A Game Theoretic Approach to Power Optimization in MIMO Interference Systems

Gürdal Arslan, M. Fatih Demirkol, and Serdar Yüksel

August 19, 2009

Abstract

A multi-link and multi-input-multi-output (MIMO) interference system is considered in which each link wishes to minimize its own power by choosing its own signal vector subject to an information theoretic Quality-of-Service (QoS) requirement. This setup leads to a multi-link game, referred to as a "power game", in which the feasible strategy set of an individual link depends on the strategies of the other links. The rates for which an equilibrium solution exists in a power game is characterized in terms of the equilibria of "capacity games" introduced in our earlier work. An example is provided where the set of equilibrium rates is properly contained in the set of achievable rates. A conservative estimate of the region of equilibrium rates is provided using a minmax approach. The results are extended to the case where the QoS requirements are relaxed. Finally, the uniqueness of equilibrium as well as the convergence of the best response dynamics (a.k.a. iterative water-filling) is shown for all rates when the interference is sufficiently small and some other mild conditions are met.

This work was supported by NSF Grant #ECCS-0547692 and Natural Sciences and Engineering Research Council of Canada (NSERC) in the form of a Discovery Grant. This paper appeared in part at the GameNets 2009 International Conference on Game Theory for Networks, 2009, Turkey.

G. Arslan is with the Department of Electrical Engineering, University of Hawaii at Manoa, 440 Holmes Hall, 2540 Dole Street, Honolulu, HI 96822, USA. E-mail: `gurdal@hawaii.edu`

M. F. Demirkol is with Turkcell Iletisim Hizmetleri, A.S., Istanbul, TURKEY. E-mail: `demirkol@ieee.org`

S. Yüksel is with the Department of Mathematics and Statistics, Queen's University, Kingston, Ontario, CANADA, K7L 3N6. E-mail: `yuksel@mast.queensu.ca`

Index Terms

Power control, MIMO systems, Co-channel interference, Ad-hoc networks, Game theory, Generalized Nash equilibrium, Iterative Water-Filling Algorithm.

I. INTRODUCTION

Multiple-Input Multiple-Output (MIMO) links use antenna arrays at both ends of a link to transmit multiple parallel streams in the same time and frequency channel [1], [2]. Signals transmitted and received by array elements at different physical locations are spatially separated by array processing algorithms. Depending on the channel scattering conditions, MIMO links can yield large gains in capacity of wireless systems.

Using antenna arrays at both ends of the links can also allow the network to accommodate multiple nearby links to transmit in the same time and frequency channel through spatial multiplexing. In this scheme, multiple links, each with different transmitter-receiver pairs, are allowed to transmit in a given range possibly through multiple streams per link.

Such a multi-access network with MIMO links is referred to as a MIMO interference system and has been considered in previous studies including [3]–[7], with the focus of finding the achievable rates of the links given their power and interference levels. Our earlier work [6] uses a multi-link game framework to analyze the MIMO interference system where each link selfishly wishes to maximize its own mutual information. In [6], the existence of Nash equilibrium associated with the multi-link game is established and sufficient conditions for the uniqueness of equilibrium are given. Decentralized algorithms are suggested as update strategies to determine the link parameters using only local information and reasonable computational burden. Since an equilibrium state does not necessarily maximize the total mutual information, a stream control approach is introduced to achieve a system-level coordination based on link negotiations. With the stream control approach, the system-wide efficiency of equilibrium is improved by imposing limits on the number of independent data streams of each link.

In this paper, we consider a MIMO interference network in which the objective of each link is to satisfy a certain Quality of Service (QoS) requirement, defined in terms of the achievable data rate of the link, with minimum possible total radiated power. In [8], an iterative method is used to determine link parameters at equilibrium for such a MIMO network. It is shown that the converged state does not necessarily yield the best network throughput. Previous work such as [9]–[11] have considered joint optimization of sets of co-channel links assuming that the base-station has an array antenna. However, the models considered in these papers involve the set of scalar power levels as the decision parameters in contrast to the model considered in this paper where the decision parameters are the covariance matrices of the transmitted signals.

We model the interactions among the links within the framework of noncooperative game theory and present a multi-link power game where each link's strategy has an effect on the strategy sets of the other links. As in [6], we follow a decentralized approach and assume availability of only local information; i.e. one link has knowledge of only its own channel and received interference conditions. We discuss the existence, uniqueness, and decentralized computation of generalized Nash equilibrium in multi-link power games.

The rest of this paper is organized as follows. Section II presents the system model. Section III introduces the power game setup. Section IV discusses the relationship between the power games and the so-called capacity games introduced in our earlier work. Section V is on obtaining a conservative estimate of the region of equilibrium rates. Section VI provides a discussion on relaxing the QoS requirements of the links. Section VII is devoted to the best response dynamics as decentralized link adjustment algorithms. Section VIII presents some simulation results. Section IX concludes the paper.

A. *Notation*

- $:=$ stands for "defined as".
- $\equiv$ stands for "identically equal to".
- $E(.)$ denotes the expectation.

- $\dagger$ denotes the conjugate transpose.
- $\mathbf{I}$ denotes an identity matrix of an appropriate dimension.
- $|\mathbf{A}|$ denotes the determinant of a square matrix.
- $\text{tr}(\mathbf{A})$ denotes the trace of a square matrix $\mathbf{A}$.
- $-k$ denotes the set of indices other than k.
- $\mathbb{R}$ denotes the set of real numbers; $\mathbb{R}^n$ denotes the $n-$dimensional Euclidian vector space; $|\mathbf{x}|$ denotes the usual Euclidian norm for a vector $\mathbf{x} \in \mathbb{R}^n$.
- $\mathbb{R}^n_+ = \{x \in \mathbb{R} : x_i \geq 0, \text{ for all } i \in \{1, \ldots, n\}\}$; $\mathbb{R}^n_{++} = \{x \in \mathbb{R} : x_i > 0, \text{ for all } i \in \{1, \ldots, n\}\}$; $\mathbb{R}^n_-$ and $\mathbb{R}^n_{--}$ are defined analogously.
- $\mathcal{H}$ denotes the Hilbert space of $m \times m$ complex Hermitian matrices, regarded as $\mathbb{R}^{m^2}$, endowed with the inner product $< A, B >:= \text{tr}(A^\dagger B)$ and the corresponding (Frobenius) norm $\|.\|_F$ (where m will be clear from the context).
- $\mathcal{H}_+$ denotes the closed convex cone of positive semi-definite matrices in $\mathcal{H}$; $\mathcal{H}^n_+$ denotes the $n-$times product $\mathcal{H}_+ \times \cdots \times \mathcal{H}_+$.
- $\lambda_{\min}(\mathbf{A})$ and $\lambda_{\max}(\mathbf{A})$ denote the minimum and maximum eigenvalue of a matrix $\mathbf{A} \in \mathcal{H}$, respectively.
- $\text{diag}(a_1, a_2, \ldots)$ denotes the diagonal matrix whose diagonal entries are the scalars $a_1, a_2, \ldots$.
- $(a)^+ = \max\{a, 0\}$ for a real a.
- $\mathcal{F} : \mathcal{X} \rightrightarrows \mathcal{Y}$ indicates that $\mathcal{F}$ is a correspondence mapping $\mathcal{X}$ into the set of subsets of $\mathcal{Y}$.
- $\text{gr}(\mathcal{F})$ denotes the graph of a correspondence $\mathcal{F} : \mathcal{X} \rightrightarrows \mathcal{Y}$, i.e., $\text{gr}(\mathcal{F}) = \{(\mathbf{X}, \mathbf{Y}) \in \mathcal{X} \times \mathcal{Y} : \mathbf{X} \in \mathcal{X},\ \mathbf{Y} \in \mathcal{F}(\mathbf{X})\}$.
- $\text{int}(\mathcal{S})$ denotes the interior of a set $\mathcal{S}$.
- $r(h) = O(h)$ means that $r(h)/\|h\|$ is bounded for all h in a neighborhood of 0.

II. SYSTEM MODEL

We consider an L-link communication system where each link is associated with a transmitter-receiver pair. Each transmitter and receiver are equipped with N_t and N_r antennas, respectively. We assume link k,

$k \in \{1, \ldots, L\}$, transmits a complex signal vector $\mathbf{x}_k$ of dimension N_t. Consequently, a complex baseband signal vector of dimension N_r denoted by $\mathbf{y}_k$ is received at the $k-$th receiver. The received signal vectors are related to the transmitted signal vectors by

$$\mathbf{y}_k = \mathbf{H}_{k,k}\mathbf{x}_k + \sum_{\ell \neq k} \mathbf{H}_{k,\ell}\mathbf{x}_\ell + \mathbf{n}_k$$

where

- $\mathbf{H}_{k,\ell}$ is the complex channel matrix of dimension $N_r \times N_t$ for the link between the $\ell-$th transmitter and the $k-$th receiver,
- $\mathbf{n}_k$ denotes the zero-mean circularly symmetric complex Gaussian noise vector at the $k-$th receiver with $E(\mathbf{n}_k\mathbf{n}_k^\dagger) = \mathbf{I}$.

To avoid trivialities, we make the following assumption throughout the paper.

Assumption 1:

$$\mathbf{H}_{k,k} \neq 0, \qquad \text{for all } k.$$

We now consider a scenario in which the $k-$th link wishes to minimize its power

$$E(\mathbf{x}_k^\dagger\mathbf{x}_k)$$

by choosing the distribution of $\mathbf{x}_k$, independently of the other links, subject to a QoS constraint

$$I(\mathbf{x}_k; \mathbf{y}_k) \geq r_k \tag{1}$$

where $I(\mathbf{x}_k; \mathbf{y}_k)$ is the mutual information between the input and the output of the channel characterized by $\mathbf{H}_{k,k}$ and $r_k \in \mathbb{R}_+$ is a given constant. The $k-$th link, not knowing the distributions of the signal vectors chosen by the other links, models the total interference $\sum_{\ell \neq k} \mathbf{H}_{k,\ell}\mathbf{x}_\ell$ at its receiver as a zero-mean circularly symmetric complex Gaussian noise vector. Under the modeling assumptions delineated above, the $k-$th link's power $E(\mathbf{x}_k^\dagger\mathbf{x}_k)$ is minimized by a zero-mean circularly symmetric complex Gaussian distribution satisfying the QoS constraint (1); see [1]. Note that if all links make the same modeling assumptions then the links can choose their optimal distributions in a manner that is mutually consistent with their modeling assumptions. We assume that this is the case, and note from [1] that the mutual information of each link can now be written as

$$I(\mathbf{x}_k;\mathbf{y}_k) = \log_2 \left|\mathbf{I} + \mathbf{R}_k^{-1/2}\mathbf{H}_{k,k}\mathbf{Q}_k\mathbf{H}_{k,k}^{\dagger}\mathbf{R}_k^{-1/2}\right| \tag{2}$$

for $k \in \{1,\ldots,L\}$, where $\mathbf{Q}_k := E(\mathbf{x}_k\mathbf{x}_k^{\dagger})$ is a Hermitian positive semi-definite matrix, and

$$\mathbf{R}_k := \mathbf{I} + \sum_{\ell \neq k} \mathbf{H}_{k,\ell}\mathbf{Q}_\ell\mathbf{H}_{k,\ell}^{\dagger} \tag{3}$$

is the covariance matrix of the total interference and noise at the $k-$th receiver[1]. We furthermore assume that both the transmitter and receiver nodes of the $k-$th link have the knowledge of the whitened channel matrix $\mathbf{R}_k^{-1/2}\mathbf{H}_{k,k}$ once the other links choose their signal vectors. Now, from the perspective of each link, the problem amounts to choosing an appropriate covariance matrix with minimum trace satisfying its own QoS constraint in the presence of the other links that also want to minimize their own powers satisfying their own QoS constraints.

III. A POWER GAME AND GENERALIZED NASH EQUILIBRIUM

The setup introduced in the previous section leads us to an $L-$link noncooperative game with cost functions

$$J_k^p(\mathbf{Q}_k) := \mathrm{tr}(\mathbf{Q}_k)$$

and feasible strategy sets

$$\mathcal{F}_k^p(\mathbf{Q}_{-k}) := \{\mathbf{Q}_k \in \mathcal{H}_+ : r_k - I(\mathbf{x}_k;\mathbf{y}_k) \leq 0\} \tag{4}$$

where

- $\mathbf{Q}_{-k} := (\mathbf{Q}_1,\ldots,\mathbf{Q}_{k-1},\mathbf{Q}_{k+1},\ldots,\mathbf{Q}_L)$,
- $I(\mathbf{x}_k;\mathbf{y}_k)$ is as in (2),
- $r_k \in \mathbb{R}_+$ is a given constant,

[1] In the remainder of the paper, we often suppress the dependence of the mutual information $I(\mathbf{x}_k;\mathbf{y}_k)$ on the link covariance matrices $\mathbf{Q}_1,\ldots,\mathbf{Q}_L$.

for $k \in \{1, \dots, L\}$. We call the above game a power game and denote it by $\Gamma^p(\mathbf{H}, \mathbf{r})$ where $\mathbf{H} := (\mathbf{H}_{k,\ell})_{1 \leq k,\ell \leq L}$, and $\mathbf{r} := (r_1, \dots, r_L)$.

A selfish link in such a strategic engagement would not be satisfied with its choice unless its cost is minimized given the choices of the other links. A steady state situation in which all link costs are mutually minimized is called a generalized Nash equilibrium[2]. For a more precise definition of equilibrium, let BR_k^p denote the $k-$th link's best response function, i.e.,

$$BR_k^p(\mathbf{Q}_{-k}) := \text{argmin}_{\mathbf{Q}_k \in \mathcal{F}_k^p(\mathbf{Q}_{-k})} J_k^p(\mathbf{Q}_k) \tag{5}$$

and let $BR^p := (BR_1^p, \dots, BR_L^p)$ denote the composite best response function; see Proposition 9 in Appendix for the fact that, for any given $\mathbf{Q} \in \mathcal{H}_+^L$, $BR^p(\mathbf{Q})$ is nonempty and single-valued. Now, a profile of link strategies $\mathbf{Q}^* = (\mathbf{Q}_1^*, \dots, \mathbf{Q}_L^*)$ is called an equilibrium of $\Gamma^p(\mathbf{H}, \mathbf{r})$ if

$$\mathbf{Q}^* = BR^p(\mathbf{Q}^*).$$

An equilibrium represents a steady-state situation in which no link has an incentive to unilaterally change its strategy. As such, equilibrium is a particularly useful notion when it is not practical to obtain and/or implement a system-wide optimal solution. For example, an equilibrium can emerge out of local optimizations performed by autonomous links in an ad-hoc wireless network without centralized coordination. Therefore, it is important to address the issue of existence and uniqueness of an equilibrium in a power game.

One approach used in [12] to establish the existence of equilibrium in single-input-single-output (SISO) power games is the framework of supermodular games which relies on the monotonicity of the best response function. However, this approach is not readily applicable to the MIMO case. For instance, BR^p is not always non-decreasing with respect to the partial order $\mathbf{Q}^1 \leq \mathbf{Q}^2 \Leftrightarrow \mathbf{Q}^2 - \mathbf{Q}^1 \in \mathcal{H}_+^L$, i.e.,

$$\mathbf{Q}^1 \leq \mathbf{Q}^2 \not\Rightarrow BR^p(\mathbf{Q}^1) \leq BR^p(\mathbf{Q}^2).$$

An example is provided below.

---

[2] We henceforth refer to a generalized Nash equilibrium simply as an equilibrium.

Example 1: Consider a power game $\Gamma^p(\mathbf{H}, \mathbf{r})$ where $L = 2,\ N_r = N_t = 2,\ \mathbf{H}_{1,1} = \mathbf{I}, \mathbf{H}_{1,2} = \mathrm{diag}(1, 0),\ r_1 = 1$. We have $BR_1^p(\mathbf{I}/2) = \mathrm{diag}(\sqrt{3} - 3/2, \sqrt{3} - 1)$, whereas $BR_1^p(\mathbf{I}) = \mathrm{diag}(0, 1)$.

Another approach to establish the existence of an equilibrium in noncooperative games is based on Kakutani's well-known fixed point theorem. Towards this end, we state a fundamental existence theorem specialized from Theorem 4.3.1 of [13] to our context.

Theorem 1 (Theorem 4.3.1 of [13]): Let us consider an $L-$player noncooperative game where the $k-$th player's strategy $\mathbf{Q}_k$ belongs to a subset $\mathcal{C}_k$ of a Euclidean space. Let $\mathcal{C} := \times_\ell \mathcal{C}_\ell$ and $\mathcal{C}_{-k} := \times_{\ell \neq k} \mathcal{C}_\ell$. Let $\mathcal{F}_k : \mathcal{C}_{-k} \rightrightarrows \mathcal{C}_k$ be the feasible strategy correspondence for player k such that $\mathcal{F}_k(\cdot) = \{\mathbf{Q}_k \in \mathcal{C}_k : g_k(\mathbf{Q}_k, \cdot) \leq 0\}$ for some $g_k : \mathcal{C} \mapsto \mathbb{R}$. Let $J_k : \mathrm{gr}(\mathcal{F}_k) \mapsto \mathbb{R}$ be the cost function for player k. Assume, for all k, that

(i) $\mathcal{C}_k$ is nonempty, convex and compact,

(ii) g_k is continuous in $\mathcal{C}$,

(iii) for any fixed $\bar{\mathbf{Q}}_{-k} \in \mathcal{C}_{-k}$, $g_k(\cdot, \bar{\mathbf{Q}}_{-k})$ is convex in $\mathcal{C}_k$,

(iv) for any fixed $\bar{\mathbf{Q}}_{-k} \in \mathcal{C}_{-k}$, there exists a $\bar{\mathbf{Q}}_k \in \mathcal{C}_k$ such that $g_k(\bar{\mathbf{Q}}_k, \bar{\mathbf{Q}}_{-k}) < 0$,

(v) J_k is continuous in $\mathrm{gr}(\mathcal{F}_k)$,

(vi) for any fixed $\bar{\mathbf{Q}}_{-k} \in \mathcal{C}_{-k}$, $J_k(\cdot, \bar{\mathbf{Q}}_{-k})$ is convex in $\mathcal{F}_k(\bar{\mathbf{Q}}_{-k})$.

Then, there exists an equilibrium.

Remark 1: Theorem 4.3.1 of [13] allows $\mathcal{F}_k$ to be arbitrary provided it is nonempty, closed, and convex valued in $\mathcal{C}_{-k}$ as well as it is both upper semi continuous and lower semi continuous in $\mathcal{C}_{-k}$. In Theorem 1, the conditions imposed on g_k are sufficient for $\mathcal{F}_k$ to satisfy such requirements [3]. Finally, if $\mathcal{F}_k \equiv \bar{\mathcal{F}}$ for some fixed nonempty, closed, and convex subset $\bar{\mathcal{F}}$ of $\mathcal{C}_k$, for all k, then assumptions (ii), (iii), (iv) of Theorem 1 are superfluous.

---

[3] In particular, see Theorem 2.2.3 of [13] for the upper semi continuity requirement, and see Theorem 12 of [14] for the lower semi continuity requirement.

It turns out that Theorem 1 is not immediately applicable to a power game $\Gamma^p(\mathbf{H}, \mathbf{r})$. Note that, if we set $\mathcal{C}_k = \mathcal{H}_+$, $g_k = r_k - I(\mathbf{x}_k; \mathbf{y}_k)$, and $J_k = J_k^p$, the assumptions (ii) through (vi) of Theorem 1 are satisfied, however, assumption (i) is not satisfied. The main difficulty stems from the fact that link feasible strategy sets $\mathcal{F}_k^p(\mathbf{Q}_{-k})$, given in (4), are unbounded. Moreover, imposing bounds on link strategies in a way that is consistent with the assumptions of Theorem 1 is not straightforward. For example, it is not clear how to impose some additional power limitations $\bar{p}_1, \ldots, \bar{p}_L$ such that the modified strategy sets

$$\{\mathbf{Q}_k \in \mathcal{H}_+ : \text{tr}(\mathbf{Q}_k) \leq \bar{p}_k, \ r_k - I(\mathbf{x}_k; \mathbf{y}_k) \leq 0\}$$

are nonempty for all k. This prompts us to follow alternative routes towards establishing the existence of equilibrium in power games. We first observe a useful relationship with a capacity game.

IV. RELATIONSHIP WITH A CAPACITY GAME

Let $\mathbf{p} := (p_1, \ldots, p_L) \in \mathbb{R}_+^L$. Consider an $L-$link noncooperative game with utility functions

$$I(\mathbf{x}_k; \mathbf{y}_k)(\mathbf{Q}) \tag{6}$$

and feasible strategy sets

$$\{\mathbf{Q}_k \in \mathcal{H}_+ : \text{tr}(\mathbf{Q}_k) \leq p_k\} \tag{7}$$

for all k, where $\mathbf{Q} = (\mathbf{Q}_1, \ldots, \mathbf{Q}_L)$, and $I(\mathbf{x}_k; \mathbf{y}_k)$ is as in (2). Here, each link wishes to maximize its utility (6) by choosing a strategy $\mathbf{Q}_k$ from its feasible strategy set (7) We call this game a *capacity game* and denote it by $\Gamma^c(\mathbf{H}, \mathbf{p})$.

The equilibria of capacity games have been studied in [6]. In particular, the existence of an equilibrium in every capacity game $\Gamma^c(\mathbf{H}, \mathbf{p})$ has been established using the framework of concave games [15][4]. The following proposition, whose proof is revealed by a little thought, relates the equilibria of power games and capacity games.

---

[4] The existence of an equilibrium in any capacity game $\Gamma^c(\mathbf{H}, \mathbf{p})$ also follows from Theorem 1.

Proposition 1: Fix $\mathbf{H}$. Consider $\tilde{\mathbf{Q}} \in \mathcal{H}_+^L$, and let

$$\tilde{\mathbf{p}} := \big(J_1^p(\tilde{\mathbf{Q}}_1), \dots, J_L^p(\tilde{\mathbf{Q}}_L)\big), \qquad \tilde{\mathbf{r}} := \big(I(\mathbf{x}_1; \mathbf{y}_1)(\tilde{\mathbf{Q}}), \dots, I(\mathbf{x}_L; \mathbf{y}_L)(\tilde{\mathbf{Q}})\big).$$

Then,

$$\tilde{\mathbf{Q}} \text{ is an equilibrium of } \Gamma^c(\mathbf{H}, \tilde{\mathbf{p}}) \qquad \Leftrightarrow \qquad \tilde{\mathbf{Q}} \text{ is an equilibrium of } \Gamma^p(\mathbf{H}, \tilde{\mathbf{r}}). \tag{8}$$

This implies that a power game $\Gamma^p(\mathbf{H}, \mathbf{r})$ would possess an equilibrium if and only if the rate profile $\mathbf{r}$ can be achieved at an equilibrium of a capacity game $\Gamma^c(\mathbf{H}, \mathbf{p})$ for some power profile $\mathbf{p}$. This leads us to the question of how we can characterize the set of rate profiles that can be achieved at an equilibrium of $\Gamma^c(\mathbf{H}, \mathbf{p})$ for some $\mathbf{p}$.

For a fixed $\mathbf{H}$, define the set of *equilibrium rates* as

$$\mathcal{R}_e(\mathbf{H}) \;:=\; \big\{\big(I(\mathbf{x}_1; \mathbf{y}_1)(\mathbf{Q}^*), \dots, I(\mathbf{x}_L; \mathbf{y}_L)(\mathbf{Q}^*)\big) : \mathbf{Q}^* \text{ is an equilibrium of } \Gamma^c(\mathbf{H}, \mathbf{p}) \text{ for some } \mathbf{p}\big\}$$

and the set of *achievable rates* as

$$\mathcal{R}_a(\mathbf{H}) := \big\{\big(I(\mathbf{x}_1; \mathbf{y}_1)(\mathbf{Q}), \dots, I(\mathbf{x}_L; \mathbf{y}_L)(\mathbf{Q})\big) : \mathbf{Q} \in \mathcal{H}_+^L\big\}.$$

Clearly, $\mathcal{R}_e(\mathbf{H}) \subset \mathcal{R}_a(\mathbf{H})$. Moreover, in some special cases, we have $\mathcal{R}_e(\mathbf{H}) = \mathcal{R}_a(\mathbf{H})$.

Proposition 2: If $N_r = N_t = 1$, then $\mathcal{R}_e(\mathbf{H}) = \mathcal{R}_a(\mathbf{H})$.

Proof: It follows from the fact that, in the SISO case, $\mathbf{p}$ is the unique equilibrium of $\Gamma^c(\mathbf{H}, \mathbf{p})$. ■

It is possible, however, to find some $\mathbf{H}$ for which $\mathcal{R}_e(\mathbf{H})$ is a proper subset of $\mathcal{R}_a(\mathbf{H})$, i.e.,

$$\mathcal{R}_e(\mathbf{H}) \subsetneq \mathcal{R}_a(\mathbf{H}).$$

Example 2: Consider the setup $L = 2,\ N_r = N_t = 2,\ \mathbf{H}_{1,1} = \mathbf{H}_{2,2} = \mathbf{I},\ \mathbf{H}_{1,2} = \mathbf{H}_{2,1} = \sqrt{\eta}\mathbf{I}$, for some $\eta \geq 0$. Any given rate profile $\mathbf{r} \in \mathbb{R}_+^2$ can be achieved by $\mathbf{Q}_1 = \mathrm{diag}(2^{r_1} - 1, 0)$, $\mathbf{Q}_2 = \mathrm{diag}(0, 2^{r_2} - 1)$. Therefore, $\mathcal{R}_a(\mathbf{H}) = \mathbb{R}_+^2$.

However, if $\eta < 1$, then the unique equilibrium of $\Gamma^c(\mathbf{H}, \mathbf{p})$ for a given $\mathbf{p} \in \mathbb{R}_+^2$ is $(p_1\mathbf{I}/2, p_2\mathbf{I}/2)$ with the corresponding rate profile $\left(\log_2\left(1 + \frac{p_1/2}{1+\eta p_2/2}\right)^2, \log_2\left(1 + \frac{p_2/2}{1+\eta p_1/2}\right)^2\right)$. This implies that, for $\eta < 1$,

$$\mathcal{R}_e(\mathbf{H}) = \left\{\mathbf{r} \in \mathbb{R}_+^2 \ : \ \left(\sqrt{2^{r_1}} - 1\right)\left(\sqrt{2^{r_2}} - 1\right) < 1/\eta^2\right\}.$$

Figure 1 illustrates $\mathcal{R}_e(\mathbf{H})$ for the case where $\eta = 1/2$.

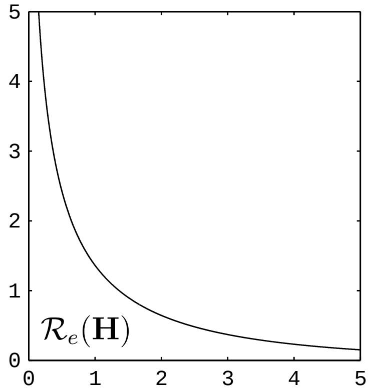

Fig. 1. An illustration of $\mathcal{R}_e(\mathbf{H})$ in Example 2 for the case $\eta = 1/2$.

V. A CONSERVATIVE ESTIMATE OF $\mathcal{R}_e(\mathbf{H})$

Here, we present an estimate of the set of equilibrium rates, which is conservative but relatively easier to compute. We recall that Theorem 1 is not applicable to establish an equilibrium in a power game mainly because feasible strategy sets are unbounded. To work around this issue, consider a nonempty, convex and compact subset $\mathcal{C} \in \mathcal{H}_+^L$ of the form

$$\mathcal{C} = \mathcal{C}_1 \times \cdots \times \mathcal{C}_L \tag{9}$$

where, for some fixed $\mathbf{p} \in \mathbb{R}_+^L$,

$$\mathcal{C}_k = \{\mathbf{Q}_k \in \mathcal{H}_+ : \text{tr}(\mathbf{Q}_k) \leq p_k\}, \qquad \text{for all } k. \tag{10}$$

Let $\mathcal{C}_{-k} := \times_{\ell \neq k} \mathcal{C}_\ell$. We now define a set of minmax rates as

$$\mathcal{R}_m(\mathbf{H}) := \bigcup_{\mathcal{C}} \left\{\mathbf{r} \in \mathbb{R}_+^L \ : \ r_k < \min_{\mathbf{Q}_{-k} \in \mathcal{C}_{-k}} \max_{\mathbf{Q}_k \in \mathcal{C}_k} I(\mathbf{x}_k; \mathbf{y}_k)(\mathbf{Q}_k, \mathbf{Q}_{-k}), \ \text{for all } k\right\}$$

where the union is taken with respect to subsets $\mathcal{C} \in \mathcal{H}_+^L$ of the form (9)-(10). Loosely speaking, $\mathcal{R}_m$ represents the rates that are achievable irrespective of the interference when the link strategies belong to *some* nonempty, convex and compact set.

Proposition 3:

$$\mathcal{R}_m(\mathbf{H}) \subset \mathcal{R}_e(\mathbf{H}), \text{ (hence } \mathcal{R}_m(\mathbf{H}) \subset \mathcal{R}_e(\mathbf{H}) \subset \mathcal{R}_a(\mathbf{H})) .$$

Proof: For a fixed $\mathbf{H}$, consider the power game $\Gamma^p(\mathbf{H}, \mathbf{r})$ with $\mathbf{r} \in \mathcal{R}_m(\mathbf{H})$. Let $\mathcal{C}^{\mathbf{r}} \in \mathcal{H}_+^L$ be a nonempty, convex and compact subset of the form (9)-(10) such that

$$r_k < \min_{\mathbf{Q}_{-k} \in \mathcal{C}^{\mathbf{r}}_{-k}} \max_{\mathbf{Q}_k \in \mathcal{C}^{\mathbf{r}}_k} I(\mathbf{x}_k; \mathbf{y}_k)(\mathbf{Q}_k, \mathbf{Q}_{-k}), \qquad \text{for all } k.$$

It is clear that

$$\mathbf{Q}_{-k} \in \mathcal{C}^{\mathbf{r}}_{-k} \Rightarrow BR^p_k(\mathbf{Q}_{-k}) \in \mathcal{C}^{\mathbf{r}}_k.$$

Hence, the composite best response function BR^p maps $\mathcal{C}^{\mathbf{r}}$ into $\mathcal{C}^{\mathbf{r}}$, i.e., $BR^p(\mathcal{C}^{\mathbf{r}}) \subset \mathcal{C}^{\mathbf{r}}$. Now, by Theorem 1, the restriction of $\Gamma^p(\mathbf{H}, \mathbf{r})$ to $\mathcal{C}^{\mathbf{r}}$ possesses an equilibrium, which is also an equilibrium of $\Gamma^p(\mathbf{H}, \mathbf{r})$. Therefore, $\mathcal{R}_m(\mathbf{H}) \subset \mathcal{R}_e(\mathbf{H})$. ■

Proposition 4: Fix $(\mathbf{H}_{k,k})_{k \in \{1,\dots,L\}}$ and $\mathbf{r} \in \mathbb{R}_+^L$. If the interference channels are sufficiently weak, i.e., $\boldsymbol{\eta} := (\mathbf{H}_{j,\ell})_{j \neq \ell}$ is sufficiently small, then

$$\mathbf{r} \in \mathcal{R}_m(\mathbf{H}), \text{ hence } \mathbf{r} \in \mathcal{R}_e(\mathbf{H}).$$

Proof: It is easy to show that, for any $\mathbf{Q}_{-k} \in \mathcal{H}_+^{L-1}$, $p_k \in \mathbb{R}_+$,

$$\max_{\mathbf{Q}_k \in \mathcal{H}_+ : \mathrm{tr}(\mathbf{Q}_k) \leq p_k} I(\mathbf{x}_k; \mathbf{y}_k)(\mathbf{Q}_k, \mathbf{Q}_{-k}) \geq \log_2 \left(1 + p_k \frac{\lambda_{\max}(\mathbf{H}^\dagger_{k,k} \mathbf{H}_{k,k})}{1 + \|\boldsymbol{\eta}\|^2 \max_{\ell \neq k} \mathrm{tr}(\mathbf{Q}_\ell)} \right) .$$

where $\|\boldsymbol{\eta}\| = \sqrt{\sum_{j \neq \ell} \|\mathbf{H}_{j,\ell}\|_F^2}$. For any $\bar{\mathbf{r}} \in \mathbb{R}_+^L$, there exists a $\bar{\mathbf{p}} \in \mathbb{R}_+^L$ such that

$$\log_2 \left(1 + \bar{p}_k \lambda_{\max}(\mathbf{H}^\dagger_{k,k} \mathbf{H}_{k,k})\right) = 2\bar{r}_k, \qquad \text{for all } k.$$

Consider the nonempty, convex and compact subset $\bar{\mathcal{C}} := \bar{\mathcal{C}}_1 \times \cdots \times \bar{\mathcal{C}}_L$ where

$$\bar{\mathcal{C}}_k = \{\mathbf{Q}_k \in \mathcal{H}_+ : \mathrm{tr}(\mathbf{Q}_k) \leq \bar{p}_k\}, \qquad \text{for all } k.$$

Clearly, if $\boldsymbol{\eta}$ is sufficiently small, then

$$\min_{\mathbf{Q}_{-k}\in\bar{\mathcal{C}}_{-k}} \max_{\mathbf{Q}_k\in\bar{\mathcal{C}}_k} I(\mathbf{x}_k;\mathbf{y}_k)(\mathbf{Q}_k,\mathbf{Q}_{-k}) > \bar{r}_k, \qquad \text{for all } k.$$

■

The following proposition, whose proof follows from the definition of $\mathcal{R}_m(\mathbf{H})$ and Proposition 3, states that all sufficiently small rates are equilibrium rates.

Proposition 5: Fix $\mathbf{H}$. There exist $\bar{\mathbf{r}} \in \mathbb{R}^L_{++}$, such that

$$\left\{\mathbf{r} \in \mathbb{R}^L_+ : 0 \le \mathbf{r} \le \bar{\mathbf{r}} \text{ (elementwise)}\right\} \subset \mathcal{R}_e(\mathbf{H}).$$

It turns out that, for some $\mathbf{H}$, $\mathcal{R}_m(\mathbf{H})$ is a proper subset of $\mathcal{R}_e(\mathbf{H})$, i.e.,

$$\mathcal{R}_m(\mathbf{H}) \subsetneq \mathcal{R}_e(\mathbf{H}).$$

Example 3: Consider the setup in Example 2. Simple calculations yield, for all $\eta \ge 0$,

$$\mathcal{R}_m(\mathbf{H}) = \left\{\mathbf{r} \in \mathbb{R}^2_+ \ : \ \left(\sqrt{2^{r_1}} - 1\right)\left(\sqrt{2^{r_2}} - 1\right) < 1/\eta^2\right\}.$$

In view of Example 2, this means that, for $\eta < 1$, $\mathcal{R}_m(\mathbf{H}) = \mathcal{R}_e(\mathbf{H})$.

However, for $\eta \ge 1$, the capacity game $\Gamma^c(\mathbf{H}, \mathbf{p})$ for some fixed $\mathbf{p} \in \mathbb{R}^2_+$ has additional equilibria (in addition to $(p_1\mathbf{I}/2, p_2\mathbf{I}/2)$). In particular, when $\eta \ge 1$ and $p_1 = p_2 = p$ for some $p > 0$, $(\mathrm{diag}(p, 0), \mathrm{diag}(0, p))$ is such an additional equilibrium with the rate profile $(\log_2(1 + p), \log_2(1 + p))$. For a sufficiently large $p > 0$, $(\log_2(1 + p), \log_2(1 + p)) \notin \mathcal{R}_m(\mathbf{H})$. As a result, for $\eta \ge 1$, $\mathcal{R}_m(\mathbf{H}) \subsetneq \mathcal{R}_e(\mathbf{H})$.

A. The Case of Diagonal Matrices

Here, we specialize to the case where all channel as well as covariance matrices are square and diagonal. We will present an explicit underestimation of the set of equilibrium rates using our minmax approach. We note that the same problem is studied in [9] using a different approach.

Consider a power game $\Gamma^p(\mathbf{H},\mathbf{r})$ where all channel matrices as well as all covariance matrices have the form: $\mathbf{H}_{k,\ell} = \text{diag}\big(h_{k,\ell}^{1,1},\dots,h_{k,\ell}^{N,N}\big)$, $\mathbf{Q}_k = \text{diag}\big(q_k^{1,1},\dots,q_k^{N,N}\big)$ for all $k,\ell \in \{1,\dots,L\}$ where $N := N_r = N_t$. Hence, we write

$$I(\mathbf{x}_k;\mathbf{y}_k)(\mathbf{Q}_k,\mathbf{Q}_{-k}) = \sum_{i=1}^{N} \log_2 \left. \quad 1 + \frac{\big|h_{k,k}^{i,i}\big|^2 q_k^{i,i}}{1+\sum_{\ell\neq k}\big|h_{k,\ell}^{i,i}\big|^2 q_\ell^{i,i}} \right) .$$

Let $\mathcal{C} = \mathcal{C}_1 \times \cdots \times \mathcal{C}_L$, where, for some $(p_1,\dots,p_L) \in \mathbb{R}_+^L$,

$$\mathcal{C}_k = \Big\{ \mathbf{Q}_k \in \mathcal{H}_+ : \text{tr}(\mathbf{Q}_k) \le p_k,\ \mathbf{Q}_k = \text{diag}\big(q_k^{1,1},\dots,q_k^{N,N}\big) \Big\}, \qquad \text{for all } k.$$

We have

$$\begin{aligned} \min_{\mathbf{Q}_{-k}\in\mathcal{C}_{-k}} \max_{\mathbf{Q}_k\in\mathcal{C}_k} I(\mathbf{x}_k;\mathbf{y}_k)(\mathbf{Q}_k,\mathbf{Q}_{-k}) \ &\ge \ \min_{\mathbf{Q}_{-k}\in\mathcal{C}_{-k}} \sum_{i=1}^{N} \log_2 \left. \quad 1 + \frac{p_k/N}{\beta_{k,k} + \sum_{\ell\neq k}\beta_{k,\ell}\ q_\ell^{i,i}} \right) \\ &= \ N \log_2 \left. \quad 1 + \frac{p_k/N}{\beta_{k,k} + \sum_{\ell\neq k}\beta_{k,\ell}\ p_\ell/N} \right) \end{aligned}$$

where, for all k,ℓ,

$$\beta_{k,\ell} := \begin{cases} \max_{i\in\{1,\dots,N\}} 1/\big|h_{k,k}^{i,i}\big|^2, & k=\ell \\ \max_{i\in\{1,\dots,N\}} \big|h_{k,\ell}^{i,i}\big|^2/\big|h_{k,k}^{i,i}\big|^2, & k\neq\ell \end{cases} .$$

This means that, if $\mathbf{r} \in \mathbb{R}_+^L$ satisfies

$$r_k < N \log_2 \left. \quad 1 + \frac{p_k/N}{\beta_{k,k} + \sum_{\ell\neq k}\beta_{k,\ell}\ p_\ell/N} \right), \qquad \text{for all } k$$

then $BR^p(\mathcal{C}) \subset \mathcal{C}$. Now, by Theorem 1, the restriction of $\Gamma^p(\mathbf{H},\mathbf{r})$ to $\mathcal{C}$ possesses an equilibrium, which is also an equilibrium of $\Gamma^p(\mathbf{H},\mathbf{r})$. Therefore, $\mathbf{r} \in \mathcal{R}_e(\mathbf{H})$, if there exists a $\mathbf{p} \in \mathbb{R}_+^L$ such that

$$(\mathbf{I}-\mathbf{A})\mathbf{p} > \mathbf{b}, \qquad \text{(elementwise)} \tag{11}$$

where

$$\mathbf{A} := \begin{pmatrix} 0 & \gamma_1\beta_{1,2} & \dots & \gamma_1\beta_{1,L} \\ \gamma_2\beta_{2,1} & 0 & \dots & \gamma_1\beta_{1,L} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_L\beta_{L,1} & \gamma_L\beta_{L,2} & \dots & 0 \end{pmatrix}, \ \mathbf{b} := N \begin{pmatrix} \gamma_1\beta_{1,1} \\ \vdots \\ \gamma_L\beta_{L,L} \end{pmatrix}$$

and $\gamma_k := \sqrt[N]{2^{r_k}} - 1$, for all k. It turns out that if $\rho(\mathbf{A}) < 1$, where $\rho(\mathbf{A})$ is the spectral radius of $\mathbf{A}$, then $(\mathbf{I} - \mathbf{A})^{-1}$ exists and $(\mathbf{I} - \mathbf{A})^{-1} \geq 0$ elementwise; see condition (N_{38}) on page 137 of [16]. This leads to the following result.

Proposition 6: Fix $\mathbf{H}$. Assume that all channel matrices and all covariance matrices are diagonal. Then,

$$\{\mathbf{r} \in \mathbb{R}_+^L : \rho(\mathbf{A}) < 1\} \subset \mathcal{R}_e(\mathbf{H})$$

where $\mathbf{A}$ is as introduced above.

Finally, we note that the condition $\rho(\mathbf{A}) < 1$ is similar to, but less stringent than, the condition in Corollary 6 of [9].

VI. RELAXING THE QOS REQUIREMENTS

We thus far viewed the QoS requirements as hard constraints. One of the difficulties with this viewpoint is that if the rates that the links seek to achieve are not achievable at an equilibrium, then the links would not be able to settle at any solution. In a practical scenario, an individual link not knowing the entire setup would not be able to easily determine whether or not its QoS requirement is achievable at an equilibrium. The lack of an equilibrium would manifest itself as persistent oscillations when the links continually adjust their covariance matrices using an update algorithm such as the best response dynamics. To overcome this difficulty, we now relax the QoS requirements by removing the hard constraint $\mathbf{Q}_k \in \mathcal{F}_k^p(\mathbf{Q}_{-k})$ and modifying the link cost functions as

$$J_k^w(\mathbf{Q}) := \mathrm{tr}(\mathbf{Q}_k) + w\big(r_k - I(\mathbf{x}_k; \mathbf{y}_k)\big)^+, \qquad \mathbf{Q}_k \in \mathcal{H}_+ \tag{12}$$

where $w \in \mathbb{R}_+$ is the cost of violating the QoS requirements. We refer to a game characterized by the cost functions $J_1^w, \ldots, J_L^w$ and the feasible strategy sets $\mathcal{H}_+ \times \cdots \times \mathcal{H}_+$ as a “weighted power game” and denote it by $\Gamma^w(\mathbf{H}, \mathbf{r})$.

We denote the $k-$th link’s best response function by BR_k^w, i.e.,

$$BR_k^w(\mathbf{Q}_{-k}) = \mathrm{argmin}_{\mathbf{Q}_k \in \mathcal{H}_+} J_k^w(\mathbf{Q}_k, \mathbf{Q}_{-k})$$

and the composite best response function by $BR^w := (BR_1^w, \ldots, BR_L^w)$; see Proposition 10 in Appendix for the fact that, for any given $\mathbf{Q} \in \mathcal{H}_+^L$, $BR^w(\mathbf{Q})$ is nonempty and single-valued. We call a profile of link strategies $\check{\mathbf{Q}} \in \mathcal{H}_+^L$ an equilibrium of $\Gamma^w(\mathbf{H}, \mathbf{r})$ if

$$\check{\mathbf{Q}} = BR^w(\check{\mathbf{Q}}).$$

We are essentially interested in the case where $w \uparrow \infty$, since, for large w, the links are expected to strive towards achieving their QoS requirements. If it is not possible for a link to achieve its QoS requirement, then this would result in a very high power level for the particular link at an equilibrium, which would perhaps prompt the link to scale down its QoS requirement. One advantage of relaxing the QoS requirements is that an equilibrium would always exist regardless of the rates that the links seek to achieve. Another advantage is that if the target rates are equilibrium rates, then the equilibria of a weighted power game for sufficiently large w contain the equilibria of the corresponding power game which by definition satisfy the QoS requirements.

Proposition 7: Fix $\mathbf{H}$, $\mathbf{r} \in \mathbb{R}_+^L$, and $w \in \mathbb{R}_+$. Then, the following statements are true.

1) The weighted power game $\Gamma^w(\mathbf{H}, \mathbf{r})$ possesses an equilibrium.
2) An equilibrium of $\Gamma^w(\mathbf{H}, \mathbf{r})$ satisfying the QoS requirement (1) is also an equilibrium of $\Gamma^p(\mathbf{H}, \mathbf{r})$.
3) There exists a $\bar{w} \in \mathbb{R}_+$ such that if $w \in [\bar{w}, \infty)$ then any equilibrium of $\Gamma^p(\mathbf{H}, \mathbf{r})$ is also an equilibrium of $\Gamma^w(\mathbf{H}, \mathbf{r})$.
4) Let w_n be an increasing positive-valued and unbounded sequence of scalars. Let $\mathbf{Q}^n$ be a corresponding sequence in $\mathcal{H}_+^L$ such that, for all $n \geq 1$, $\mathbf{Q}^n$ is an equilibrium of $\Gamma^{w_n}(\mathbf{H}, \mathbf{r})$.
   a) If $\sup_{n \geq 1} \sum_{k=1}^L \text{tr}(\mathbf{Q}_k^n) < \infty$, then there exists an $\bar{n} \geq 1$ such that, for all $n \geq \bar{n}$, $\mathbf{Q}^n$ is an equilibrium of $\Gamma^p(\mathbf{H}, \mathbf{r})$.
   b) If $\mathbf{r} \notin \mathcal{R}_e(\mathbf{H})$, then $\sup_{n \geq 1} \sum_{k=1}^L \text{tr}(\mathbf{Q}_k^n) = \infty$.

Proof:

1) We observe that, for any $\mathbf{Q}_{-k} \in \mathcal{H}_+^{L-1}$,

$$\mathrm{tr}\big(BR_k^w(\mathbf{Q}_k)\big) \leq J_k^w\big(BR_k^w(\mathbf{Q}_k), \mathbf{Q}_{-k}\big) \leq J_k^w(0, \mathbf{Q}_{-k}) = wr_k.$$

Therefore, without loss of generality, we can restrict the link strategies to $\mathcal{F}^w := \mathcal{F}_1^w \times \cdots \times \mathcal{F}_L^w$ where, for all k,

$$\mathcal{F}_k^w := \{\mathbf{Q}_k \in \mathcal{H}_+ : \mathrm{tr}(\mathbf{Q}_k) \leq wr_k\}.$$

Clearly, $\mathcal{F}_k^w$ is nonempty, convex and compact. Also, J_k^w is continuous in $\mathcal{F}^w$, and $J_k^w(\cdot, \mathbf{Q}_{-k})$ is convex in $\mathcal{F}_k^w$ for each fixed $\mathbf{Q}_{-k} \in \mathcal{F}_{-k}^w := \times_{\ell \neq k} \mathcal{F}_\ell^w$. Now, the existence of an equilibrium follows from Theorem 1.

2) Obvious.

3) Consider any equilibrium $\mathbf{Q}^*$ of the power game $\Gamma^p(\mathbf{H}, \mathbf{r})$. Proposition 10 in Appendix shows that there exists a $\bar{w} \in \mathbb{R}_+$ such that, for all $w \in [\bar{w}, \infty)$,

$$BR^w(\mathbf{Q}^*) = BR^p(\mathbf{Q}^*) = \mathbf{Q}^*.$$

4)
   a) Since $\mathbf{Q}^n$ belongs to a compact subset of $\mathcal{H}_+^L$, there exists a $\bar{w} \geq 0$ such that, for any $w \geq \bar{w}$, $BR^w(\mathbf{Q}^n) = BR^p(\mathbf{Q}^n)$, for all $n \geq 1$; see Proposition 10 in Appendix . This implies that there exists an $\bar{n} \geq 1$, such that, for all $n \geq \bar{n}$, $\mathbf{Q}^n = BR^{w_n}(\mathbf{Q}^n) = BR^p(\mathbf{Q}^n)$.
   b) Suppose that $\sup_{n\geq 1} \sum_{k=1}^L \mathrm{tr}(\mathbf{Q}_k) < \infty$. Then, part a) implies that, for a sufficiently large n, $\mathbf{Q}^n$ is an equilibrium of $\Gamma^p(\mathbf{H}, \mathbf{r})$, which contradicts $\mathbf{r} \notin \mathcal{R}_e(\mathbf{H})$.

■

VII. THE BEST RESPONSE DYNAMICS

Here, we consider a situation in which the links are iteratively adjusting their covariance matrices to minimize their cost functions in a weighted power game. During iteration $t+1$, any individual link k knows nothing about the setup except it can compute its own best response $BR_k^w(\mathbf{Q}_{-k}(t))$ to the decisions $\mathbf{Q}_{-k}(t)$ made by the other links at the previous iteration t. For this, it is sufficient for an individual link k

to know its own channel matrix $\mathbf{H}_{k,k}$, its own QoS requirement r_k, and to measure the covariance matrix $\mathbf{R}_k$ of the total noise and interference corresponding to $\mathbf{Q}_{-k}(t)$. The actual computation of $BR_k^w(\mathbf{Q}_{-k}(t))$ can be done as shown in Proposition 10 in Appendix . Once $BR_k^w(\mathbf{Q}_{-k}(t))$ is obtained, any link k updates its own covariance matrix according to

$$\mathbf{Q}_k(t+1) = (1-\alpha(t))\mathbf{Q}_k(t) + \alpha(t) BR_k^w(\mathbf{Q}_{-k}(t)) \tag{13}$$

where $0 \leq \alpha(t) \leq 1$ is a parameter that represents the k-th link's willingness to optimize (in other words, $1-\alpha(t)$ is the k-th link's inertia) at step t. The inertia prevents the links from overreacting and generally helps with the convergence of the updates.

We also consider the case where the links are engaged in a power game and therefore update their covariance matrices according to

$$\mathbf{Q}_k(t+1) = (1-\alpha(t))\mathbf{Q}_k(t) + \alpha(t) BR_k^p(\mathbf{Q}_{-k}(t)) \tag{14}$$

where $0 \leq \alpha(t) \leq 1$ again represents the k-th link's willingness to optimize at step t; see Proposition 9 in Appendix for the actual computation of $BR_k^p(\mathbf{Q}_{-k}(t))$. We should point out that the covariance matrices generated during the iterations of (14) need not satisfy the QoS requirements of the links. More precisely, $\mathbf{Q}_k(t)$ need not belong to $\mathcal{F}_k^p(\mathbf{Q}_{-k}(t))$, even though $\mathbf{Q}_k(t) \in \mathcal{F}_k^p(\mathbf{Q}_{-k}(t-1))$ by construction. However, if the best response dynamics (14) converge to some limiting covariance matrices $\tilde{\mathbf{Q}}$, then $\tilde{\mathbf{Q}}$ must be an equilibrium and satisfy all QoS requirements.

The best response dynamics (13)-(14) can be generalized by allowing the links to update intermittently as long as they do not completely stop updating until convergence. In our numerical simulations, both best response dynamics (13)-(14) typically converge. Moreover, for sufficiently small interference, convergence can be proven along the lines of Proposition 4.1 in [6] by showing that the best response functions are contractions.

Proposition 8: Fix $(\mathbf{H}_{k,k})_{k\in\{1,\ldots,L\}}$, $\mathbf{r} \in \mathbb{R}_+^L$, and $w \in \mathbb{R}_+$. Assume

(i) $\text{rank}(\mathbf{H}_{k,k}) = N_t$, for all k,

(ii) $w > \max_{k\in\{1,\ldots,L\}} \left(2^{r_k}/\lambda_{\min}(\mathbf{H}_{k,k}^{\dagger}\mathbf{H}_{k,k})\right)$,

(iii) $\mathbf{Q}_k(0) \in \mathcal{F}_k^w$, where $\mathcal{F}_k^w = \{\mathbf{Q}_k \in \mathcal{H}_+ : \text{tr}(\mathbf{Q}_k) \leq w r_k\}$, for all k,

(iv) $\lim_{t\to\infty} \alpha(t) = 0$ and $\sum_{k=0}^{\infty} \alpha(t) = \infty$, for all k.

If the interference channels are sufficiently weak, i.e., $(\mathbf{H}_{k,\ell})_{k\neq\ell}$ is sufficiently small, then the best response dynamics (13) and (14) generate identical trajectories converging to a limiting point, say $\mathbf{Q}^* \in \mathcal{H}_+^L$. Moreover, $\mathbf{Q}^*$ is an equilibrium of $\Gamma^p(\mathbf{H}, \mathbf{r})$ (as well as $\Gamma^w(\mathbf{H}, \mathbf{r})$) that is unique within $\mathcal{F}^w$.

Proof:

Proposition 10 shows that $BR^w \equiv BR^p$ in $\mathcal{F}^w$ and both are contractions in $\mathcal{F}^w$. With this, the desired result follows from Theorem 6.9 in [17].

■

Remark 2: An example of $\alpha(t)$ satisfying the assumption (iv) in Proposition 8 is $\alpha(t) = 1/t$. Moreover, assumption (iv) in Proposition 8 can be relaxed, e.g., it can be relaxed to

$$\alpha(t) = \bar{\alpha}, \qquad \text{for all } t \geq 0,\ \text{ for some } \bar{\alpha} \in (0, 1]$$

due to the well-known contraction mapping theorem.

Remark 3: For the results of proposition 8 to hold, the interference channels should be weak enough so that the composite best response functions are contractions. Following along the steps of the proof, one can obtain a bound on how large the interference channels could be for the composite best response functions to be contractions. However, such a bound would be clearly dependent on the parameters of the model whose values are never accurately known in practice. Therefore, we do not see much practical value in such a bound, and as a result, we content ourselves with showing that, if the interference becomes sufficiently weak, then the best response dynamics will converge to a unique equilibrium.

We now provide some examples of nonconvergent cases when the conditions of Proposition 8 are not met.

Example 4: The best response dynamics (14) with $\alpha(t) = 1$ starting from 0 may diverge even if an equilibrium exists. Consider the setup $L = 2,\ N_r = N_t = 2,\ \mathbf{H}_{1,1} = \mathbf{H}_{2,2} = \mathbf{I},\ \mathbf{H}_{1,2} = \mathbf{H}_{2,1} = \sqrt{\eta}\mathbf{I}$, for some $\eta \geq 0$, and $r_1 = r_2 = 1$. We first compute that, for any $\rho \geq 0$, $BR_k^p(\rho\mathbf{I}) = \gamma(1 + \rho\eta)\mathbf{I}$, where $\gamma := \sqrt{2} - 1$. Therefore, starting from 0, the best response dynamics would generate

$$0 \to \gamma(\mathbf{I}, \mathbf{I}) \to \gamma(1 + \gamma\eta)(\mathbf{I}, \mathbf{I}) \to \gamma(1 + \gamma\eta + \gamma^2\eta^2)(\mathbf{I}, \mathbf{I}) \to \dots$$

which diverges when $\gamma\eta \geq 1$.

Example 5: The best response dynamics (14) may not converge even if the initial condition is feasible and arbitrarily close to an equilibrium. Consider the setup given in the previous example with $\eta = 1$, in which $(\text{diag}(1, 0), \text{diag}(0, 1))$ is an equilibrium. We first compute that, for any $\theta \geq 0$ and small $\epsilon > 0$,

$$\begin{aligned} BR_k^p(\text{diag}(\theta, \theta + 1 - \epsilon)) &= \text{diag}(\bar{\theta} + 1 - \epsilon, \bar{\theta}) \\ BR_k^p(\text{diag}(\theta + 1 - \epsilon, \theta)) &= \text{diag}(\bar{\theta}, \bar{\theta} + 1 - \epsilon) \end{aligned}$$

for some $\bar{\theta} > 0$. Now, consider the initial condition $\mathbf{Q}^0 = (\text{diag}(1, \epsilon), \text{diag}(0, 1 + \epsilon))$ for some small $\epsilon > 0$. Note that $\mathbf{Q}^0$ is feasible, i.e., $I(\mathbf{x}_k; \mathbf{y}_k)(\mathbf{Q}^0) \geq 1$. Starting from $\mathbf{Q}^0$, the best response dynamics would generate

$$\begin{aligned} \mathbf{Q}^1 &= (\text{diag}(1, 0), \text{diag}(\theta_1, \theta_1 + 1 - \epsilon)) \\ \mathbf{Q}^2 &= (\text{diag}(\theta_2 + 1 - \epsilon, \theta_2), \text{diag}(0, 1)) \\ \mathbf{Q}^3 &= (\text{diag}(1, 0), \text{diag}(\theta_3, \theta_3 + 1 - \epsilon)) \\ &\vdots \end{aligned}$$

for some $\theta_1 > 0,\ \theta_2 > 0,\ \theta_3 > 0, \dots$.

VIII. SIMULATION RESULTS

In this section, we numerically verify some of our theoretical results. We consider the following setup: $L = 3,\ N_r = N_t = 2$,

$$\mathbf{H}_{1,1} = \mathbf{H}_{2,2} = \mathbf{H}_{3,3} = 2\mathbf{I},\ \mathbf{H}_{1,2} = \mathbf{H}_{2,1} = \mathbf{H}_{3,1} = \frac{1}{2}\begin{pmatrix} 1 & i \\ 0 & 1 \end{pmatrix},\ \mathbf{H}_{1,3} = \mathbf{H}_{2,3} = \mathbf{H}_{3,2} = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ i & 1 \end{pmatrix}$$

with two different target rate profiles $\mathbf{r}^1 = (4, 3, 2)$ and $\mathbf{r}^2 = (7.2, 5.4, 3.6)$. We simulate the best response dynamics (14) and (13) with $\alpha(t) \equiv 1$, some random initial conditions, and different values of w. The rate and power trajectories generated by (14) with $\mathbf{r}^1$ is shown in Figure 2. When this simulation is repeated with (13) using the same initial conditions and $w \geq 5 \ln 2$, the results are identical to those in Figure 2. Clearly, the results in Figure 2 are convergent and $\mathbf{r}^1$ is achieved at equilibrium.

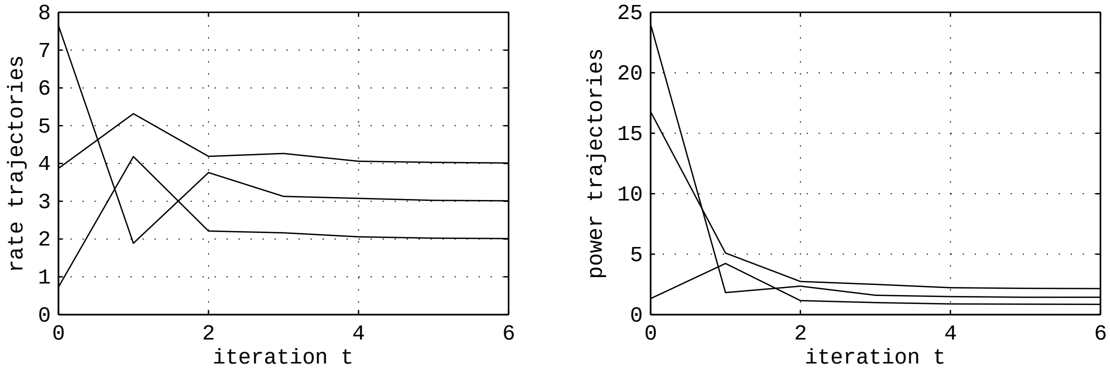

Fig. 2. The rate and power trajectories generated by (14) with $\mathbf{r}^1$.

Fig 3 shows the rate and power trajectories generated by (14) with $\mathbf{r}^2$ where the dashed lines represent $\mathbf{r}^2$. Our repeated simulations of this case with random initial conditions generate trajectories similar to those in Fig 3 where the rate trajectories settle at values that are slightly lower than $\mathbf{r}^2$ but the power trajectories diverge. However, target rate profiles that are slightly smaller than $\mathbf{r}^2$ (elementwise) are achievable at equilibrium with bounded power trajectories (the results are not shown here). This suggest that $\mathbf{r}^2$ is outside but not too far from $\mathcal{R}_e(\mathbf{H})$.

When we simulate (13) with $\mathbf{r}^2$ and increasing values of w, the results are always convergent in contrast to those in Fig 3. We only present the case where $w = 20 \ln 2$ in Figure 4 where the dashed lines again

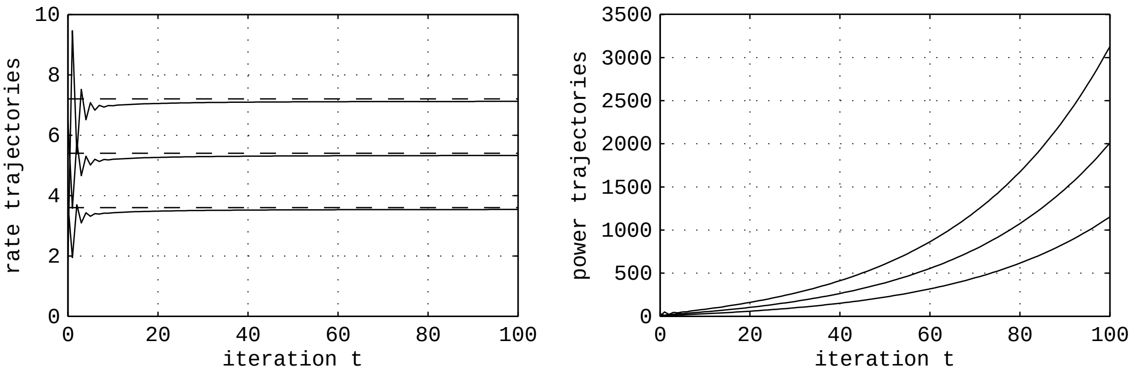

Fig. 3. The rate and power trajectories generated by (14) with $\mathbf{r}^2$.

represent $\mathbf{r}^2$. As w gets larger, the rate trajectories settle at values that are closer to but still below $\mathbf{r}^2$ whereas the power trajectories settle at higher values, which is consistent with our theoretical results.

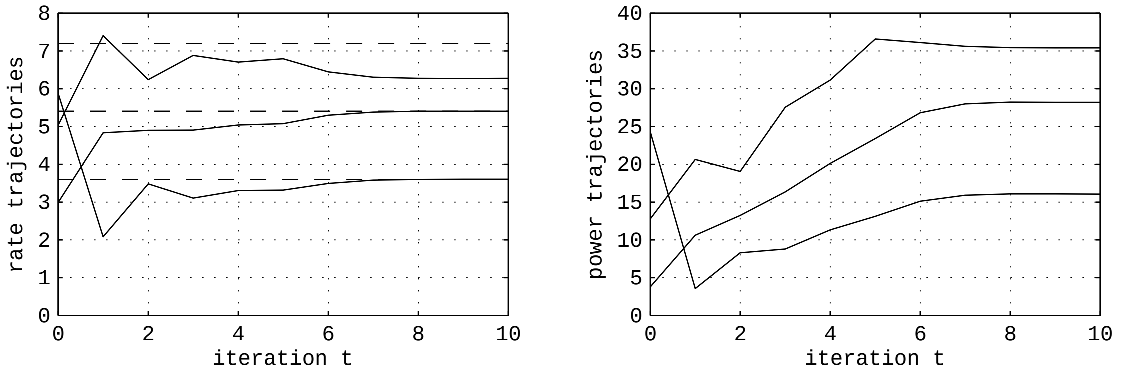

Fig. 4. The rate and power trajectories generated by (13) with $\mathbf{r}^2$.

IX. CONCLUSIONS

We studied a power control problem in MIMO interference systems within the framework of multi-link games. We established a relationship between power control and capacity control problems from a game theoretic perspective. We illustrated on an example that equilibrium may not exist for all achievable rates. Using a minmax approach, we obtained a conservative estimate of the rate region for which an equilibrium exists in a power game. We extended our results to the case where the QoS requirements are relaxed. We presented some sufficient conditions for the uniqueness of equilibrium as well as for the convergence of the best response process (iterative water-filling). Improving the efficiency of equilibrium using stream control as in [6] remains as a significant future work.

APPENDIX

Proposition 9:

1) Fix any $\mathbf{H}$ and $\mathbf{r} \in \mathbb{R}_+^L$.
    a) The $k-$th link's best response in $\Gamma^p(\mathbf{H},\mathbf{r})$ to any $\mathbf{Q}_{-k} \in \mathcal{H}_+^{L-1}$ is unique and given as

    $$BR_k^p(\mathbf{Q}_{-k}) = \mathbf{V}_k \mathrm{diag}\left(q_k^{1,1},\dots,q_k^{n_k,n_k},0,\dots,0\right)\mathbf{V}_k^\dagger = \sum_{i=1}^{n_k} q_k^{i,i}\mathbf{v}_k^i\left(\mathbf{v}_k^i\right)^\dagger$$

    where

    - $\sigma_k^1 \geq \dots \geq \sigma_k^{n_k} > \sigma_k^{n_k+1} = \dots = \sigma_k^{N_t} = 0$ are the eigenvalues of $\mathbf{H}_{k,k}^\dagger \mathbf{R}_k^{-1}\mathbf{H}_{k,k}$ [5],
    - $\mathbf{V}_k$ is a unitary matrix such that its columns $\mathbf{v}_k^1,\dots,\mathbf{v}_k^{N_t}$ are a set of orthonormal eigenvectors of $\mathbf{H}_{k,k}^\dagger \mathbf{R}_k^{-1}\mathbf{H}_{k,k}$ corresponding to $\sigma_k^1,\dots,\sigma_k^{N_t}$, respectively,
    - $q_k^{i,i} = (\mu_k/\ln 2 - 1/\sigma_k^i)^+$, for all $i \in \{1,\dots,n_k\}$,
    - μ_k is such that $r_k = \sum_{i=1}^{n_k}\log_2(1+q_k^{i,i}\sigma_k^i)$, that is,

    $$r_k = \sum_{i=1}^{n_k}(\log_2(\sigma_k^i\mu_k/\ln 2))^+. \tag{15}$$

    Note that this is essentially the well-known water-filling solution [1] except that the water level is chosen to satisfy the rate constraint instead of a given power constraint.

    b) For any $\mathbf{Q}_{-k} \in \mathcal{H}_+^{L-1}$,

    $$BR_k^p(\mathbf{Q}_{-k}) = \mathrm{argmin}_{\mathbf{Q}_k \in \mathcal{H}_+}\mathrm{tr}(\mathbf{Q}_k) + \mu_k(\mathbf{Q}_{-k})\big(r_k - I(\mathbf{x}_k;\mathbf{y}_k)(\mathbf{Q}_k,\mathbf{Q}_{-k})\big) \tag{16}$$

    where $\mu_k(\mathbf{Q}_{-k})$ is the unique solution of (15) for μ_k.

2) Fix any $(\mathbf{H}_{k,k})_{k\in\{1,\dots,L\}}$, $\mathbf{r} \in \mathbb{R}_+^L$, and convex and compact subset $\mathcal{C} = \mathcal{C}_1 \times \dots \times \mathcal{C}_L \subset \mathcal{H}_+^L$. $\Big($Recall the notation $\boldsymbol{\eta} = (\mathbf{H}_{j,\ell})_{j\neq\ell}$, $\|\boldsymbol{\eta}\| = \sqrt{\sum_{j\neq\ell}\|\mathbf{H}_{j,\ell}\|_F^2}$, and $\mathcal{C}_{-k} = \times_{\ell\neq k}\mathcal{C}_\ell\Big)$.

---

[5] Note that $\sigma_k^1 > 0$ due to Assumption 1. Also, since $\mathbf{R}_k^{-1}$ is nonsingular for any $\mathbf{Q}_{-k} \in \mathcal{H}_+^{L-1}$, the number $n_k \in \{1,\dots,\min\{N_r,N_t\}\}$ of non-zero eigenvalues of $\mathbf{H}_{k,k}^\dagger \mathbf{R}_k^{-1}\mathbf{H}_{k,k}$ is the same for any $\mathbf{Q}_{-k} \in \mathcal{H}_+^{L-1}$.

a) For sufficiently small $\boldsymbol{\eta}$, μ_k is Lipschitz in $\mathcal{C}_{-k}$ with modulus $O(\|\boldsymbol{\eta}\|^2)$.

b) For sufficiently small $\boldsymbol{\eta}$, BR^p is Lipschitz in $\mathcal{C}$ with modulus $O(\|\boldsymbol{\eta}\|^2)$, provided

  (i) $\text{rank}(\mathbf{H}_{k,k}) = N_t$, for all k, and

  (ii) $\mathcal{C} \subset BR^p(\mathcal{C})$.

Proof:

1) a) Recall that $BR^p_k(\mathbf{Q}_{-k})$ is the set of solutions of the problem

$$\min_{\mathbf{Q}_k \in \mathcal{H}_+} \text{tr}(\mathbf{Q}_k) \qquad \text{subject to} \qquad g_k(\mathbf{Q}_k, \mathbf{Q}_{-k}) \in \mathbb{R}_- \tag{17}$$

where $g_k(\mathbf{Q}_k, \mathbf{Q}_{-k}) := r_k - I(\mathbf{x}_k; \mathbf{y}_k)(\mathbf{Q}_k, \mathbf{Q}_{-k})$. Let the Lagrangian of the problem (17) be denoted by

$$L_k(\mathbf{Q}_k, \mu_k, \mathbf{Q}_{-k}) := \text{tr}(\mathbf{Q}_k) + \mu_k g_k(\mathbf{Q}_k, \mathbf{Q}_{-k}), \qquad \mu_k \in \mathbb{R}.$$

Note that the problem (17) is convex (in the sense of Definition 2.163 in [18]) and has a finite optimal value. Moreover, the regularity condition $0 \in \text{int}\big(r_k - I(\mathbf{x}_k; \mathbf{y}_k)(\mathcal{H}_+, \mathbf{Q}_{-k}) - \mathbb{R}_-\big)$ is satisfied; see (3.12) in [18]). Hence, Theorem 3.4 in [18]) implies that a feasible point $\bar{\mathbf{Q}}_k$ is a solution of the problem (17), i.e., $\bar{\mathbf{Q}}_k \in BR^p_k(\mathbf{Q}_{-k})$, if and only if there exists a Lagrange multiplier $\bar{\mu}_k$ satisfying

$$\bar{\mathbf{Q}}_k \in \text{argmin}_{\mathbf{Q}_k \in \mathcal{H}_+} L_k(\mathbf{Q}_k, \bar{\mu}_k, \mathbf{Q}_{-k}) \quad \text{and} \quad g_k(\bar{\mathbf{Q}}_k, \mathbf{Q}_{-k}) \in \mathbb{R}_-, \ \bar{\mu}_k \in \mathbb{R}_+, \ \bar{\mu}_k g_k(\bar{\mathbf{Q}}_k, \mathbf{Q}_{-k}) = 0. \tag{18}$$

We first solve

$$\min_{\mathbf{Q}_k \in \mathcal{H}_+} L_k(\mathbf{Q}_k, \mu_k, \mathbf{Q}_{-k}), \qquad \text{for any fixed } \mu_k \in \mathbb{R}_+. \tag{19}$$

Let $\boldsymbol{\Sigma} := \text{diag}(\sigma_k^1, \ldots, \sigma_k^{n_k}, 0, \ldots, 0)$. Using the determinant identity $|\mathbf{I} + \mathbf{AB}| = |\mathbf{I} + \mathbf{BA}|$, we write

$$L_k(\mathbf{Q}_k, \mu_k, \mathbf{Q}_{-k}) = \text{tr}(\mathbf{V}_k^\dagger \mathbf{Q}_k \mathbf{V}_k) + \mu_k \left(r_k - \log_2 \left| \mathbf{I} + \boldsymbol{\Sigma}_k^{1/2} \mathbf{V}_k^\dagger \mathbf{Q}_k \mathbf{V}_k \boldsymbol{\Sigma}_k^{1/2} \right| \right).$$

We can equivalently perform the minimization over $\tilde{\mathbf{Q}}_k := \mathbf{V}_k^\dagger \mathbf{Q}_k \mathbf{V}_k$. Hence, we focus on

$$\min_{\tilde{\mathbf{Q}}_k \in \mathcal{H}_+} \text{tr}(\tilde{\mathbf{Q}}_k) + \mu_k \big(r_k - \log_2 \left| \mathbf{I} + \boldsymbol{\Sigma}_k^{1/2} \tilde{\mathbf{Q}}_k \boldsymbol{\Sigma}_k^{1/2} \right| \big). \tag{20}$$

Let $\tilde{q}_k^{1,1}, \ldots, \tilde{q}_k^{N_t,N_t}$ be the diagonal entries of $\tilde{\mathbf{Q}}_k$. By Hadamard's inequality for Hermitian positive definite matrices, $\left|\mathbf{I} + \mathbf{\Sigma}_k^{1/2} \tilde{\mathbf{Q}}_k \mathbf{\Sigma}_k^{1/2}\right| \le \prod_{i=1}^{n_k} (1 + \tilde{q}_k^{i,i} \sigma_k^i)$ with equality if and only if $\mathbf{I} + \mathbf{\Sigma}_k^{1/2} \tilde{\mathbf{Q}}_k \mathbf{\Sigma}_k^{1/2}$ is diagonal. This implies that the minimum in (20) can be achieved only by diagonal matrices in $\mathcal{H}_+$ of the form $\mathrm{diag}\big(\tilde{q}_k^{1,1}, \ldots, \tilde{q}_k^{n_k,n_k}, 0, \ldots, 0\big)$. Accordingly, we refocus on

$$\min_{\tilde{q}_k^{1,1} \ge 0, \ldots, \tilde{q}_k^{n_k,n_k} \ge 0} \sum_{i=1}^{n_k} \tilde{q}_k^{i,i} + \mu_k \left. \quad r_k - \sum_{i=1}^{n_k} \log_2(1 + \tilde{q}_k^{i,i} \sigma_k^i) \right) . \tag{21}$$

The unique solution of (21) is obtained as

$$\check{q}_k^{i,i} = \big(\mu_k / \ln 2 - 1/\sigma_k^i\big)^+, \qquad i \in \{1, \ldots, n_k\}.$$

Therefore, the unique[6] solution of (19) is given as

$$\mathbf{V}_k \mathrm{diag}\big(\check{q}_k^{1,1}, \ldots, \check{q}_k^{n_k,n_k}, 0, \ldots, 0\big) \mathbf{V}_k^\dagger. \tag{22}$$

Finally, in view of the solution (22), the unique Lagrange multiplier $\bar{\mu}_k$ satisfying (18) is such that $r_k = \sum_{i=1}^{n_k} \big(\log_2(\sigma_k^i \bar{\mu}_k / \ln 2)\big)^+$.

b) Follows from part a).

2) a)

Step 1:

By a temporary abuse of notation, we let $\mu_k(\boldsymbol{\sigma}_k)$ denote the unique solution of (15) for a given $\boldsymbol{\sigma}_k = (\sigma_k^1, \ldots, \sigma_k^{n_k}) \in \mathbb{R}_{++}^{n_k}$. Fix any $\boldsymbol{\sigma}_k \in \mathbb{R}_{++}^{n_k}$ and $\mathbf{h}_k \in \mathbb{R}^{n_k}$. Find a permutation π of $\{1, \ldots, n_k\}$ such that, for all small $\alpha \ge 0$, $\sigma_k^{\pi(1)} + \alpha h_k^{\pi(1)} \ge \cdots \ge \sigma_k^{\pi(n_k)} + \alpha h_k^{\pi(n_k)} > 0$. It turns out that, for all small $\alpha \ge 0$, we have

$$\mu_k(\boldsymbol{\sigma}_k + \alpha \mathbf{h}_k) = \ln 2 \left. \quad \frac{2^{r_k}}{\Pi_{i=1}^{m_k(0^+)} \big(\sigma_k^{\pi(i)} + \alpha h_k^{\pi(i)}\big)} \right)^{1/m_k(0^+)} \tag{23}$$

where $m_k(0^+) = \lim_{\alpha \downarrow 0} m_k(\alpha)$ and

$$m_k(\alpha) := \max \left\{ 1 \le \ell \le n_k : \left. \quad \frac{2^{r_k}}{\Pi_{i=1}^{\ell} \big(\sigma_k^{\pi(i)} + \alpha h_k^{\pi(i)}\big)} \right)^{1/\ell} \big(\sigma_k^{\pi(\ell)} + \alpha h_k^{\pi(\ell)}\big) > 1 \right\}.$$

---

[6] Even though the choice of the eigenvectors $\mathbf{v}_k^1, \ldots, \mathbf{v}_k^{n_k}$ is not unique, for an $m-$repeated eigenvalue, say $\sigma_k^1 = \cdots = \sigma_k^m > 0$, the matrix $\sum_{i=1}^m \mathbf{v}_k^i \big(\mathbf{v}_k^i\big)^\dagger$ is unique, i.e., independent of the choice of the eigenvectors $\mathbf{v}_k^1, \ldots, \mathbf{v}_k^m$.

To see this, let $f_k^\ell(\boldsymbol{\sigma}_k+\alpha\mathbf{h}_k) := \left(\frac{2^{r_k}}{\Pi_{i=1}^{\ell}\left(\sigma_k^{\pi(i)}+\alpha h_k^{\pi(i)}\right)}\right)^{1/\ell}\left(\sigma_k^{\pi(\ell)}+\alpha h_k^{\pi(\ell)}\right)$, for $\ell \in \{1,\dots,n_k\}$. We note that, for all small $\alpha \geq 0$, $f_k^{\ell+1}(\boldsymbol{\sigma}_k+\alpha\mathbf{h}_k) \geq 1$ implies $f_k^{\ell}(\boldsymbol{\sigma}_k+\alpha\mathbf{h}_k) \geq f_k^{\ell+1}(\boldsymbol{\sigma}_k+\alpha\mathbf{h}_k)$. Hence, there are two possibilities. First, suppose

$$f_k^1(\boldsymbol{\sigma}_k) \geq \cdots \geq f_k^{m_k(0)}(\boldsymbol{\sigma}_k) > 1 > \max_{\ell\in\{m_k(0)+1,\dots,n_k\}} f_k^\ell(\boldsymbol{\sigma}_k).$$

Then, the continuity of f_k^ℓ implies

$$f_k^1(\boldsymbol{\sigma}_k+\alpha\mathbf{h}_k) \geq \cdots \geq f_k^{m_k(0)}(\boldsymbol{\sigma}_k+\alpha\mathbf{h}_k) > 1 > \max_{\ell\in\{m_k(0)+1,\dots,n_k\}} f_k^\ell(\boldsymbol{\sigma}_k+\alpha\mathbf{h}_k)$$

for all small $\alpha \geq 0$. This means that $m_k(0^+)$ exists and equals $m_k(0)$. Second, suppose, for some $\bar{m}_k(0) \in \{m_k(0)+1,\dots,n_k\}$,

$$f_k^1(\boldsymbol{\sigma}_k) \geq \cdots \geq f_k^{m_k(0)}(\boldsymbol{\sigma}_k) > 1 = f_k^{m_k(0)+1}(\boldsymbol{\sigma}_k) = \cdots = f_k^{\bar{m}_k(0)}(\boldsymbol{\sigma}_k) > \max_{\ell\in\{\bar{m}_k(0)+1,\dots,n_k\}} f_k^\ell(\boldsymbol{\sigma}_k).$$

Since $f_k^l(\boldsymbol{\sigma}_k+\alpha\mathbf{h}_k)$ is real analytic[7] in α near 0, we have, for some $\tilde{m}_k(0) \in \{m_k(0),\dots,\bar{m}_k(0)\}$,

$$f_k^1(\boldsymbol{\sigma}_k+\alpha\mathbf{h}_k) \geq \cdots \geq f_k^{\tilde{m}_k(0)}(\boldsymbol{\sigma}_k+\alpha\mathbf{h}_k) > 1 \geq \max_{\ell\in\{\tilde{m}_k(0)+1,\dots,n_k\}} f_k^\ell(\boldsymbol{\sigma}_k+\alpha\mathbf{h}_k)$$

for all small $\alpha > 0$.[8] This means that $m_k(0^+)$ exists and equals $\tilde{m}_k(0)$. One can now directly verify that the right-hand-side of (23) satisfies (15) for all small $\alpha \geq 0$. This allows us to compute the directional derivative of μ_k as

$$\lim_{\alpha\downarrow 0}\frac{\mu_k(\boldsymbol{\sigma}_k+\alpha\mathbf{h}_k)-\mu_k(\boldsymbol{\sigma}_k)}{\alpha} = -(\ln 2)\mu_k(\boldsymbol{\sigma}_k)\frac{1}{m_k(0^+)}\sum_{i=1}^{m_k(0^+)}\frac{h_k^{\pi(i)}}{\sigma_k^{\pi(i)}}.$$

Since $|\mu_k(\boldsymbol{\sigma}_k)| \leq 2^{r_k}/\min_{i\in\{1,\dots,n_k\}}\sigma_k^i$, the directional derivative above is upper bounded in magnitude by $2^{r_k}|\mathbf{h}_k|/\left(\min_{i\in\{1,\dots,n_k\}}\sigma_k^i\right)^2$. By 9.13 Theorem in [21], μ_k is continuous at any $\boldsymbol{\sigma}_k \in \mathbb{R}_{++}^{n_k}$[9]. Using Corollary 3.7 in [22], we now conclude that, for any fixed $\sigma_k^{\min} > 0$, μ_k is Lipschitz in $\{\boldsymbol{\sigma}_k \in \mathbb{R}_{++}^{n_k} : \min_{i\in\{1,\dots,n_k\}}\sigma_k^i > \sigma_k^{\min}\}$ with modulus $2^{r_k}/\left(\sigma_k^{\min}\right)^2$.

Step 2:

---

[7] See Exercise 13 on page 96 of [19] for a sufficient condition for real analyticity.

[8] The zeros of a nonconstant real analytic function must be isolated; see the remark on page 103 of [20].

[9] Note that the subderivative in condition (f) of 9.13 Theorem in [21] is upper bounded by the directional derivative computed above.

For any $\mathbf{Q}_{-k}$, $\bar{\mathbf{Q}}_{-k} \in \mathcal{C}_{-k}$, let $\mathbf{R}_k$, $\bar{\mathbf{R}}_k$ denote the corresponding covariance matrices for total noise and interference, respectively. Also, let $\boldsymbol{\sigma}_k$, $\bar{\boldsymbol{\sigma}}_k \in \mathbb{R}^{n_k}_{++}$ denote the vectors of non-zero eigenvalues of $\mathbf{H}^\dagger_{k,k}\mathbf{R}_k^{-1}\mathbf{H}_{k,k}$, $\mathbf{H}^\dagger_{k,k}\bar{\mathbf{R}}_k^{-1}\mathbf{H}_{k,k}$, respectively, both arranged in decreasing order. Note that

$$\|\mathbf{R}_k^{-1}\|_F \le \sqrt{N_r} \quad \text{and} \quad \|\mathbf{R}_k - \bar{\mathbf{R}}_k\|_F \le \sum_{\ell \ne k} \|\mathbf{H}_{k,\ell}\|_F^2 \|\mathbf{Q}_\ell - \bar{\mathbf{Q}}_\ell\|_F. \tag{24}$$

We have, for sufficiently small $\boldsymbol{\eta}$,

$$\begin{aligned} |\boldsymbol{\sigma}_k - \bar{\boldsymbol{\sigma}}_k| &\le \|\mathbf{H}_{k,k}\|_F^2 \|\mathbf{R}_k^{-1} - \bar{\mathbf{R}}_k^{-1}\|_F \\ &\le \|\mathbf{H}_{k,k}\|_F^2 \frac{\|\mathbf{R}_k^{-1}\|_F^2 \|\mathbf{R}_k - \bar{\mathbf{R}}_k\|_F}{1 - \|\mathbf{R}_k^{-1}\|_F \|\mathbf{R}_k - \bar{\mathbf{R}}_k\|_F} \\ &\le O(\|\boldsymbol{\eta}\|^2) \sqrt{\sum_{\ell \ne k} \|\mathbf{Q}_\ell - \bar{\mathbf{Q}}_\ell\|_F^2} \end{aligned}$$

where the first inequality follows from the Hoffman-Wielandt inequality for Hermitian matrices (see 6.3.8 Corollary in [23]) and submultiplicative property of $\|.\|_F$, the second inequality follows from a well known bound on inverses of perturbed matrices (see Section 5.8 in [23]), and the last inequality follows from (24).

Step 3:

There exists a $\sigma_k^{\min} > 0$ such that, for all sufficiently small $\boldsymbol{\eta}$ and for all $\mathbf{Q}_{-k} \in \mathcal{C}_{-k}$, the ordered non-zero eigenvalues of the corresponding $\mathbf{H}^\dagger_{k,k}\mathbf{R}_k^{-1}\mathbf{H}_{k,k}$ satisfy $\sigma_k^1 \ge \cdots \ge \sigma_k^{n_k} > \sigma_k^{\min}$. With this, from Steps 1 and 2 above, we finally conclude that μ_k is a Lipschitz function of $\mathbf{Q}_{-k}$ in $\mathcal{C}_{-k}$ with modulus $O(\|\boldsymbol{\eta}\|^2)$.

b) For any given $\mathbf{Q}_{-k} \in \mathcal{H}_+^{L-1}$, $BR_k^p(\mathbf{Q}_{-k})$ is the unique minimizer of $L_k(\cdot, \mu_k(\mathbf{Q}_{-k}), \mathbf{Q}_{-k})$ in $\mathcal{H}_+$. Also, $L_k(\cdot, \mu_k(\mathbf{Q}_{-k}), \mathbf{Q}_{-k})$ is twice differentiable in $\mathcal{H}_+$ and its second differential for any $\mathbf{Z}_k \in \mathcal{H}$ equals

$$\left.\frac{\partial^2}{\partial \alpha^2} L_k(\mathbf{Q}_k + \alpha \mathbf{Z}_k, \mu_k(\mathbf{Q}_{-k}), \mathbf{Q}_{-k})\right|_{\alpha=0} = \frac{\mu_k(\mathbf{Q}_{-k})}{\ln 2} \mathrm{tr}\left(\mathbf{H}^\dagger_{k,k}(\mathbf{R}_k + \mathbf{H}_{k,k}\mathbf{Q}_k\mathbf{H}^\dagger_{k,k})^{-1}\mathbf{H}_{k,k}\mathbf{Z}_k\right)^2.$$

Hence, there exists a $c > 0$ such that, for all small $\boldsymbol{\eta}$ and $(\mathbf{Q}_k, \mathbf{Q}_{-k}) \in \mathcal{C}$,

$$\left.\frac{\partial^2}{\partial \alpha^2} L_k(\mathbf{Q}_k + \alpha \mathbf{Z}_k, \mu_k(\mathbf{Q}_{-k}), \mathbf{Q}_{-k})\right|_{\alpha=0} \ge 2c\|\mathbf{Z}\|_F^2.$$

This implies the following second-order growth condition: for all small $\boldsymbol{\eta}$ and $(\mathbf{Q}_k, \mathbf{Q}_{-k}) \in \mathcal{C}$,

$$L_k(\mathbf{Q}_k, \mu_k(\mathbf{Q}_{-k}), \mathbf{Q}_{-k}) - L_k(BR_k^p(\mathbf{Q}_{-k}), \mu_k(\mathbf{Q}_{-k}), \mathbf{Q}_{-k}) \geq c\|\mathbf{Q}_k - BR_k^p(\mathbf{Q}_{-k})\|_F^2.$$

Theorem 4.32 in [18] implies, for any small $\boldsymbol{\eta}$ and $\mathbf{Q}_{-k}, \bar{\mathbf{Q}}_{-k} \in \times_{\ell \neq k} \mathcal{C}_\ell$,

$$\|BR_k^p(\mathbf{Q}_{-k}) - BR_k^p(\bar{\mathbf{Q}}_{-k})\|_F \leq c^{-1}\tilde{k}$$

where $\tilde{k}$ is the Lipschitz modulus of $L_k(\cdot, \mu_k(\mathbf{Q}_{-k}), \mathbf{Q}_{-k}) - L_k(\cdot, \mu_k(\bar{\mathbf{Q}}_{-k}), \bar{\mathbf{Q}}_{-k})$ in $\mathcal{C}_k$. We bound $\tilde{k}$ as

$$\begin{aligned}
\tilde{k} &\leq \max_{\mathbf{Q}_k \in \mathcal{C}_k} \|\mu_k(\mathbf{Q}_{-k})\nabla_{\mathbf{Q}_k} I(\mathbf{x}_k; \mathbf{y}_k)(\mathbf{Q}_k, \mathbf{Q}_{-k}) - \mu_k(\bar{\mathbf{Q}}_{-k})\nabla_{\mathbf{Q}_k} I(\mathbf{x}_k; \mathbf{y}_k)(\mathbf{Q}_k, \bar{\mathbf{Q}}_{-k})\|_F \\
&\leq \max_{\mathbf{Q}_k \in \mathcal{C}_k} \big\{|\mu_k(\mathbf{Q}_{-k})|.\|\nabla_{\mathbf{Q}_k} I(\mathbf{x}_k; \mathbf{y}_k)(\mathbf{Q}_k, \mathbf{Q}_{-k}) - \nabla_{\mathbf{Q}_k} I(\mathbf{x}_k; \mathbf{y}_k)(\mathbf{Q}_k, \bar{\mathbf{Q}}_{-k})\|_F \\
&\qquad + \|\nabla_{\mathbf{Q}_k} I(\mathbf{x}_k; \mathbf{y}_k)(\mathbf{Q}_k, \mathbf{Q}_{-k})\|_F.|\mu_k(\mathbf{Q}_{-k}) - \mu_k(\bar{\mathbf{Q}}_{-k})|\big\}
\end{aligned}$$

where $\nabla_{\mathbf{Q}_k} I(\mathbf{x}_k; \mathbf{y}_k)(\mathbf{Q}_k, \mathbf{Q}_{-k}) = \mathbf{H}_{k,k}^\dagger\big(\mathbf{R}_k + \mathbf{H}_{k,k}\mathbf{Q}_k\mathbf{H}_{k,k}^\dagger\big)^{-1}\mathbf{H}_{k,k}$. We recall that, there exists a $\sigma_k^{\min} > 0$ such that, for all sufficiently small $\boldsymbol{\eta}$ and for all $\mathbf{Q}_{-k} \in \mathcal{C}_{-k}$, the ordered non-zero eigenvalues of the corresponding $\mathbf{H}_{k,k}^\dagger\mathbf{R}_k^{-1}\mathbf{H}_{k,k}$ satisfy $\sigma_k^1 \geq \cdots \geq \sigma_k^{n_k} > \sigma_k^{\min}$, and hence $|\mu_k(\mathbf{Q}_{-k})| \leq 2^{r_k}/\sigma_k^{\min}$. Also, for all small $\boldsymbol{\eta}$ and $(\mathbf{Q}_k, \mathbf{Q}_{-k}) \in \mathcal{C}$,

$$\begin{aligned}
\|\nabla_{\mathbf{Q}_k} I(\mathbf{x}_k; \mathbf{y}_k)(\mathbf{Q}_k, \mathbf{Q}_{-k})\|_F &\leq \|\mathbf{H}_{k,k}\|_F^2\sqrt{N_r} \\
\|\nabla_{\mathbf{Q}_k} I(\mathbf{x}_k; \mathbf{y}_k)(\mathbf{Q}_k, \mathbf{Q}_{-k}) - \nabla_{\mathbf{Q}_k} I(\mathbf{x}_k; \mathbf{y}_k)(\mathbf{Q}_k, \bar{\mathbf{Q}}_{-k})\|_F &\leq O(\|\boldsymbol{\eta}\|^2)\sqrt{\sum_{\ell \neq k} \|\mathbf{Q}_\ell - \bar{\mathbf{Q}}_\ell\|_F^2}.
\end{aligned}$$

Hence, $\tilde{k} \leq O(\|\boldsymbol{\eta}\|^2)\sqrt{\sum_{\ell \neq k} \|\mathbf{Q}_\ell - \bar{\mathbf{Q}}_\ell\|_F^2}$, which implies the desired result.

■

Proposition 10: Fix $\mathbf{H}$, $\mathbf{r} \in \mathbb{R}_+^L$, and $w \in \mathbb{R}_+$.

1) The $k-$th link's best response in $\Gamma^w(\mathbf{H}, \mathbf{r})$ to any fixed $\mathbf{Q}_{-k} \in \mathcal{H}_+^{L-1}$ is unique and given as

$$BR_k^w(\mathbf{Q}_{-k}) = \mathbf{V}_k \mathrm{diag}\left(q_k^{1,1}, \ldots, q_k^{n_k,n_k}, 0, \ldots, 0\right)\mathbf{V}_k^\dagger = \sum_{i=1}^{n_k} q_k^{i,i}\mathbf{v}_k^i\left(\mathbf{v}_k^i\right)^\dagger$$

where

$$q_k^{i,i} = \left(\min\{w, \mu_k\}/\ln 2 - 1/\sigma_k^i\right)^+, \qquad i \in \{1, \ldots, n_k\}$$

and μ_k, $\sigma_k^1, \ldots, \sigma_k^{N_t}$, $\mathbf{V}_k = (\mathbf{v}_k^1, \ldots, \mathbf{v}_k^{N_t})$ are as in Proposition 9. Note that $BR_k^w(\mathbf{Q}_{-k}) = BR_k^p(\mathbf{Q}_{-k})$ whenever $w \geq \mu_k$.

2) For sufficiently small $\boldsymbol{\eta}$, $BR^w \equiv BR^p$ in $\mathcal{F}^w$ and both are contractions in $\mathcal{F}^w$, provided

   (i) $\text{rank}(\mathbf{H}_{k,k}) = N_t$, for all k, and

   (ii) $w > \max_{k\in\{1,\ldots,L\}} \left(2^{r_k}/\lambda_{\min}(\mathbf{H}_{k,k}^{\dagger}\mathbf{H}_{k,k})\right)$.

Proof:

1) Ignoring $(.)^+$ in the cost function leads to a problem of the form (19) with the unique minimizer

$$\mathbf{V}_k\text{diag}\left(\check{q}_k^{1,1}, \ldots, \check{q}_k^{n_k,n_k}, 0, \ldots, 0\right)\mathbf{V}_k^{\dagger} \tag{25}$$

where $\check{q}_k^{i,i} = (w/\ln 2 - 1/\sigma_k^i)^+$, $i \in \{1, \ldots, n_k\}$. Since ignoring $(.)^+$ lower bounds the cost, if $r_k \geq \sum_{i=1}^{n_k} \log_2(1+\check{q}_k^{i,i}\sigma_k^i)$, which is equivalent to $w \leq \mu_k$, then $BR_k^w(\mathbf{Q}_{-k})$ is given by (25). On the other hand, if $r_k < \sum_{i=1}^{n_k} \log_2(1+\check{q}_k^{i,i}\sigma_k^i)$, which is equivalent to $w \geq \mu_k$, then $BR_k^w(\mathbf{Q}_{-k}) = BR_k^p(\mathbf{Q}_{-k})$.

2) We first show that, for sufficiently small $\boldsymbol{\eta}$, $BR^w \equiv BR^p$ in $\mathcal{F}^w$, under the hypothesis. To see this, we need to show $w \geq \mu_k(\mathbf{Q}_{-k})$ for all k and $\mathbf{Q}_{-k} \in \mathcal{F}_{-k}^w$. This is implied by

$$\mu_k(\mathbf{Q}_{-k}) \leq \frac{2^{r_k}}{\lambda_{\min}(\mathbf{H}_{k,k}^{\dagger}\mathbf{H}_{k,k})}\left(1 + w\|\boldsymbol{\eta}\|^2 \max_{\ell\neq k} r_\ell\right).$$

Recall that $\mathcal{F}^w \subset BR^w(\mathcal{F}^w) = BR^p(\mathcal{F}^w)$. Proposition 9 implies that, for sufficiently small $\boldsymbol{\eta}$, BR^p is a contraction in $\mathcal{F}^w$.

■

REFERENCES

[1] E. Telatar, "Capacity of multi-antenna gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, 1999.

[2] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, pp. 311–335, 1998.

[3] M. F. Demirkol and M. A. Ingram, "Power-controlled capacity for interfering MIMO links," in *Proceedings of the IEEE Vehicular Technology Conference*, vol. 1, Oct. 2001, pp. 187–191.

[4] ——, “Stream control in networks with interfering MIMO links,” in *Proceedings of the IEEE Wireless Communications and Networking Conference*, vol. 1, Mar. 2003, pp. 343–348.

[5] S. Ye and R. S. Blum, “Optimized signalling for MIMO interference systems with feedback,” *IEEE Transactions on Signal Processing*, vol. 51, no. 11, pp. 2839–2848, 2003.

[6] G. Arslan, M. F. Demirkol, and Y. Song, “Equilibrium efficiency improvement in MIMO interference systems: a decentralized stream control approach,” *IEEE Transactions on Wireless Communications*, vol. 6, no. 8, pp. 2984–2993, 2007.

[7] G. Scutari, D. P. Palomar, and S. Barbarossa, “Competitive design of multiuser mimo systems based on game theory: a unified view,” *Journal of Selected Areas in Communications*, vol. 26, no. 7, pp. 1089–1103, Sept. 2008.

[8] M. F. Demirkol and M. A. Ingram, “Control using capacity constraints for interfering MIMO links,” in *Proceedings of the International Symposium on Personal, Indoor and Mobile Radio Communications*, vol. 3, Sept. 2001, pp. 1032–1036.

[9] J.-S. Pang, G. Scutari, F. Facchinei, and C. Wang, “Distributed power allocation with rate constraints in Gaussian parallel interference channels,” *IEEE Transaction on Information Theory*, vol. 54, no. 8, pp. 3471–3489, Aug. 2008.

[10] F. R. Farrokhi, K. J. R. Liu, and L. Tassiulas, “Transmit beamforming and power control for cellular wireless systems,” *Journal of Selected Areas in Communications*, vol. 16, no. 8, pp. 1437–1450, Oct. 1998.

[11] G. J. Foschini and Z. Miljanic, “A simple distributed autonomous power control algorithm and its convergence,” *IEEE Transactions on Vehicular Technology*, vol. 42, no. 4, pp. 641–646, Nov. 1993.

[12] E. Altman and Z. Altman, “S-modular games and power control in wireless networks,” *IEEE Transactions on Automatic Control*, vol. 48, no. 5, pp. 839–842, 2003.

[13] T. Ichiishi, *Game Theory for Economic Analysis*. New York, NY: Academic Press, 1983.

[14] W. W. Hogan, “Point-to-set mappings in mathematical programming,” *SIAM Review*, vol. 15, no. 13, pp. 591–603, July 1973.

[15] J. B. Rosen, “Existence and uniqueness of equilibrium points for concave $n-$person games,” *Econometrica*, vol. 33, no. 3, pp. 520–534, 1965.

[16] A. Berman and R. J. Plemmons, *Nonnegative matrices in the mathematical sciences*. New York, NY, USA: Academic Press, 1979.

[17] M. Benaim, “Dynamics of stochastic approximation algorithms,” in *Seminaire de Probabilites XXXIII*, e. a. J. Azema, Ed. Springer-Verlag Lecture Notes in Mathematics, 1999, vol. 1709, pp. 1–68.

[18] J. F. Bonnans and A. Shapiro, *Perturbation Analysis of Optimization Problems*. New York, NY, USA: Springer-Verlag, 2000.

[19] R. E. Greene and S. G. Krantz, *Function theory of one complex variable*, 2nd ed. Providence, RI, USA: American Mathematical Society, 2002.

[20] S. G. Krantz, *Function theory of several complex variables*, 2nd ed. Providence, RI, USA: American Mathematical Society, 2001.

[21] R. T. Rockafellar and R. J.-B. Wets, *Variational analysis*. Germany: Springer-Verlag, 1998.

[22] F. H. Clarke, R. J. Stern, and P. R. Wolenski, “Subgradient criteria for monotonicity, the lipschitz condition, and convexity,” *Canadian Journal of Mathematics*, vol. 45, no. 6, pp. 1167–1183, 1993.

[23] R. A. Horn and C. R. Johnson, *Matrix analysis*. New York, NY, USA: Cambridge University Press, 1985.