamma}(\mathbf{p}^\star))\tilde{F}) = 1. \tag{4.7}$$

Proof. The definition of $\tilde{F}$ implies (4.7) 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 $\mathrm{diag}(\boldsymbol{\gamma}(\mathbf{p}^\star))\tilde{F}$ is zero. Let F_1 be the submatrix of F obtained by deleting the lth row and column. Let $\boldsymbol{\gamma}_1$ be the vector obtained from $\boldsymbol{\gamma}$ by deleting the lth coordinate. Hence, the characteristic polynomial of $\mathrm{diag}(\boldsymbol{\gamma})F$, $\det(xI - \mathrm{diag}(\boldsymbol{\gamma})F$, is equal to $x \det(xI - \mathrm{diag}(\boldsymbol{\gamma}_1)F_1)$. Therefore, $\rho(\mathrm{diag}(\boldsymbol{\gamma})F) = \rho(\mathrm{diag}(\boldsymbol{\gamma}_1)F_1)$. Continuing in this manner, we deduce the lemma. □

Corollary 4.3 and Lemma 4.4 imply that if the optimizer of (4.5) $\boldsymbol{\gamma}'$ satisfies $P(\boldsymbol{\gamma}') \preceq \bar{\mathbf{p}}$, then $P(\boldsymbol{\gamma}')$ is also the global optimizer of (3.1).

We now state a stronger relaxation problem than (4.5). Consider the transformation $\boldsymbol{\gamma}(\mathbf{p}, \mathbf{n})$ given in (2.1) as a function of $\mathbf{p}$ and the noise $\mathbf{n} > \mathbf{0}$. Clearly,

$\boldsymbol{\gamma}(\mathbf{p}, \mathbf{n}_1) \gneq \boldsymbol{\gamma}(\mathbf{p}, \mathbf{n}_2)$ for any $\mathbf{n}_2 \gneq \mathbf{n}_1 > \mathbf{0}$ and $\mathbf{p} \geq \mathbf{0}$. Hence, it would be useful to consider the limiting case $\delta(\mathbf{p}) := \boldsymbol{\gamma}(\mathbf{p}, \mathbf{0})$:

$$\beta_l(\mathbf{p}) := \frac{g_{ll} p_l}{\sum_{j \neq l} g_{lj} p_j}, \ l = 1, \ldots, L, \ \boldsymbol{\beta}(\mathbf{p}) = (\beta_1(\mathbf{p}), \ldots, \beta_L(\mathbf{p}))^\top, \mathbf{p} \gneq \mathbf{0}. \qquad (4.8)$$

Hence, $\beta(\mathbf{p})$ can be viewed as communication in *noiseless channels* [1]. Another way to look at the noiseless channels is to allow channels with unlimited power $\mathbf{p}$, i.e. $\bar{p}_i = \infty, i = 1, \ldots, L$. This remark is implied by the identity

$$\boldsymbol{\gamma}(t\mathbf{p}, \mathbf{n}) = \boldsymbol{\gamma}(\mathbf{p}, \frac{1}{t}\mathbf{n}) \text{ for any } t > 0.$$

Denote by $\Pi_L = \{\mathbf{p} \in \mathbb{R}_+^L, \ \mathbf{1}^\top \mathbf{p} = 1\}$ the set of probability vectors in $\mathbb{R}_+^L$. Since $\beta(\mathbf{p}) = \beta(t\mathbf{p})$ for any $t > 0$, we deduce

$$\sup_{\mathbf{p} \in \mathbb{R}_+^L \setminus \{\mathbf{0}\}} \Phi_{\mathbf{w}}(\boldsymbol{\beta}(\mathbf{p})) = \max_{\mathbf{p} \in \Pi_L} \Phi_{\mathbf{w}}(\boldsymbol{\beta}(\mathbf{p})). \qquad (4.9)$$

Clearly, the maximum in (4.9) is greater than the maximum in (3.1).

Note that the definition of F given by (2.2) implies that

$$\mathrm{diag}(\boldsymbol{\beta}(\mathbf{p})) F \mathbf{p} = \mathbf{p}, \text{ and } \rho(\mathrm{diag}(\boldsymbol{\beta}(\mathbf{p})) F) = 1 \text{ for any } \mathbf{p} \gneq \mathbf{0}. \qquad (4.10)$$

Hence, the maximum (4.9) is equal to

$$\max_{\boldsymbol{\beta} \in \mathbb{R}_+^L, \rho(\mathrm{diag}(\boldsymbol{\beta}) F) \leq 1} \Phi_{\mathbf{w}}(\boldsymbol{\beta}) = \max_{\boldsymbol{\beta} \in \mathbb{R}_+^L, \rho(\mathrm{diag}(\boldsymbol{\beta}) F) = 1} \Phi_{\mathbf{w}}(\boldsymbol{\beta}). \qquad (4.11)$$

Since $F \lneq \tilde{F}$ and $\tilde{F}$ is positive, it follows that $\rho(\mathrm{diag}(\boldsymbol{\gamma}) F) < \rho(\mathrm{diag}(\boldsymbol{\gamma}) \tilde{F})$ for any $\boldsymbol{\gamma} \gneq \mathbf{0}$ [9]. Hence, the maximum in (4.11) is greater than the maximum in (4.5).

Since it is generally difficult to determine precisely the spectral radius of a given matrix [20], the relaxed problems given by (4.5) and (4.11) enables one to find an upper bound to (3.1) quickly when L is fairly large.

We next modify the function $\Phi_{\mathbf{w}}(e^{\tilde{\boldsymbol{\gamma}}})$ appearing in (5.4). Consider the function $\log(1 + e^x)$. If e^x is relatively small, then $\log(1 + e^x)$ is close to zero. If e^x is assumed to be bigger than 1, then $\log(1 + e^x) \approx \log e^x = x$. Hence, it is reasonable to approximate $\log(1 + e^x)$ by x (Note that $\log(1 + e^x) > x$).

Theorem 4.5 *Let $\mathbf{w} = (w_1, \ldots, w_L)^\top$ be a positive probability vector. Then*

$$\max_{\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_L)^\top > \mathbf{0}, \rho(\mathrm{diag}(\boldsymbol{\gamma}) \tilde{F}) \leq 1} \sum_{l=1}^{L} w_l \log \gamma_l = \sum_{l=1}^{L} w_l \log \gamma_l^\star, \qquad (4.12)$$

where $\boldsymbol{\gamma}^\star = (\gamma_1^\star, \ldots, \gamma_L^\star)^\top > \mathbf{0}$ is the unique vector satisfying the following conditions. Let $A^\star = \mathrm{diag}(\boldsymbol{\gamma}^\star) \tilde{F}$. Then $\rho(A^\star) = 1$ and $\mathbf{x}(A^\star) \circ \mathbf{y}(A^\star) = \mathbf{w}$.

Proof. In the maximum problem (4.12), it is enough to consider $\boldsymbol{\gamma} > \mathbf{0}$ such that $\rho(\mathrm{diag}(\boldsymbol{\gamma}) \tilde{F}) = 1$. Then each of such $\boldsymbol{\gamma}$ is of the form $\mathbf{p} \circ (F\mathbf{p})^{-1}$, for some $\mathbf{0} < \mathbf{z}$ ($= \mathbf{x}(\mathrm{diag}(\boldsymbol{\gamma}) \tilde{F})$ in Theorem 6.3. Use Theorem 6.3 and Corollary 6.6 to deduce the theorem. □

Combining Theorem 6.3 and Corollary 6.9 we deduce the following result.

Theorem 4.6 *Let* $\mathbf{w} = (w_1, \ldots, w_L)^\top$ *be a positive probability vector satisfying the conditions (6.6). Let* $\hat{F}$ *be either equal to* F *or to* $F + (1/\bar{p}_l)\mathbf{v}\mathbf{e}_l^\top$ *for some* $l \in \langle L \rangle$*, respectively. Then*

$$\max_{\boldsymbol{\gamma}=(\gamma_1,\ldots,\gamma_L)^\top > \mathbf{0}, \rho(\mathrm{diag}(\boldsymbol{\gamma})\hat{F}) \le 1} \sum_{l=1}^{L} w_l \log \gamma_l = \sum_{l=1}^{L} w_l \log \gamma_l^\star, \tag{4.13}$$

where $\boldsymbol{\gamma}^\star = (\gamma_1^\star, \ldots, \gamma_L^\star)^\top > \mathbf{0}$ *is a vector satisfying the following conditions. Let* $A^\star = \mathrm{diag}(\boldsymbol{\gamma}^\star)\hat{F}$*. Then* $\rho(A^\star) = 1$ *and* $\mathbf{x}(A^\star) \circ \mathbf{y}(A^\star) = \mathbf{w}$*.*

Note that the choice $\hat{F} = F + (1/\bar{p}_l)\mathbf{v}\mathbf{e}_l^\top$ applies, if we assume that the optimal solution $\mathbf{p}^\star = (p_1^\star, \ldots, p_L^\star)^\top$ of the maximal problem (3.1) satisfies $p_l^\star = \bar{p}_l$ and $p_j^\star < \bar{p}_j$ for $j \neq l$.

The above last two theorems enable us to choose $\mathbf{w}$ for which we know the solution to the maximal problems (4.12) and (4.13). Namely, choose $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2 > \mathbf{0}$ such that $A_1 = \mathrm{diag}(\boldsymbol{\beta}_1)\tilde{F}, A_2 = \mathrm{diag}(\boldsymbol{\beta}_2)\hat{F}$ have spectral radius one. Let $\mathbf{w}_i = \mathbf{x}(A_i) \circ \mathbf{y}(A_i)$ for $i = 1, 2$. Then for $\mathbf{w}_1$, (4.12) has the unique maximal solution $\boldsymbol{\gamma}^\star = \boldsymbol{\beta}_1$. For $\mathbf{w}_2$, (4.13) has a maximal solution $\boldsymbol{\gamma}^\star = \boldsymbol{\beta}_2$. Note that in view of Theorem 6.3, $\mathbf{w}_2$ does not have to satisfy the conditions (6.6).

5 Algorithms for sum rate maximization

In this section, we outline three algorithms for finding and estimating the maximal sum rates. In view of Lemma 3.1, it is enough to consider the maximal problem (3.1) in the case where $\mathbf{w} > \mathbf{0}$. Theorem 3.5 gives rise to the following algorithm, which is the gradient algorithm in the variable $\mathbf{p}$ in the compact polyhedron $[\mathbf{0}, \bar{\mathbf{p}}]$.

Algorithm 5.1

1. *Choose* $\mathbf{p}_0 \in [\mathbf{0}, \bar{\mathbf{p}}]$*:*
   (a) *Either at random;*
   (b) *or* $\mathbf{p}_0 = \bar{\mathbf{p}}$*.*
2. *Given* $\mathbf{p}_k = (p_{1,k}, \ldots, p_{L,k})^\top \in [\mathbf{0}, \bar{\mathbf{p}}]$ *for* $k \ge 0$*, compute* $\mathbf{a} = (a_1, \ldots, a_L)^\top = \nabla_{\mathbf{p}} \Phi_{\mathbf{w}}(\boldsymbol{\gamma}(\mathbf{p}_k))$*. If* $\mathbf{a}$ *satisfies the conditions (3.3) for* $\mathbf{p}^\star = \mathbf{p}_k$*, then* $\mathbf{p}_k$ *is the output. Otherwise let* $\mathbf{b} = (b_1, \ldots, b_L)^\top$ *be defined as follows.*
   (a) $b_i = 0$ *if* $p_{i,k} = 0$ *and* $a_i < 0$*;*
   (b) $b_i = 0$ *if* $p_{i,k} = \bar{p}_i$ *and* $a_i > 0$*;*
   (c) $b_i = a_i$ *if* $0 < p_i < \bar{p}_i$*.*

   Then $\mathbf{p}_{k+1} = \mathbf{p}_k + t_k\mathbf{b}$*, where* $t_k > 0$ *satisfies the conditions* $\mathbf{p}_{k+1} \in [\mathbf{0}, \bar{\mathbf{p}}]$ *and* $\Phi_{\mathbf{w}}(\boldsymbol{\gamma}(\mathbf{p}_k + t_k\mathbf{b}_k))$ *increases on the interval* $[0, t_k]$*.*

The problem with the gradient method, and its variations as a conjugate gradient method is that it is hard to choose the optimal value of t_k in each step, e.g. [2]. We now use the reformulation of the maximal problem given by (3.4). Since $\mathbf{w} > \mathbf{0}$, the function $\Phi_{\mathbf{w}}(e^{\tilde{\gamma}})$ is strictly convex. Thus, the maximum is achieved only on the boundary of the convex set

$$D(\{F\}) = \{\tilde{\boldsymbol{\gamma}} \in \mathbb{R}^L, \quad \log \rho(\operatorname{diag}(e^{\tilde{\boldsymbol{\gamma}}})(F + (1/\bar{p}_l)\mathbf{v}\mathbf{e}_l^\top)) \leq 0, \ \forall\, l\}. \tag{5.1}$$

If one wants to use numerical methods and software for finding the maximum value of convex functions on bounded closed convex sets , e.g., [15], then one needs to consider the maximization problem (3.4) with additional constraints:

$$D(\{F\}, K) = \{\tilde{\boldsymbol{\gamma}} \in D(\{F\}), \quad \tilde{\boldsymbol{\gamma}} \geq -K\mathbf{1}\}. \tag{5.2}$$

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

Lemma 5.2 *Let $\bar{\mathbf{p}} > \mathbf{0}$ be given and let R be defined as in Lemma 3.6. Assume that $K > \log R$. Let $\underline{\mathbf{p}} = P(e^{-K}\mathbf{1}) = (e^K I - F)^{-1}\mathbf{v}$. Then $D(\{F\}, K) \subseteq \log \boldsymbol{\gamma}([\underline{\mathbf{p}}, \bar{\mathbf{p}}])$.*

Proof. From the definition of K, we have that $e^K > R$. Hence, $\rho(e^{-K}B_l) < 1$ for $l = 1, \ldots, L$. Thus $-K\mathbf{1} \in D(\{F\})$. Let $\underline{\boldsymbol{\gamma}} = e^{-K}\mathbf{1}$. Assume that $\tilde{\boldsymbol{\gamma}} \in D(\{F\}, K)$. Then $\tilde{\boldsymbol{\gamma}} \geq -K\mathbf{1}$. Hence, $\boldsymbol{\gamma} = e^{\tilde{\boldsymbol{\gamma}}} \geq \underline{\boldsymbol{\gamma}}$. Since $\rho(\operatorname{diag}(\boldsymbol{\gamma})F) < 1$, Claim 2.2 yields that $\mathbf{p} = P(\boldsymbol{\gamma}) \geq P(\underline{\boldsymbol{\gamma}}) = \underline{\mathbf{p}}$, where P is defined by (2.5). The inequality $P(\boldsymbol{\gamma}) \leq \bar{\mathbf{p}}$ follows from Corollary 2.4. □

Thus, we can apply the numerical methods to find the maximum of the strictly convex function $\Phi_{\mathbf{w}}(e^{\tilde{\gamma}})$ on the closed bounded set $D(\{F\}, K)$, e.g. [15]. In particular, we can use the gradient method. It takes the given boundary point $\tilde{\boldsymbol{\gamma}}_k$ to another boundary point of $\tilde{\boldsymbol{\gamma}}_{k+1} \in D(\{F\}, K)$, in the direction induced by the gradient of $\Phi_{\mathbf{w}}(e^{\tilde{\gamma}})$. However, the complicated boundary of $D(\{F\}, K)$ will make any algorithm expensive.

Furthermore, even though the constraint set in (3.2) can be transformed into a strict convex set, it is in general difficult to determine precisely the spectral radius of a given matrix [20]. To make the problem simpler and to enable fast algorithms, we approximate the convex set $D(\{F\}, K)$ by a bigger polyhedral convex sets as follows. Choose a finite number of points $\boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_M$ on the boundary of $D(\{F\})$, which preferably lie in $D(\{F\}, K)$. Let $\mathrm{H}_1(\boldsymbol{\xi}), \ldots, \mathrm{H}_N(\boldsymbol{\xi}), \boldsymbol{\xi} \in \mathbb{R}^L$ be the N supporting hyperplanes of $D(\{F\}$. (Note that we can have more than one supporting hyperplane at $\boldsymbol{\zeta}_i$, and at most L supporting hyperplanes.) So each $\boldsymbol{\xi} \in D(\{F\}, K)$ satisfies the inequality $H_j(\boldsymbol{\xi}) \leq 0$ for $j = 1, \ldots, N$. Let $\bar{\boldsymbol{\gamma}}$ be defined by (2.17). Define

$$D(\boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_M, K) = \{\boldsymbol{\xi} \in \mathbb{R}^L, \ -K\mathbf{1} \leq \boldsymbol{\xi} \leq \log \bar{\boldsymbol{\gamma}}, \ \mathrm{H}_j(\boldsymbol{\xi}) \leq 0 \text{ for } j = 1, \ldots, N\}. \tag{5.3}$$

Hence, $D(\boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_M, K)$ is a polytope which contains $D(\{F\}, K)$. Thus

$$\max_{\tilde{\boldsymbol{\gamma}} \in D(\boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_M, K)} \Phi_{\mathbf{w}}(e^{\tilde{\boldsymbol{\gamma}}}) \geq \tag{5.4}$$

$$\max_{\tilde{\boldsymbol{\gamma}} \in D(\{F\}, K)} \Phi_{\mathbf{w}}(e^{\tilde{\boldsymbol{\gamma}}}). \tag{5.5}$$

Since $\Phi_{\mathbf{w}}(e^{\tilde{\gamma}})$ is strictly convex, the maximum in (5.4) is achieved only at the extreme points of $D(\boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_M, K)$. The maximal solution can be found using a variant of a simplex algorithm [6]. More precisely, one starts at some extreme point of $\boldsymbol{\xi} \in D(\boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_M, K)$. Replace the strictly convex function $\Phi_{\mathbf{w}}(e^{\tilde{\boldsymbol{\gamma}}})$ by its first order Taylor expansion $\Psi_{\boldsymbol{\xi}}$ at $\boldsymbol{\xi}$. Then we find another extreme point $\boldsymbol{\eta}$ of $D(\boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_M, K)$, such that $\Psi_{\boldsymbol{\xi}}(\boldsymbol{\eta}) > \Psi_{\boldsymbol{\xi}}(\boldsymbol{\xi}) = \Phi_{\mathbf{w}}(e^{\boldsymbol{\xi}})$. Then we replace $\Phi_{\mathbf{w}}(e^{\tilde{\boldsymbol{\gamma}}})$ by its first order Taylor expansion $\Psi_{\boldsymbol{\eta}}$ at $\boldsymbol{\eta}$ and continue the algorithm. Our second proposed algorithm for finding an optimal $\tilde{\boldsymbol{\gamma}}^\star$ that maximizes (5.4) is given as follows.

Algorithm 5.3

1. *Choose an arbitrarily extreme point* $\boldsymbol{\xi}_0 \in D(\boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_M, K)$.
2. *Let* $\Psi_{\boldsymbol{\xi}_k}(\boldsymbol{\xi}) = \Phi_{\mathbf{w}}(e^{\boldsymbol{\xi}_k}) + (\mathbf{w} \circ (\mathbf{1} + e^{\boldsymbol{\xi}_k})^{-1} \circ e^{\boldsymbol{\xi}_k})^\top (\boldsymbol{\xi} - \boldsymbol{\xi}_k)$. *Solve the linear program* $\max_{\boldsymbol{\xi}} \Psi_{\boldsymbol{\xi}_k}(\boldsymbol{\xi})$ *subject to* $\boldsymbol{\xi} \in D(\boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_M, K)$ *using the simplex algorithm in [6] by finding an extreme point* $\boldsymbol{\xi}_{k+1}$ *of* $D(\boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_M, K)$, *such that* $\Psi_{\boldsymbol{\xi}_k}(\boldsymbol{\xi}_{k+1}) > \Psi_{\boldsymbol{\xi}_k}(\boldsymbol{\xi}_k) = \Phi_{\mathbf{w}}(e^{\boldsymbol{\xi}_k})$.
3. *Compute* $\mathbf{p}_k = P(e^{\boldsymbol{\xi}_{k+1}})$. *If* $\mathbf{p}_k \in [0, \bar{\mathbf{p}}]$, *compute* $\mathbf{a} = (a_1, \ldots, a_L)^\top = \nabla_{\mathbf{p}} \Phi_{\mathbf{w}}(\boldsymbol{\gamma}(\mathbf{p}_k))$. *If* $\mathbf{a}$ *satisfies the conditions (3.3) for* $\mathbf{p}^\star = \mathbf{p}_k$, *then* $\mathbf{p}_k$ *is the output. Otherwise, go to Step 2 using* $\Psi_{\boldsymbol{\xi}_{k+1}}(\boldsymbol{\xi})$.

As in the previous section, it would be useful to consider the following related maximal problem:

$$\max_{\tilde{\boldsymbol{\gamma}} \in D(\boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_M, K)} \mathbf{w}^\top \tilde{\boldsymbol{\gamma}}. \tag{5.6}$$

This problem given by (5.6) is a standard linear program, which can be solved in polynomial time by the classical ellipsoid algorithm [10]. Our third proposed algorithm for finding an optimal $\tilde{\boldsymbol{\gamma}}^\star$ that maximizes (5.6) is given as follows.

Algorithm 5.4

1. *Solve the linear program* $\max_{\tilde{\boldsymbol{\gamma}}} \mathbf{w}^\top \tilde{\boldsymbol{\gamma}}$ *subject to* $\tilde{\boldsymbol{\gamma}} \in D(\boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_M, K)$ *using the ellipsoid algorithm in [10].*
2. *Compute* $\mathbf{p} = P(e^{\tilde{\boldsymbol{\gamma}}})$. *If* $\mathbf{p} \in [0, \bar{\mathbf{p}}]$, *then* $\mathbf{p}$ *is the output. Otherwise, project* $\mathbf{p}$ *onto* $[0, \bar{\mathbf{p}}]$.

We note that $\tilde{\boldsymbol{\gamma}} \in D(\boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_M, K)$ in Algorithm 5.4 can be replaced by the set of supporting hyperplane $D(\tilde{F}, K) = \{\tilde{\boldsymbol{\gamma}} \in \rho(\mathrm{diag}(e^{\tilde{\boldsymbol{\gamma}}})\tilde{F}) \le 1, \quad \tilde{\boldsymbol{\gamma}} \ge -K\mathbf{1}\}$ (extending (4.6)) or, if $L \ge 3$ and $\mathbf{w}$ satisfies the conditions (6.6), $D(F, K) = \{\tilde{\boldsymbol{\gamma}} \in \rho(\mathrm{diag}(e^{\tilde{\boldsymbol{\gamma}}})F) \le 1, \quad \tilde{\boldsymbol{\gamma}} \ge -K\mathbf{1}\}$ based on the relaxed maximal problems in Section 4. Then Theorem 4.5 and 4.6 quantify the closed-form solution $\tilde{\boldsymbol{\gamma}}$ computed by Algorithm 5.4.

We conclude this section by showing how to compute the supporting hyperplanes $\mathrm{H}_j, j = 1, \ldots, N$, which define $D(\boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_M, K)$. To do that, we give a characterization of supporting hyperlanes of $D(\{F\})$ at a boundary point $\boldsymbol{\zeta} \in \partial D(\{F\})$.

Theorem 5.5 *Let* $\bar{\mathbf{p}} = (\bar{p}_1, \ldots \bar{p}_L)^\top > \mathbf{0}$ *be given. Consider the convex set (5.1). Let* $\boldsymbol{\zeta}$ *be a boundary point of* $\partial D(\{F\})$. *Then* $\boldsymbol{\zeta} = \log \boldsymbol{\gamma}(\mathbf{p})$, *where* $\mathbf{0} \le \mathbf{p} = (p_1, \ldots, p_L)^\top \le \bar{\mathbf{p}}$. *The set* $\mathcal{B} := \{l \in \langle L \rangle, \ p_l = \bar{p}_l\}$ *is nonempty. For each* $B_l = (F + (1/\bar{p}_l)\mathbf{v}\mathbf{e}_l^\top))$ *let* $\mathrm{H}_l(\xi)$ *be defined as in Theorem 6.2, where* $B = B_l$ *and* $\boldsymbol{\eta} = \boldsymbol{\zeta}$. *Then* $\mathrm{H}_l \le 0$, *for* $l \in \mathcal{B}$, *are the supporting hyperplanes of* $D(\{F\})$ *at* $\boldsymbol{\zeta}$.

Proof. Let $\mathbf{p} = P(e^{\boldsymbol{\zeta}})$. Theorem 2.3 implies the set $\mathcal{B}$ is nonempty. Furthermore, $\rho(e^{\boldsymbol{\zeta}} B_l) = 1$ if and only if $p_l = \bar{p}_l$. Hence, ζ lies exactly at the intersection of the hypersurfaces $\log \rho(e^{\boldsymbol{\zeta}} B_l) = 0, l \in \mathcal{B}$. Theorem 6.2 implies that the supporting hyperplanes of $D(\{F\})$ at $\boldsymbol{\zeta}$ are $\mathrm{H}_l(\boldsymbol{\xi}) \leq 0$ for $l \in \mathcal{B}$. $\square$

We now show how to choose the boundary points $\boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_M \in \partial D(\{F\})$ and to compute the supporting hyperplanes of $D(\{F\})$ 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 5.2. Choose $M_i \geq 2$ equidistant points in each interval $[\underline{p}_i, \bar{p}_i]$.

$$p_{j_i,i} = \frac{j_i \underline{p}_i + (M_i - j_i)\bar{p}_i}{M_i} \text{ for } j_i = 1, \ldots, M_i, \text{ and } i = 1, \ldots, L. \tag{5.7}$$

Let

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

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

$$\{\boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_M\} = \log \boldsymbol{\gamma}(\mathcal{P}).$$

The supporting hyperplanes of $D(\{F\})$ at each $\boldsymbol{\zeta}_i$ are given by Theorem 5.5.

6 Appendix: Friedland-Karlin results

In this section, we recall some results from [8] and state the extensions of these results, and then illustrate their applications in this paper. We first state the following extension of [8, Theorem 3.1]:

Theorem 6.1 *Let $A \in \mathbb{R}_+^{L \times L}$ be an irreducible matrix. Assume that $\mathbf{x}(A) = (x_1(A), \ldots, x_L(A))^\top, \mathbf{y}(A) = (y_1(A), \ldots, y_L(A))^\top > \mathbf{0}$ are left and right Perron-Frobenius eigenvectors of A, normalized such that $\mathbf{x}(A) \circ \mathbf{y}(A)$ is a probability vector Suppose $\boldsymbol{\gamma}$ is a nonnegative vector. Then*

$$\rho(A) \prod_l \gamma_l^{(\mathbf{x}(A)\circ\mathbf{y}(A))_l} \leq \rho(\mathrm{diag}(\boldsymbol{\gamma})A). \tag{6.1}$$

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:

$$\rho(A) \leq \prod_{l=1}^{L} \left(\frac{(A\mathbf{z})_l}{z_l}\right)^{(\mathbf{x}(A)\circ\mathbf{y}(A))_l}. \tag{6.2}$$

If A is an irreducible matrix with positive diagonal elements, then equality holds in (6.2) if and only if $\mathbf{z} = t\mathbf{x}(A)$ for some positive t.

Proof. Theorem 3.1 in [8] makes the following assumptions. First, in the inequality (6.1), it assumes that $\boldsymbol{\gamma} > \mathbf{0}$. Second, in (6.2), it assumes that $\rho(A) = 1$. Third, the equality case in (6.2) for $\mathbf{z} > \mathbf{0}$ is stated for a positive matrix A. We now show how to deduce the stronger version of Theorem 3.1 claimed here.

First, by using the continuity argument, we deduce the validity of (6.1) for any $\boldsymbol{\gamma} \geq \mathbf{0}$. Second, by replacing A by tA, where $t > 0$, we deduce that it is enough to show (6.2) in the case $\rho(A) = 1$.

Third, to deduce the equality case in (6.2) for $\mathbf{z} > \mathbf{0}$, we need to examine the proof of Lemma 3.2 in [8]. The proof of the Lemma 3.2 applies if the following condition holds. For any sequence of probability vectors $\mathbf{z}_i = (z_{1,i}, \ldots, z_{L,i})^\top, i = 1, \ldots,$ which converges to a probability vector $\boldsymbol{\zeta} = (\zeta_1, \ldots, \zeta_L)^\top$, where $\boldsymbol{\zeta}$ has at least one zero coordinate, the function $\prod_{l=1}^{L} \left(\frac{(A\mathbf{z})_l}{z_l}\right)^{(\mathbf{x}(A)\circ\mathbf{y}(A))_l}$ tends to ∞ on the sequence $\mathbf{z}_i, i = 1, \ldots$. Assume that $\mathcal{A} = \{l \in \langle L\rangle, \zeta_l = 0\}$. Note that the complement of $\mathcal{A}$ in $\langle L\rangle$, denoted by $\mathcal{A}^c$ is nonempty.

Since $A = [a_{ij}]$ has positive diagonal entries, it follows that $\frac{(A\mathbf{z})_l}{z_l} \geq a_{ll} > 0$ for each $l \in \langle L\rangle$. Since A is irreducible, we deduce that there exist $l \in \mathcal{A}$ and $m \in \mathcal{A}^c$ such that $a_{lm} > 0$. Hence, $\lim_{i\to\infty} \frac{(A\mathbf{z}_i)_l}{z_{l,i}} = \infty$. This shows that the unboundedness condition holds.

□

The following result gives an interpretation of the inequality (6.1) in terms of the supporting hyperplane of the convex function $\log \rho(e^{\boldsymbol{\xi}} B)$, where $B \in \mathbb{R}_+^{L\times L}$ is irreducible and $\boldsymbol{\xi} \in \mathbb{R}^L$.

Theorem 6.2 *Let $B \in \mathbb{R}_+^{L\times L}$ be an irreducible matrix. Let $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_L)^\top \in \mathbb{R}^L$ satisfy the condition $\rho(e^{\boldsymbol{\eta}} B) = 1$. Denote $A = e^{\boldsymbol{\eta}} B$ and assume that $\mathbf{x}(A) = (x_1(A), \ldots, x_L(A))^\top, \mathbf{y}(A) = (y_1(A), \ldots, y_L(A))^\top > \mathbf{0}$ are left and right Perron-Frobenius eigenvectors of A, normalized such that $\mathbf{x}(A)\circ\mathbf{y}(A)$ is a probability vector. Let*

$$\mathrm{H}(\boldsymbol{\xi}) = \sum_{l=1}^{L} x_l(A) y_l(A)(\xi_l - \eta_l). \tag{6.3}$$

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

Proof. Let $\boldsymbol{\xi} \in \mathbb{R}^L$. Then $e^{\boldsymbol{\xi}} B = e^{\boldsymbol{\xi}-\boldsymbol{\eta}} A$. Theorem 6.1 implies that $\mathrm{H}(\boldsymbol{\xi}) \leq \log \rho(e^{\boldsymbol{\xi}} B)$. Thus, $\mathrm{H}(\boldsymbol{\xi}) \leq 0$ if $\log \rho(e^{\boldsymbol{\xi}} B) \leq 0$. Clearly, $\mathrm{H}(\boldsymbol{\eta}) = 0$. Hence, $\mathrm{H}(\boldsymbol{\xi}) \leq 0$ is a supporting hyperplane of the convex set $\log \rho(e^{\boldsymbol{\xi}} B) \leq 0$. Since the function $\log \rho(e^{\boldsymbol{\xi}} B)$ is a smooth function of $\boldsymbol{\xi}$, it follows that $\mathrm{H}(\boldsymbol{\xi}) \leq 0$ is unique. □

We now give an application of (6.2) in Theorem 6.1.

Theorem 6.3 *Let $B \in \mathbb{R}_+^{L\times L}$ be an irreducible matrix. Let $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_L)^\top \in \mathbb{R}^L$. Let $\mathbf{w} = \mathbf{x}(\mathrm{diag}(e^{\boldsymbol{\eta}})B) \circ \mathbf{y}(\mathrm{diag}(e^{\boldsymbol{\eta}})B) = (w_1, \ldots, w_L)^\top$ be a probability vector. Then for any positive vector $\mathbf{z} = (z_1, \ldots, z_L)^\top$*

$$\sum_{l=1}^{L} w_l \log \frac{z_l}{(B\mathbf{z})_l} \leq -\log \rho(\mathrm{diag}(e^{\boldsymbol{\eta}})B) + \sum_{l=1}^{L} w_l \eta_l. \tag{6.4}$$

If B has a positive diagonal, then equality holds if and only if $\mathbf{z} = t\mathbf{x}(\mathrm{diag}(e^{\boldsymbol{\eta}})B)$ for some $t > 0$.

Proof. Let $A = \mathrm{diag}(e^{\boldsymbol{\eta}})B$. Then

$$\sum_{l=1}^{L} w_l \log \frac{(B\mathbf{z})_l}{z_l} = \sum_{l=1}^{L} w_l \log \frac{(A\mathbf{z})_l}{z_l} - \sum_{l=1} w_l \eta_l.$$

Use (6.2) to deduce (6.4). The equality case follows from the equality case in (6.2). □

We now study the following inverse problem.

Problem 6.4 *Let $B \in \mathbb{R}_+^{L\times L}, \mathbf{w} \in \mathbb{R}_+^L$ be given irreducible matrix and positive probability vector, respectively. When does there exist $\boldsymbol{\eta} \in \mathbb{R}^L$ such that $\mathbf{x}(\mathrm{diag}(e^{\boldsymbol{\eta}})B) \circ \mathbf{y}(\mathrm{diag}(e^{\boldsymbol{\eta}})B) = \mathbf{w}$? If such $\boldsymbol{\eta}$ exists, when is it unique up to an addition $t\mathbf{1}$?*

To solve the inverse problem, we recall Theorem 3.2 in [8].

Theorem 6.5 *Let $A \in \mathbb{R}_+^{L\times L}, \mathbf{u}, \mathbf{v} \in \mathbb{R}_+^L$ be given, where A is irreducible with positive diagonal elements and $\mathbf{u}, \mathbf{v}$ are positive. Then there exists $D_1, D_2 \in \mathbb{R}_+^{L\times L}$ such that*

$$D_1 A D_2 \mathbf{u} = \mathbf{u},\ \mathbf{v}^\top D_1 A D_2 = \mathbf{v}^\top,\ D_1 = \mathrm{diag}(\mathbf{f}),\ D_2 = \mathrm{diag}(\mathbf{g}) \text{ and } \mathbf{f}, \mathbf{g} > \mathbf{0}. \quad (6.5)$$

The pair (D_1, D_2) is unique to the change $(tD_1, t^{-1}D_2)$ for any $t > 0$. There exist $\boldsymbol{\eta} \in \mathbb{R}^L$ such that $\mathbf{x}(\mathrm{diag}(e^{\boldsymbol{\eta}}B)) \circ \mathbf{y}(\mathrm{diag}(e^{\boldsymbol{\eta}}B)) = \mathbf{w}$. Furthermore, $\boldsymbol{\eta}$ is unique up to an addition $t\mathbf{1}$.

Corollary 6.6 *Let $B \in \mathbb{R}_+^{L\times L}, \mathbf{w} \in \mathbb{R}_+^L$ be given irreducible matrix with positive diagonal elements and positive probability vector, respectively. Then there exists $\boldsymbol{\eta} \in \mathbb{R}^L$ such that $\mathbf{x}(\mathrm{diag}(e^{\boldsymbol{\eta}})B) \circ \mathbf{y}(\mathrm{diag}(e^{\boldsymbol{\eta}})B) = \mathbf{w}$. Furthermore, $\boldsymbol{\eta}$ is unique up to an addition of $t\mathbf{1}$.*

Proof. Let $\mathbf{u} = \mathbf{1}, \mathbf{v} = \mathbf{w}$. Then there exists D_1, D_2 two diagonal matrices with positive diagonal entries such that $D_1 B D_2 \mathbf{1} = \mathbf{1}, \mathbf{w}^\top D_1 B D_2 = \mathbf{w}^\top$. Consider the matrix $D_2 D_1 B = D_2(D_1 B D_2)D_2^{-1}$. It is straightforward to see that $\mathbf{x}(D_2 D_1 B) \circ \mathbf{y}(D_2 D_1 B) = \mathbf{w}$. Hence, $\boldsymbol{\eta}$ is the unique solution of $\mathrm{diag}(e^{\boldsymbol{\eta}}) = D_2 D_1$.

Assume that $\boldsymbol{\zeta} \in \mathbb{R}^L$ satisfies $\mathbf{x}(\mathrm{diag}(e^{\boldsymbol{\zeta}})B) \circ \mathbf{y}(\mathrm{diag}(e^{\boldsymbol{\zeta}})B) = \mathbf{w}$. By considering $\tilde{\boldsymbol{\zeta}} = \boldsymbol{\zeta} + t\mathbf{1}$, we may assume that $\rho(\mathrm{diag}(e^{\boldsymbol{\zeta}})B) = 1$. Let $D_4 = \mathrm{diag}(\mathbf{x}(\mathrm{diag}(e^{\boldsymbol{\zeta}})B)$. Then $(D_4^{-1}\mathrm{diag}(e^{\boldsymbol{\zeta}})BD_4)\mathbf{1} = \mathbf{1}$. Let $D_3 = D_4^{-1}\mathrm{diag}(e^{\boldsymbol{\zeta}})$. Hence, $\mathbf{y}(D_3 B D_4) = \mathbf{w}$. In view of Theorem 6.5, $\mathrm{diag}(e^{\boldsymbol{\zeta}}) = D_4 D_3 = D_2 D_1 = \mathrm{diag}(e^{\boldsymbol{\eta}})$. □

Unfortunately, the matrix F defined in (2.2) have zero diagonal entries and positive off-diagonal entries. For $L = 2$, it is easy to show that $\mathbf{x}(F)\circ\mathbf{y}(F) = (\frac{1}{2}, \frac{1}{2})^\top$. In particular, for $L = 2$, Problem 6.4 is not solvable for $\mathbf{w} \neq (\frac{1}{2}, \frac{1}{2})^\top$. Similarly, given positive $\mathbf{u}, \mathbf{v} \in \mathbb{R}^2$ such that $\mathbf{u} \circ \mathbf{v} \neq t(1, 1)$ for any positive t, (6.5) does not hold for $A = F$. For $L \geq 3$, the situation is different.

Theorem 6.7 *Let $L \geq 3, A \in \mathbb{R}_+^{L\times L}, \mathbf{u} = (u_1, \ldots, u_L)^\top, \mathbf{v} = (v_1, \ldots, v_L)^\top \in \mathbb{R}_+^L$ be given, where A is a matrix with zero diagonal entries and positive off-diagonal* *elements, and* $\mathbf{u}, \mathbf{v}$ *are positive. Assume that* $\mathbf{w} = \mathbf{u} \circ \mathbf{v}$ *is a probability vector satisfying the condition*

$$\sum_{j \in \langle L \rangle \backslash \{l\}} w_j > w_l \text{ for all } l \in \langle L \rangle. \tag{6.6}$$

Then there exists $D_1, D_2 \in \mathbb{R}_+^{L \times L}$ *such that (6.5) holds.*

Proof. Let $A_i = A + \frac{1}{i} I, i = 1, \ldots$, where I is the $L \times L$ identity matrix. Theorem 6.5 implies

$$B_i = D_{1,i} A_i D_{2,i},\ D_{1,i} = \mathrm{diag}(\mathbf{f}_i),\ D_{2,i} = \mathrm{diag}(\mathbf{g}_i),\ B_i \mathbf{u} = \mathbf{u},\ \mathbf{v}^\top B_i = \mathbf{v}^\top,$$
$$\mathbf{f}_i = (f_{1,i}, \ldots, f_{L,i})^\top,\ \mathbf{w}_i = (g_{1,i}, \ldots, g_{L,i})^\top,\ s_i = \max_{j \in \langle L \rangle} f_{j,i} = \max_{j \in \langle L \rangle} g_{j,i},\ i = 1, \ldots.$$

Note that each entry of B_i is bounded by $\frac{\max_j u_j}{\min_j u_j}$. By passing to the subsequence $B_{i_k}, \mathbf{f}_{i_k}, \mathbf{g}_{i_k}, 1 \le i_1 < i_2 < \ldots$, we can assume that the first subsequence converges to B, and the last two subsequence converge in generalized sense:

$$\lim_{k \to \infty} B_{i_k} = B = [b_{jl}] \in \mathbb{R}_+^{L \times L},\ \lim_{k \to \infty} \mathbf{f}_{i_k} = \mathbf{f} = (f_1, \ldots, f_L)^\top,\ \lim_{k \to \infty} \mathbf{g}_{i_j} = \mathbf{g} = (g_1, \ldots, g_L)^\top,$$
$$f_j, g_j \in [0, \infty], j = 1, \ldots, L,\ \lim_{k \to \infty} s_{i_k} = s = \max_{j \in \langle L \rangle} f_j = \max_{j \in \langle L \rangle} g_j \in [0, \infty].$$

Note that

$$B\mathbf{u} = \mathbf{u}, \quad \mathbf{v}^\top B = \mathbf{v}^\top. \tag{6.7}$$

Assume first that $s < \infty$. Then $B = \mathrm{diag}(\mathbf{f}) A \,\mathrm{diag}(\mathbf{g})$. In view of (6.7), $\mathbf{f} \circ \mathbf{g} > \mathbf{0}$. This proves the theorem in this case.

Assume now that $s = \infty$. Let

$$\mathcal{F}_\infty = \{j \in \langle L \rangle,\ f_j = \infty\},\ \mathcal{F}_+ = \{j \in \langle L \rangle,\ f_j \in (0, \infty),\}\ \mathcal{F}_0 = \{j \in \langle L \rangle,\ f_j = 0\},$$
$$\mathcal{G}_\infty = \{j \in \langle L \rangle,\ g_j = \infty\},\ \mathcal{G}_+ = \{j \in \langle L \rangle,\ g_j \in (0, \infty),\}\ \mathcal{G}_0 = \{j \in \langle L \rangle,\ g_j = 0\}.$$

Since off-diagonal entries of A are positive, and $B \in \mathbb{R}_+^{L \times L}$ it follows that $\mathcal{F}_\infty = \mathcal{G}_\infty = \{l\}$ for some $l \in \langle L \rangle$. Furthermore, $\mathcal{F}_+ = \mathcal{G}_+ = \emptyset$. So $\mathcal{F}_0 = \mathcal{G}_0 = \langle L \rangle \backslash \{l\}$. Assume first that $l = 1$. Then the principal submatrix $[b_{jl}]_{j=l=2}^L$ is zero. (6.7) yields that

$$b_{j1} = \frac{u_j}{u_1},\ b_{1j} = \frac{v_j}{v_1} \text{ for } j = 2, \ldots, L,\ b_{11} u_1 v_1 + \sum_{j=2}^L u_j v_j = u_1 v_1.$$

Since $b_{11} \ge 0$, the above last equality contradicts the condition (6.6) for $l = 1$. Similar argument implies the impossibility of $\mathcal{F}_\infty = \mathcal{G}_\infty = \{l\}$ for any $l \ge 2$. Hence, $s < \infty$ and we conclude the theorem. □

We do not know whether, under the conditions of Theorem 6.7, the diagonal matrices (D_1, D_2) are unique up to the transformation $(tD_1, t^{-1} D_2)$. We now generalize the above theorem.

Theorem 6.8 *Let*

$$L \ge 3,\ A = [a_{jl}]_{j=l=1}^L \in \mathbb{R}_+^{L \times L},\ \mathbf{0} < \mathbf{u} = (u_1, \ldots, u_L)^\top,\ \mathbf{v} = (v_1, \ldots, v_L)^\top \in \mathbb{R}_+^L$$

be given. Assume that A has positive off-diagonal elements, and $\mathbf{w} = \mathbf{u} \circ \mathbf{v}$ is a probability vector satisfying the condition

$$\sum_{j \in \langle L \rangle \setminus \{l\}} w_j > w_l \tag{6.8}$$

for each l such that $a_{ll} = 0$. Then there exists $D_1, D_2 \in \mathbb{R}_+^{L \times L}$ such that (6.5) holds.

Proof. In view of Theorems 6.5 and 6.7, it is enough to assume that A has positive and zero diagonal entries. Apply the proof of Theorem 6.7 and the following observation. If $\mathcal{F}_\infty = \mathcal{G}_\infty = \{l\}$ then $a_{ll} = 0$. □

Corollary 6.9 *Let $B = [b_{jl}]_{j=l=1}^{L} \in \mathbb{R}_+^{L \times L}, \mathbf{w} \in \mathbb{R}_+^L$ be given matrix with positive off-diagonal elements and a positive probability vector, respectively. Assume that $L \geq 3$ and $\mathbf{w}$ satisfies the conditions (6.8) for each l such that $b_{ll} = 0$. Then there exists $\boldsymbol{\eta} \in \mathbb{R}^L$ such that $\mathbf{x}(\mathrm{diag}(e^{\boldsymbol{\eta}})B) \circ \mathbf{y}(\mathrm{diag}(e^{\boldsymbol{\eta}})B) = \mathbf{w}$.*

References

[1] J. M. Aein, Power balancing in systems employing frequency reuse, *COMSAT Technical Review*, 3 (1973), 277-299, 1973.

[2] M. Avriel, *Nonlinear Programming: Analysis and Methods*, Dover Publishing, 2003.

[3] A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, Academic Press, 1979.

[4] M. Charafeddine, A. Sezgin and A. Paulraj, Rate region frontiers for n-user interference channel with interference as noise, *Proc. of 45th Allerton*, 2007.

[5] 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, 2007.

[6] G. B. Dantzig, *Linear Programming and Extensions*, Princeton, NJ: Princeton University Press, 1963.

[7] , M. Ebrahimi and M. A. Maddah-Ali and A. K. Khandani, Power Allocation and Asymptotic Achievable Sum-Rates in Single-Hop Wireless Networks, *Proc. of IEEE 40th Annual Conference on Information Sciences and Systems*, 2006.

[8] 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.

[9] F. R. Gantmacher, *The Theory of Matrices*, Vol. II, Chelsea Publ. Co., New York, 1974.

[10] L. G. Khachian, A Polynomial Algorithm in Linear Programming, *Dokl. Akad. Nauk SSSR* 244 (1979), 1093-1096. English translation in *Soviet Math. Dokl.* 20, 191-194, 1979.

[11] J. F. C. Kingman, A convexity property of positive matrices, *Quart. J. Math. Oxford Ser.* 12 (2), 283-284, 1961.

[12] H. W. Kuhn and A. W. Tucker, *Nonlinear programming*, Proceedings of 2nd Berkeley Symposium: 481-492, Berkeley: University of California Press, 1951.

[13] Z.-Q. Luo and W. Yu, An Introduction to Convex Optimization for Communications and Signal Processing, (Tutorial Paper), *IEEE Journal on Selected Areas in Communications*, 24 (8), 1426-1438, 2006.

[14] 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.

[15] J. Nocedal and S. J. Wright, *Numerical Optimization*, Springer, 1999.

[16] R. T. Rockafeller, *Convex Analysis*, Princeton University Press, 1970.

[17] R. Srikant, *The Mathematics of Internet Congestion Control*, Birkhauser, Boston, 2003.

[18] C. W. Tan, D. P. Palomar and M. Chiang, Solving Nonconvex Power Control Problems in Wireless Networks: Low SIR Regime and Distributed Algorithms, *Proc. of IEEE Globecom*, 2005.

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

[20] R. S. Varga, Matrix iterative analysis, *Prentice Hall*, 1963.

[21] H. Wielandt, Unzerlegbare nicht-negative Matrizen, *Math. Z.* 52 (1), 642-648, 1950.

[22] 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.