

# Distributed Precoding for MISO Interference Channels with Channel Mean Feedback: Algorithms and Analysis

Minhua Ding\*, Olav Tirkkonen\*, Randall A. Berry<sup>‡</sup>, and Sennur Ulukus\*\*

\*Dept. of COMNET, Aalto University, Finland

<sup>‡</sup>Dept. of EECS, Northwestern University, USA

\*\*Dept. of ECE, University of Maryland, College Park, USA

Email: minhua.ding@ieee.org, olav.tirkkonen@aalto.fi, rberry@ece.northwestern.edu, ulukus@umd.edu

**Abstract**—This work focuses on the design and analysis of distributed stochastic precoding algorithms for multiple-input single-output (MISO) interference channels, where each transmitter is provided with *mean information* of its intended channel and that of interfering channels. Unlike in cases where exact channel gains are known as in most existing works, here general-rank precoding is required for optimality instead of the rank-one beamforming. An efficient algorithm for the distributed implementation of the Nash equilibrium precoding is first proposed. A sufficient condition for this algorithm to converge to the unique equilibrium is derived for the two-user case based on stochastic ordering, and is valid for a wide range of system parameters. To improve the sum-rate performance under medium to strong interference, a pricing-based algorithm is also provided and its convergence analyzed. The two algorithms are compared in terms of sum-rate and system overhead.

## I. INTRODUCTION

Wireless interference channels have been subject of recent intensive research due to important applications such as cognitive radios [1]. Despite the efforts, the capacity region of interference channels remains unknown except for a few special cases. However, tight bounds have been developed for general and cognitive interference channels [2][3]. Degrees of freedom and interference alignment in multi-antenna interference channels have also been investigated in e.g. [4].

A pragmatic approach in signal processing for interference channels is to treat the interference as additive noise [5][6]. Under this assumption, the competition for shared media among transceivers (users) has been modeled by noncooperative strategic games in digital subscriber lines and multi-antenna wireless interference channels [5]-[7]. For multiple-input single-output (MISO) interference channels, the maximal simultaneously achievable rate tuple of all users, i.e., the Pareto boundary of the achievable rate region, has been characterized [8].

Most existing investigations assume perfect channel state information (CSI) at transmitters [5]-[8]. In practice, however, it is challenging to measure a large amount of CSI accurately and in a timely manner. Robust system designs which consider

imperfect CSI are thus required. Instead of adopting the worst-case design as in [9], here we focus on stochastic designs to utilize the available statistical channel knowledge such as channel covariance information (CCI) and channel mean information (CMI) [10][11]. The CCI and CMI at the transmitters are synonymous to channel covariance feedback and channel mean feedback, respectively [10][11]. Regarding the studies on MISO interference channels with CCI, to date, only a necessary condition has been derived in [12] for precoding to achieve rates on the Pareto boundary; specific precoding strategies have not been proposed. Assuming channel mean feedback, stochastic zero forcing and Nash equilibrium precoding strategies have been proposed in [13]; however, the uniqueness of the equilibrium and its efficient implementation have not been addressed.

In this work, we investigate distributed stochastic precoding algorithms for MISO interference channels with channel mean feedback, treating interference as noise at receivers. An efficient algorithm to implement the Nash equilibrium precoding is first proposed with reduced complexity compared to the method in [13], and a novel sufficient condition for it to converge to the unique fixed point (equilibrium) is derived for the two-user case based on stochastic ordering [14]. Although the Nash equilibrium precoding is a natural steady-state solution in a competitive scenario, it is rate-inefficient under medium to strong interference [7]. To improve the sum-rate, we further introduce interference compensation as in [15][16], provide a distributed pricing-based algorithm and characterize its convergence. Compared to [15][16], our analysis features stochastic optimization and general-rank matrix precoding required for optimality with statistical CSI. The two algorithms are then compared in terms of sum-rate and system overhead, based on both analytical and simulation results.

## II. SYSTEM MODEL AND PROBLEM STATEMENT

We consider a MISO interference channel with  $K$  parallel transmission links. Each transmitter is equipped with  $N$  antennas and each receiver has a single antenna. The transmission of the  $N \times 1$  signal vector  $\mathbf{x}_i$  from the  $i$ -th transmitter is interfered by the transmission of the signal vector  $\mathbf{x}_k$  in the  $k$ -th link,  $\forall k \neq i$ . The observed signal at the  $i$ -th receiver can

This work was supported by the Academy of Finland and the U.S. National Science Foundation under the WiFiUS framework.

be written as

$$y_i = \sqrt{\gamma_i} \mathbf{h}_{ii}^\dagger \mathbf{x}_i + \sum_{k=1, k \neq i}^K \sqrt{\gamma_k} \mathbf{h}_{ki}^\dagger \mathbf{x}_k + n_i, \quad i = 1, \dots, K$$

where  $\dagger$  denotes the complex conjugate transpose, the  $N \times 1$  vector  $\mathbf{h}_{ki}$  denotes the channel from the  $k$ -th transmitter to the  $i$ -th receiver, and  $n_i$  denotes circularly symmetric complex Gaussian noise with zero mean and unit variance (i.e., distributed as  $\mathcal{N}_c(0, 1)$ ). The source signal  $\mathbf{x}_i$  is assumed to be Gaussian with its covariance matrix  $\mathbf{E}\{\mathbf{x}_i \mathbf{x}_i^\dagger\} = \mathbf{Q}_i$ ,  $\forall i = 1, \dots, K$ , and  $\mathbf{E}\{\mathbf{x}_i \mathbf{x}_k^\dagger\} = \mathbf{0}, \forall k \neq i$ . Due to the total transmit power constraint  $\gamma_i$  at the  $i$ -th transmitter, we have  $\text{tr}(\mathbf{Q}_i) \leq 1, \forall i$ . Since the noise variance is normalized,  $\gamma_i$  also denotes the nominal transmit *signal-to-noise ratio* (SNR) of link  $i, \forall i$ . All channels are subject to Rician fading:

$$\mathbf{h}_{ki} = \boldsymbol{\mu}_{ki} + \sqrt{\alpha_{ki}} \mathbf{h}_{w,ki}, \quad i, k \in \{1, \dots, K\} \quad (1)$$

where  $\boldsymbol{\mu}_{ki}$  is the statistical mean of  $\mathbf{h}_{ki}$ ,  $\alpha_{ki} (> 0)$  is a scaling factor, and the  $N \times 1$  vector  $\mathbf{h}_{w,ki}$  denotes the scattering component distributed as  $\mathcal{N}_c(\mathbf{0}, \mathbf{I}_N)$  [10][11]. In general,  $\mathbf{h}_{ki}$  and  $\mathbf{h}_{lm}$  are independent,  $\forall k \neq l$  or  $i \neq m$ . We refer to  $\{\boldsymbol{\mu}_{ki}, \alpha_{ki}\}$  as the CMI of  $\mathbf{h}_{ki}$ .

Each receiver is assumed to have perfect CSI of its own channel only. Each transmitter  $i$  is provided with the CMI of its own channel  $\mathbf{h}_{ii}$  and that of all the interfering channels  $\mathbf{h}_{ki}$  (i.e.,  $\{\boldsymbol{\mu}_{ki}, \alpha_{ki}\}, \forall k$ ), which is referred to as *channel mean feedback* at transmitter  $i$ , for all  $i$ . Receivers here perform single-user decoding and treat interference as noise. With statistical feedback, the performance measure of interest is the ergodic rate in the  $i$ -th link [12]:

$$\bar{R}_i = \mathbf{E} \left\{ \log \left( 1 + \frac{\gamma_i \mathbf{h}_{ii}^\dagger \mathbf{Q}_i \mathbf{h}_{ii}}{1 + \sum_{k=1, k \neq i}^K \gamma_k \mathbf{h}_{ki}^\dagger \mathbf{Q}_k \mathbf{h}_{ki}} \right) \right\} \quad (2)$$

$\forall i = 1, \dots, K$ , where the expectation is with respect to the distributions of  $\{\mathbf{h}_{ki}\}_{k=1}^K$ , and  $\log$  denotes natural logarithm. The rate here is not the optimal information-theoretic rate; it is an achievable rate which we adopt as the payoff function later.

Our goal is to design distributed and efficient precoding algorithms for the above MISO interference channel with channel mean feedback. By *precoding*, we mean the design of  $\{\mathbf{Q}_i\}_{i=1}^K$ . Below two precoding algorithms are analyzed.

### III. DISTRIBUTED IMPLEMENTATION OF NASH EQUILIBRIUM PRECODING

Here we model precoding for MISO interference channels as a noncooperative strategic game as in [5]–[7], [13], in which the players are the  $K$  users (transmitters), the set of actions of transmitter  $i$  is given by  $\{\mathbf{Q}_i : \mathbf{Q}_i \succeq \mathbf{0}, \text{tr}(\mathbf{Q}_i) \leq 1\}$ , and the payoff function of each user is the ergodic rate in (2). Note that  $\mathbf{Q}_i \succeq \mathbf{0}$  means that  $\mathbf{Q}_i$  is positive semi-definite,  $\forall i$ .

The following lemma has been shown in [13] based on the result in [10].

*Lemma 1:* [13] A Nash equilibrium precoding strategy exists for the above game, and has the structure:  $\mathbf{Q}_i^\circ = \mathbf{U}_i^\circ \boldsymbol{\Sigma}_i^\circ \mathbf{U}_i^{\circ \dagger}$ , where

$$\mathbf{U}_i^\circ = [\boldsymbol{\mu}_{ii} / \|\boldsymbol{\mu}_{ii}\| \quad \mathbf{u}_{i2} \dots \mathbf{u}_{iN}], \quad (3)$$

$$\boldsymbol{\Sigma}_i^\circ = \text{diag} \left\{ 1 - p_i^\circ, \frac{p_i^\circ}{N-1}, \dots, \frac{p_i^\circ}{N-1} \right\}, \quad p_i^\circ \in [0, 1], \quad (4)$$

and  $\mathbf{u}_{i2} \dots \mathbf{u}_{iN}$  are arbitrarily chosen orthonormal vectors orthogonal to  $\boldsymbol{\mu}_{ii}, i = 1, \dots, K$ .

From *Lemma 1*, the beamforming directions of  $\mathbf{Q}_i^\circ$  are specified up to a unitary transform of  $\mathbf{u}_{i2} \dots \mathbf{u}_{iN}$ , and the equilibrium power allocation in  $\mathbf{Q}_i^\circ$  is parameterized by a single variable  $p_i^\circ$ , as a result of the following optimization

$$\text{maximize}_{\mathbf{Q}_i \succeq \mathbf{0}, \text{tr}(\mathbf{Q}_i) \leq 1} \bar{R}_i \quad (5)$$

at user  $i, \forall i = 1, \dots, K$  [13]. Clearly, the original precoding game ultimately reduces to a power allocation game.

In [13], the variables  $p_1^\circ \dots p_K^\circ$  have been determined by finding an intersection of the best response functions using a computationally complex multi-dimensional exhaustive search method. The following proposition provides an important step towards reducing this complexity.

*Proposition 1:* In the Nash equilibrium precoding,

$$p_i^\circ \in [0, (N-1)/N], \quad \forall i = 1, \dots, K. \quad (6)$$

*Proof:* See Appendix I.  $\blacksquare$

Without loss of optimality, *Proposition 1* shortens the search range of the equilibrium power allocation by  $1/N$  (see (4) and (6)). Later it will also play an important role in establishing a sufficient condition for a distributed implementation of the equilibrium precoding to converge to its unique fixed point.

For a distributed implementation of the Nash equilibrium precoding, an algorithm can start from any feasible  $\{\mathbf{Q}_l\}_{l=1}^K$  ( $\mathbf{Q}_l \succeq \mathbf{0}, \text{tr}(\mathbf{Q}_l) \leq 1, \forall l$ ) and then sequentially update the precoder of each user (i.e., solving (5) given  $\{\mathbf{Q}_k\}_{k \neq i}$  in a round-robin fashion,  $\forall i$ ). From *Lemma 1*, after at most  $K$  updates, the optimal transmit beams for each user will be fixed according to (3) and the iterations will become power updates only. The powers will be in the range specified by (6). Based on (3), the optimal beamforming directions for each user can be fixed and announced in advance; in fact, only the mean vectors  $\{\boldsymbol{\mu}_{ii}\}_{i=1}^K$  need to be announced in advance, since  $\mathbf{u}_{i2} \dots \mathbf{u}_{iN}, \forall i$ , can be chosen arbitrarily by each user as long as they are orthonormal among each other and orthogonal to  $\boldsymbol{\mu}_{ii}$ . At each iteration only the new power update is announced.

The above simplified but equivalent procedure is summarized as *Algorithm 1*, which requires only one-dimensional (1-D) search (efficient due to concavity) [17][10][11] in each iteration and typically converges in a few iterations (see Section V). In contrast, the method in [13] requires  $K$ -dimensional search for the  $K$ -user case. Since the complexity of exhaustive search grows exponentially with dimensionality, *Algorithm 1* effectively reduces the complexity of implementing the equilibrium precoding.

**Algorithm I: Distributed implementation of Nash equilibrium precoding**

- 1) Announce  $\{\mu_{li}\}_{l=1}^K$  among users, so that  $\{\mathbf{Q}_l\}_{l=1}^K$  are available at all users according to (3). Initialize  $p_l \in [0, 1], \forall l = 1, \dots, K$ . Set **counter** = mod(any nonnegative integer,  $K$ ). Set  $i = \mathbf{counter}$  if **counter**  $\neq 0$ ; otherwise, set  $i = K$ .  $\forall k \neq i$ , user  $k$  announces  $p_k$ . Set  $p_i^{prev} = p_i$ .
- 2) User  $i$  solves the following using efficient 1-D search (Note:  $\bar{R}_i$  is concave in  $p_i$ ):  $p_i = \arg \max_{p_i \in [0, (N-1)/N]} \bar{R}_i$  (given  $p_k, \forall k \neq i$ ), and then announces the new  $p_i$ .
- 3) Stop if the termination condition is met. Otherwise, set **counter** = mod(**(counter + 1)**,  $K$ ). Also set  $i = \mathbf{counter}$  if **counter**  $\neq 0$ ; otherwise, set  $i = K$ . Go back to 2).

The following proposition presents a sufficient condition<sup>1</sup> for Algorithm I to converge to its unique equilibrium power allocation in the *two-user* case.

*Proposition 2:* Consider the MISO interference channel described in Section II with two users. If the following holds:

$$\frac{\alpha_{21}\alpha_{12}}{\alpha_{11}\alpha_{22}} \cdot \mathcal{C}_1\mathcal{C}_2 < 1, \quad (7)$$

then the Nash equilibrium power allocation is unique and Algorithm I converges to the unique equilibrium (up to a unitary transform of  $\mathbf{u}_{i2} \dots \mathbf{u}_{iN}, \forall i$ ). In the above,

$$\mathcal{C}_1 = \sqrt{\frac{\mathbb{E}\left\{\frac{(W_2 - W_1)^2}{[1 + \gamma_1 \alpha_{11} Y_{\min} + \gamma_2 \alpha_{21} W_{\min}]^2}\right\}}{\mathbb{E}\left\{\frac{(Y_2 - Y_1)^2}{[1 + \gamma_1 \alpha_{11} Y_{\max} + \gamma_2 \alpha_{21} W_{\max}]^2}\right\}}} \quad (8)$$

$$\mathcal{C}_2 = \sqrt{\frac{\mathbb{E}\left\{\frac{(Z_2 - Z_1)^2}{[1 + \gamma_2 \alpha_{22} T_{\min} + \gamma_1 \alpha_{12} Z_{\min}]^2}\right\}}{\mathbb{E}\left\{\frac{(T_2 - T_1)^2}{[1 + \gamma_2 \alpha_{22} T_{\max} + \gamma_1 \alpha_{12} Z_{\max}]^2}\right\}}} \quad (9)$$

are constants, where  $W_i, Y_i, Z_i, T_i, i = 1, 2$ , are all independent random variables resulting from alternative expressions of  $\bar{R}_1$  and  $\bar{R}_2$  (see Appendix II), and  $W_{\min}, W_{\max}, Y_{\min}, Y_{\max}, Z_{\min}, Z_{\max}, T_{\min}, T_{\max}$  are also random variables. Specifically, denote  $\boldsymbol{\beta} \triangleq \mathbf{U}_2^{\dagger} \boldsymbol{\mu}_{21} = (\beta_1 \dots \beta_N)^T$  and  $\boldsymbol{\xi} \triangleq \mathbf{U}_1^{\dagger} \boldsymbol{\mu}_{12} = (\xi_1 \dots \xi_N)^T$ , where  $\mathbf{U}_1^{\dagger}$  and  $\mathbf{U}_2^{\dagger}$  are given in (3), and then

$$W_1 = \left| b_{w,1} + \frac{\beta_1}{\sqrt{\alpha_{21}}} \right|^2, \quad Y_1 = \left| \tilde{b}_{w,1} + \frac{\|\boldsymbol{\mu}_{11}\|}{\sqrt{\alpha_{11}}} \right|^2 \quad (10)$$

$$W_2 = \frac{\sum_{i=2}^N \left| b_{w,i} + \frac{\beta_i}{\sqrt{\alpha_{21}}} \right|^2}{N-1}, \quad Y_2 = \frac{\sum_{i=2}^N |\tilde{b}_{w,i}|^2}{N-1} \quad (11)$$

$$Z_1 = \left| \hat{b}_{w,1} + \frac{\xi_1}{\sqrt{\alpha_{12}}} \right|^2, \quad T_1 = \left| \check{b}_{w,1} + \frac{\|\boldsymbol{\mu}_{22}\|}{\sqrt{\alpha_{22}}} \right|^2 \quad (12)$$

$$Z_2 = \frac{\sum_{i=2}^N \left| \hat{b}_{w,i} + \frac{\xi_i}{\sqrt{\alpha_{12}}} \right|^2}{N-1}, \quad T_2 = \frac{\sum_{i=2}^N |\check{b}_{w,i}|^2}{N-1} \quad (13)$$

where  $\|\cdot\|$  denotes the (Euclidean) vector norm;  $b_{w,i}, \tilde{b}_{w,i}, \hat{b}_{w,i}, \check{b}_{w,i}, i = 1, \dots, N$ , are all independent and identically distributed (i.i.d.)  $\mathcal{N}_c(0, 1)$  random variables. In addition,  $W_{\min} =$

<sup>1</sup>Sufficient conditions for the uniqueness of Nash equilibrium have been presented previously, e.g., in [5][6] and references therein, assuming perfect CSI.

$\min\{W_1, \frac{W_1 + (N-1)W_2}{N}\}$ ,  $W_{\max} = \max\{W_1, \frac{W_1 + (N-1)W_2}{N}\}$ ;  $Y_{\min}, Y_{\max}, Z_{\min}, Z_{\max}, T_{\min}$ , and  $T_{\max}$  are defined accordingly.

*Proof:* From any feasible power allocation with fixed optimal transmit beams, after at most two iterations ( $K = 2$ ), the powers will be in  $[0, (N-1)/N]$ . Let  $p_1^{\circ}$  and  $p_2^{\circ}$  be a Nash equilibrium power allocation in  $\mathbf{Q}_1^{\circ}$  and  $\mathbf{Q}_2^{\circ}$ , respectively. Then at time  $t + 1$  ( $t > 2$ ),

$$\begin{aligned} |p_1^{t+1} - p_1^{\circ}| &\leq \left( \max_{(p_1, p_2) \in [0, \frac{N-1}{N}]} \left| \frac{dp_1}{dp_2} \right| \right) |p_2^t - p_2^{\circ}| \\ &\leq \frac{\gamma_2 \alpha_{21}}{\gamma_1 \alpha_{11}} \cdot \mathcal{C}_1 \cdot |p_2^t - p_2^{\circ}| \end{aligned} \quad (14)$$

$$\begin{aligned} &\leq \frac{\gamma_2 \alpha_{21}}{\gamma_1 \alpha_{11}} \cdot \mathcal{C}_1 \left( \max_{(p_1, p_2) \in [0, \frac{N-1}{N}]} \left| \frac{dp_2}{dp_1} \right| \right) |p_1^{t-1} - p_1^{\circ}| \\ &\leq \frac{\gamma_2 \alpha_{21}}{\gamma_1 \alpha_{11}} \cdot \mathcal{C}_1 \cdot \frac{\gamma_1 \alpha_{12}}{\gamma_2 \alpha_{22}} \cdot \mathcal{C}_2 \cdot |p_1^{t-1} - p_1^{\circ}|, \end{aligned} \quad (15)$$

where  $\frac{dp_1}{dp_2}$  denotes the slope of the optimized value of  $p_1$  in  $\mathbf{Q}_1$  as a function of the given  $p_2$  in  $\mathbf{Q}_2$ , and  $\frac{dp_2}{dp_1}$  is defined accordingly. In (14)-(15), we have used:  $\forall p_1, p_2 \in [0, \frac{N-1}{N}]$ ,

$$\max_{(p_1, p_2)} \left| \frac{dp_1}{dp_2} \right| \leq \frac{\gamma_2 \alpha_{21}}{\gamma_1 \alpha_{11}} \cdot \mathcal{C}_1, \quad \max_{(p_1, p_2)} \left| \frac{dp_2}{dp_1} \right| \leq \frac{\gamma_1 \alpha_{12}}{\gamma_2 \alpha_{22}} \cdot \mathcal{C}_2. \quad (16)$$

We prove (16) in Appendix II. When (7) is satisfied, from (15), the iterative updates constitute a pseudocontraction and  $p_1^t \rightarrow p_1^{\circ}$ . Similarly,  $p_2^t \rightarrow p_2^{\circ}$ . From [18, Prop. 1.2], the updates lead to a unique fixed point which is the Nash equilibrium power allocation  $(p_1^{\circ}, p_2^{\circ})$  itself. This establishes the uniqueness of the equilibrium power allocation as well as the convergence of the algorithm. ■

The condition in (7) is effective for a wide range of system parameters in the two-user case, and can be conveniently evaluated either using numerical integration or using Monte Carlo simulations (see Section V). A subtle point is that *Proposition 1* helps significantly to tighten the upper-bounds on  $|\frac{dp_1}{dp_2}|$  and  $|\frac{dp_2}{dp_1}|$  (see (16)); otherwise (7) would be too stringent to be practically useful (see (28) in Appendix II).

#### IV. DISTRIBUTED PRICING-BASED PRECODING

To improve the sum-rate performance of the precoding, we introduce interference compensation as in [15]. Consider the non-convex problem of maximizing the sum-rate:

$$\text{maximize}_{\{\mathbf{Q}_i \succeq \mathbf{0}, \text{tr}(\mathbf{Q}_i) \leq 1\}_{i=1}^K} \sum_{i=1}^K \bar{R}_i. \quad (17)$$

The associated Lagrangian is given by

$$\mathcal{L}(\{\mathbf{Q}_i, \theta_i, \boldsymbol{\Theta}_i\}_{i=1}^K) = \sum_i [\bar{R}_i + \theta_i(1 - \text{tr}(\mathbf{Q}_i)) + \text{tr}(\boldsymbol{\Theta}_i \mathbf{Q}_i)]$$

and the Karush-Kuhn-Tucker conditions are

$$\frac{\partial \mathcal{L}}{\partial \mathbf{Q}_i} = \frac{\partial \bar{R}_i}{\partial \mathbf{Q}_i} + \sum_{k \neq i} \frac{\partial \bar{R}_k}{\partial \mathbf{Q}_i} - \theta_i \mathbf{I}_N + \boldsymbol{\Theta}_i = \mathbf{0}, \quad (18)$$

$$\theta_i(1 - \text{tr}(\mathbf{Q}_i)) = 0, \quad \text{tr}(\boldsymbol{\Theta}_i \mathbf{Q}_i) = 0, \quad (19)$$

$$\text{tr}(\mathbf{Q}_i) \leq 1, \quad \mathbf{Q}_i \succeq \mathbf{0}, \quad \boldsymbol{\Theta}_i \succeq \mathbf{0}, \quad \theta_i \geq 0, \quad \forall i. \quad (20)$$

Though in principle (17) can be solved in a centralized fashion using exhaustive searches [12] or potentially by the method in [19], the computational complexity would be very high.

Here we are interested in distributed and computationally less complex solutions. Thus, parallel to the development in [15], we consider a pricing-based approach and introduce the interference prices as  $\mathbf{\Pi}_{ki} = -\frac{\partial \bar{R}_k}{\partial \mathbf{Q}_i}$ ,  $\forall k \neq i$ , according to the second term on the right-hand side of (18). Introducing interference prices allows for altruistic behaviors among users [15]. In contrast to the scalar prices in [15][16], here

$$\mathbf{\Pi}_{ki} = \mathbb{E} \left\{ \frac{\gamma_k \mathbf{h}_{kk}^\dagger \mathbf{Q}_k \mathbf{h}_{kk}}{1 + \sum_l \gamma_l \mathbf{h}_{lk}^\dagger \mathbf{Q}_l \mathbf{h}_{lk}} \cdot \frac{\gamma_i \mathbf{h}_{ik} \mathbf{h}_{ik}^\dagger}{1 + \sum_{l \neq k} \gamma_l \mathbf{h}_{lk}^\dagger \mathbf{Q}_l \mathbf{h}_{lk}} \right\}$$

is a Hermitian positive semidefinite matrix evaluated at user  $k$ ,  $\forall k \neq i$ . Different from the cases with perfect CSI [15][16], here the interference prices are not separable from the instantaneous channel coefficients due to the assumption of channel mean feedback, and general-rank matrix precoding is considered for optimality instead of the rank-one beamforming.

After evaluating and exchanging the interference prices, user  $i$  maximizes the surplus, which is  $\bar{R}_i$  minus the cost for causing interference to other users:

$$\underset{\substack{\mathbf{Q}_i \succeq \mathbf{0} \\ \text{tr}(\mathbf{Q}_i) \leq 1}}{\text{maximize}} \quad \bar{R}_i - \sum_{k \neq i} \text{tr}(\mathbf{Q}_i \mathbf{\Pi}_{ki}), \quad \forall i. \quad (21)$$

For each user  $i$ , this is a convex stochastic program which does not admit a closed-form solution in general. A well-established approach in stochastic optimization is the Sample Average Approximation (SAA) [20, Ch. 5]. Following the SAA method,  $\bar{R}_i$  in (21) (i.e., that in (2)) is replaced with

$$\widehat{\bar{R}}_i = \frac{1}{M} \sum_{m=1}^M \log \left( 1 + \frac{\gamma_i \mathbf{h}_{ii}^{[m]\dagger} \mathbf{Q}_i \mathbf{h}_{ii}^{[m]}}{1 + \sum_{k=1, k \neq i}^K \gamma_k \mathbf{h}_{ki}^{[m]\dagger} \mathbf{Q}_k \mathbf{h}_{ki}^{[m]}} \right) \quad (22)$$

and then user  $i$  solves

$$\underset{\substack{\mathbf{Q}_i \succeq \mathbf{0} \\ \text{tr}(\mathbf{Q}_i) \leq 1}}{\text{maximize}} \quad \widehat{\bar{R}}_i - \sum_{k \neq i} \text{tr}(\mathbf{Q}_i \mathbf{\Pi}_{ki}) \quad (23)$$

instead. In (22), for each  $m$ ,  $\mathbf{h}_{ki}^{[m]}$  denotes an independent realization of the random vector  $\mathbf{h}_{ki}$ ,  $\forall k, i \in \{1, \dots, K\}$ . By the law of large numbers,  $\widehat{\bar{R}}_i$  is a good approximation to  $\bar{R}_i$  when  $M$  is sufficiently large [20, Ch. 5]. Note that (23) is a convex program, and thus can be efficiently solved with standard convex optimization tools [21].

A distributed pricing-based sequential precoding algorithm, Algorithm II, can be developed based on [16]. By setting all prices to be zero matrices, it also serves as a distributed implementation of the Nash equilibrium precoding in which (5) is solved using the SAA method with  $\bar{R}_i$  replaced by  $\widehat{\bar{R}}_i$  in (22).

*Proposition 3:* Assuming that user  $i$  solves (21) in the precoder update of Step 2) in Algorithm II, the sum-rate is nondecreasing after each update of the precoder.

---

#### Algorithm II: Pricing-based sequential precoding

---

- 1) Initialize  $\mathbf{Q}_l$ :  $\mathbf{Q}_l \succeq \mathbf{0}$ ,  $\text{tr}(\mathbf{Q}_l) \leq 1, \forall l = 1, \dots, K$ ;  
Set **counter** = mod(any nonnegative integer,  $K$ ).  
Set  $i = \mathbf{counter}$  if **counter**  $\neq 0$ ; otherwise, set  $i = K$ .  
 $\forall k \neq i$ , user  $k$  announces  $\mathbf{Q}_k$ .
  - 2) [**Price update**]  $\forall k \neq i$ , user  $k$  calculates and announces  $\mathbf{\Pi}_{ki}$ .  
[**Precoder update**] User  $i$  solves (23) (given  $\{\mathbf{Q}_k, \mathbf{\Pi}_{ki}\}_{k \neq i}$ ) and announces the new  $\mathbf{Q}_i$ .
  - 3) Stop if the termination criterion is met. Otherwise, set **counter** = mod( $(\mathbf{counter} + 1), K$ ). Also set  $i = \mathbf{counter}$  if **counter**  $\neq 0$ ; otherwise, set  $i = K$ . Go back to 2).
- 

*Proof:* The proof follows similar lines to those in [16] and hinges on the fact that  $\bar{R}_k$  is a convex function of  $\mathbf{Q}_i, \forall i \neq k$  [21, p. 70]. Details are omitted due to space constraints. ■

Since the updated sum-rates form a nondecreasing sequence and are bounded from above, Algorithm II converges, assuming that (21) is solved with sufficient accuracy by using (23) and a sufficiently large  $M$  in (22).

#### V. NUMERICAL EXAMPLES AND DISCUSSIONS

Here the number of users is set to be  $K = 2$ . The number of samples in the SAA method (22) is set to be  $M = 2000$ . Monte Carlo simulations are used to evaluate  $\mathcal{C}_1$  and  $\mathcal{C}_2$  in (7) using  $10^6$  samples.

##### A. Comparison of the two algorithms in terms of sum-rate

**Example 1:** (Weak mean interference)  $N = 4, \gamma_1 = \gamma_2$ ;  $\alpha_{11} = \alpha_{22} = 0.2$ ;  $\alpha_{12} = \alpha_{21} = 0.3$ ;  $j = \sqrt{-1}$   
 $\boldsymbol{\mu}_{11} = [0.6541 - 0.1027j; -0.0207 + 0.3992j; 0.1800 - 0.3236j; 0.0960 + 1.2008j]$ ; ( $\|\boldsymbol{\mu}_{11}\| = 1.4787$ )  
 $\boldsymbol{\mu}_{22} = [0.3929 + 0.6897j; 0.8930 - 0.8766j; -0.3805 - 0.7926j; 0.4719 + 0.3142j]$ ; ( $\|\boldsymbol{\mu}_{22}\| = 1.8139$ )  
 $\boldsymbol{\mu}_{12} = [0.0504 - 0.1457j; 0.0360 - 0.0576j; 0.0033 - 0.1824j; -0.1544 - 0.1624j]$ ; ( $\|\boldsymbol{\mu}_{12}\| = 0.3345$ )  
 $\boldsymbol{\mu}_{21} = [0.0557 + 0.0053j; 0.1166 + 0.0881j; 0.0951 + 0.0741j; 0.0752 - 0.0333j]$ ; ( $\|\boldsymbol{\mu}_{21}\| = 0.2140$ ).

**Example 2:** (Strong mean interference)  $N = 4, \gamma_1 = \gamma_2$ ;  $\alpha_{11}, \alpha_{22}, \alpha_{12}, \alpha_{21}, \boldsymbol{\mu}_{11}$  and  $\boldsymbol{\mu}_{22}$  are the same as in **Example 1**;  $\boldsymbol{\mu}_{21}$  and  $\boldsymbol{\mu}_{12}$  are five times their counterparts in **Example 1**, respectively. Thus,  $\|\boldsymbol{\mu}_{12}\| = 1.6724$  and  $\|\boldsymbol{\mu}_{21}\| = 1.0699$ .

We first examine the condition in (7). From Table I, we can see that (7) is satisfied in all the cases here. In Fig. 1 we present the convergence curves corresponding to **Example 2** with  $\gamma_1 = \gamma_2 = 24$  dB. (While we can fix the optimal beaming directions in Algorithm I, here we actually let both algorithms start from the same arbitrarily chosen feasible  $\{\mathbf{Q}_i\}_{i=1}^2$ ). The convergence curves for **Example 1** are similar. Both algorithms typically converge within a few iterations from any feasible point in all the simulations conducted here.

The sum-rates obtained for both examples are given in Fig. 2. Since the two examples differ only in the mean strength of  $\boldsymbol{\mu}_{12}$  and  $\boldsymbol{\mu}_{21}$ , clearly, without interference cancellation, stronger interference can only lead to worse performance. When the mean interference is relatively weak, the

Table I: Values of  $\alpha_{21}\alpha_{12}C_1C_2/(\alpha_{11}\alpha_{22})$  in (7) in both examples

$\gamma_1 = \gamma_2$	0 dB	3 dB	6 dB	9 dB	12 dB
<b>Example 1</b>	0.0777	0.1077	0.1374	0.1608	0.1768
<b>Example 2</b>	0.3677	0.5250	0.6816	0.8079	0.8916
$\gamma_1 = \gamma_2$	15 dB	18 dB	21 dB	24 dB	—
<b>Example 1</b>	0.1862	0.1910	0.1938	0.1951	—
<b>Example 2</b>	0.9434	0.9706	0.9884	0.9941	—

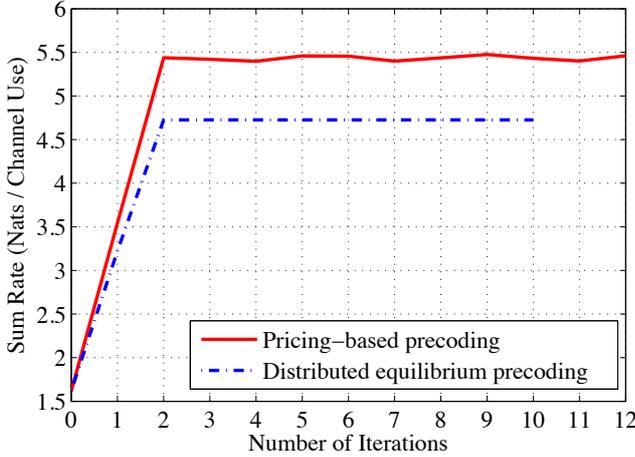


Fig. 1: Convergence of both algorithms in **Example 2** with  $\gamma_1 = \gamma_2 = 24$  dB). (Similar curves have been obtained using other values of  $\gamma_1$  and  $\gamma_2$  and for **Example 1**.)

two algorithms share a similar performance, which can be analytically shown and is seen in Fig. 2. With strong mean interference, the performance of the equilibrium precoding is clearly interference-limited, whereas the pricing-based algorithm effectively improves the sum-rate, especially at medium to high transmit SNRs.

### B. Comparison of the two algorithms in terms of system overhead

Both algorithms require the same amount of CSI at transmitters and receivers (see Section II).

During each iteration, Algorithm II requires one user to announce the transmit covariance matrix and requires  $(K-1)$  users to update and announce the price matrices. Since all the involved matrices are Hermitian positive semi-definite, the amount of updates can be roughly halved. In comparison, using Algorithm I, only a real scalar needs to be announced at each iteration, which constitutes very little overhead.

## VI. CONCLUDING REMARKS

Two distributed precoding algorithms have been investigated for MISO interference channels with channel mean feedback. For the distributed implementation of the Nash equilibrium precoding, we have provided an efficient algorithm and derived a sufficient condition for it to converge to the unique equilibrium in the two-user case. A pricing-based precoding algorithm has also been provided together with

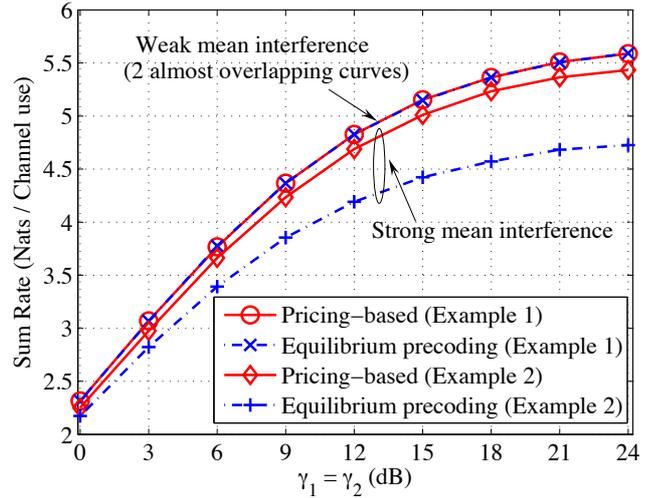


Fig. 2: Comparison of the sum-rates obtained from both algorithms in both examples.

its convergence property, which outperforms the equilibrium precoding in terms of sum-rate under strong interference at the cost of higher overhead.

## APPENDIX I PROOF OF Proposition 1

*Lemma 2:* [14, Theorem 5.A.4] Let  $V_1$  and  $V_2$  be two nonnegative random variables. Then  $V_1$  is smaller than  $V_2$  in the Laplace transform order [i.e.,  $\mathbb{E}\{\exp\{-sV_1\}\} \geq \mathbb{E}\{\exp\{-sV_2\}\}, \forall s > 0$ , by definition, denoted  $V_1 \leq_{\text{LT}} V_2$ ], if and only if  $\mathbb{E}\{\psi(V_1)\} \leq \mathbb{E}\{\psi(V_2)\}$  for all differentiable functions  $\psi$  defined on  $[0, \infty)$  with a completely monotone derivative [i.e.,  $(-1)^{n-1}\psi^{(n)}(v) \geq 0, \forall v > 0, \forall n = 1, 2, \dots$ ], provided that the expectations exist.

*Lemma 3:* Let  $\varphi$  be a differentiable function defined on  $[0, \infty)$  with a completely monotone derivative. Define  $\lambda = (\lambda_1 \dots \lambda_N)^T$ . Consider the following problem

$$\text{maximize}_{\substack{\lambda_i \geq 0, \forall i \\ \sum_i \lambda_i = 1}} \mathbb{E} \left\{ \varphi \left( \sum_{i=1}^N \lambda_i |q_{w,i} + \tau_i|^2 \right) \right\}, \quad (24)$$

where  $q_{w,i}, i = 1, \dots, N$ , are i.i.d.  $\mathcal{N}_c(0, 1)$  random variables, and  $\tau_i, i = 1, \dots, N$ , are constants. If  $|\tau_k| \geq |\tau_m|, \forall k \neq m, k, m \in \{1, \dots, N\}$ , then  $\lambda_k \geq \lambda_m$  in the optimum  $\lambda$ .

*Proof:* Define the random variable  $S_1$  as

$$S_1 = \sum_{i \neq k, m} \lambda_i |q_{w,i} + \tau_i|^2 + \lambda_m |q_{w,k} + \tau_k|^2 + \lambda_k |q_{w,m} + \tau_m|^2$$

with  $|\tau_k| \geq |\tau_m|$  and  $\lambda_k \geq \lambda_m$ . Define  $S_2 = \sum_{i=1}^N \lambda_i |q_{w,i} + \tau_i|^2$ , which is the same as  $S_1$  except that the positions of  $\lambda_k$  and  $\lambda_m$  are swapped. Then we need to show  $\mathbb{E}\{\varphi(S_1)\} \leq \mathbb{E}\{\varphi(S_2)\}$ . Let the Laplace transforms of the probability density functions of  $S_1$  and  $S_2$  be  $\mathcal{L}_1(s)$  and  $\mathcal{L}_2(s)$ , respectively, which can be obtained from [22, p. 44, Eq. (2-1-117)].

According to *Lemma 2*, it is sufficient to show  $S_1 \leq_{\text{LT}} S_2$ , or equivalently,  $\mathcal{L}_2(s) \leq \mathcal{L}_1(s)$ ,  $\forall s > 0$ . But  $\mathcal{L}_i(s) > 0$ ,  $\forall s > 0$ ,  $i = 1, 2$  and

$$\log \frac{\mathcal{L}_2(s)}{\mathcal{L}_1(s)} = \frac{s(\lambda_m - \lambda_k)(|\tau_k|^2 - |\tau_m|^2)}{(1 + \lambda_k s)(1 + \lambda_m s)} \leq 0, \forall s > 0,$$

since  $|\tau_k| \geq |\tau_m|$  and  $\lambda_k \geq \lambda_m$ . This concludes the proof. ■

Now we are ready to prove *Proposition 1*.

*Proof:* Note that  $\tilde{\varphi}_i(\tilde{v})$  below is a differentiable function defined on  $[0, +\infty)$  with a completely monotone derivative:

$$\tilde{\varphi}_i(\tilde{v}) \triangleq \mathbb{E}_{\{\mathbf{h}_{ki}\}_{k \neq i}} \left\{ \log \left( 1 + \frac{\gamma_i \alpha_{ii} \cdot \tilde{v}}{1 + \sum_{k \neq i} \gamma_k \mathbf{h}_{ki}^\dagger \mathbf{Q}_k \mathbf{h}_{ki}} \right) \right\}$$

$\forall i$ . Expand  $\frac{1}{\alpha_{ii}} \mathbf{h}_{ii}^\dagger \mathbf{Q}_i^\circ \mathbf{h}_{ii}$  using (3)-(4), and then we obtain  $\bar{R}_i = \mathbb{E}_{\mathbf{h}_{ii}} \left\{ \tilde{\varphi}_i \left( \frac{1}{\alpha_{ii}} \mathbf{h}_{ii}^\dagger \mathbf{Q}_i^\circ \mathbf{h}_{ii} \right) \right\}$  which is equivalent to

$$\mathbb{E} \left\{ \tilde{\varphi}_i \left( (1 - p_i^\circ) \left| \eta_{w,1} + \frac{\|\boldsymbol{\mu}_{ii}\|^2}{\alpha_{ii}} \right| + \frac{p_i^\circ}{N-1} \sum_{i=2}^N |\eta_{w,i}|^2 \right) \right\}$$

where  $\eta_{w,i}$ ,  $i = 1, \dots, N$ , are i.i.d.  $\mathcal{N}_c(0, 1)$  random variables. Applying *Lemma 3*, we obtain  $(1 - p_i^\circ) \geq p_i^\circ / (N - 1)$ ,  $\forall i$ . ■

## APPENDIX II PROOF OF (16)

*Proof:* Due to space constraints, we provide only an outline of the first inequality in (16); the second one follows using similar arguments. Suppose that in the previous iteration  $p_2$  in  $\mathbf{Q}_2$  has been updated. Now we need to find the optimal  $p_1$  in  $\mathbf{Q}_1$  which maximizes  $\bar{R}_1$ , given  $p_2$ . Based on the known optimal eigen-structures of  $\{\mathbf{Q}_i\}_{i=1}^2$  (see (3)-(4)), we obtain an equivalent form of  $\bar{R}_1$  given below:

$$\bar{R}_1 = \mathbb{E} \left\{ \log \left( 1 + \frac{\gamma_1 \alpha_{11} [(1 - p_1) Y_1 + p_1 Y_2]}{1 + \gamma_2 \alpha_{21} [(1 - p_2) W_1 + p_2 W_2]} \right) \right\}$$

where the random variables  $W_1, W_2, Y_1$  and  $Y_2$  have been introduced in (10)-(11). Since  $\bar{R}_1$  is concave in  $p_1$ , the necessary and sufficient condition for  $p_1 \in [0, 1]$  to be optimal for  $\bar{R}_1$  is either i)  $\frac{\partial \bar{R}_1}{\partial p_1} = 0$ , or ii)  $p_1 = 0$  and  $\frac{\partial \bar{R}_1}{\partial p_1} |_{p_1=0} < 0$  (and thus  $\frac{\partial \bar{R}_1}{\partial p_1} < 0$ ,  $\forall p_1 \in [0, 1]$ ). In either case, define

$$f(p_1, p_2) \triangleq \frac{\partial \bar{R}_1}{\partial p_1} = \mathbb{E} \left\{ \frac{\gamma_1 \alpha_{11} (Y_2 - Y_1)}{1 + \gamma_1 \alpha_{11} Y_{p_1} + \gamma_2 \alpha_{21} W_{p_2}} \right\} \quad (25)$$

with  $Y_{p_1} \triangleq (1 - p_1) Y_1 + p_1 Y_2$  and  $W_{p_2} \triangleq (1 - p_2) W_1 + p_2 W_2$ . By setting  $f(p_1, p_2)$  in (25) at any valid value, we obtain  $p_1$  as an implicit (smooth) function of the given  $p_2$ , from which we have  $\frac{\partial f}{\partial p_1} dp_1 + \frac{\partial f}{\partial p_2} dp_2 = 0$  or  $\frac{dp_1}{dp_2} = -\frac{\partial f}{\partial p_2} / \left( \frac{\partial f}{\partial p_1} \right)$ . Thus,

$$\left| \frac{dp_1}{dp_2} \right| = \frac{\gamma_2 \alpha_{21} \mathbb{E} \left\{ \frac{(W_2 - W_1)(Y_2 - Y_1)}{[1 + \gamma_1 \alpha_{11} Y_{p_1} + \gamma_2 \alpha_{21} W_{p_2}]^2} \right\}}{\gamma_1 \alpha_{11} \mathbb{E} \left\{ \frac{(Y_2 - Y_1)^2}{[1 + \gamma_1 \alpha_{11} Y_{p_1} + \gamma_2 \alpha_{21} W_{p_2}]^2} \right\}} \quad (26)$$

$$\leq \frac{\gamma_2 \alpha_{21}}{\gamma_1 \alpha_{11}} \sqrt{\frac{\mathbb{E} \left\{ \frac{(W_2 - W_1)^2}{[1 + \gamma_1 \alpha_{11} Y_{p_1} + \gamma_2 \alpha_{21} W_{p_2}]^2} \right\}}{\mathbb{E} \left\{ \frac{(Y_2 - Y_1)^2}{[1 + \gamma_1 \alpha_{11} Y_{p_1} + \gamma_2 \alpha_{21} W_{p_2}]^2} \right\}}} \quad (27)$$

$$\leq \frac{\gamma_2 \alpha_{21}}{\gamma_1 \alpha_{11}} \sqrt{\frac{\mathbb{E} \left\{ \frac{(W_2 - W_1)^2}{[1 + \gamma_1 \alpha_{11} Y_{\min} + \gamma_2 \alpha_{21} W_{\min}]^2} \right\}}{\mathbb{E} \left\{ \frac{(Y_2 - Y_1)^2}{[1 + \gamma_1 \alpha_{11} Y_{\max} + \gamma_2 \alpha_{21} W_{\max}]^2} \right\}}} = \frac{\gamma_2 \alpha_{21}}{\gamma_1 \alpha_{11}} \mathcal{C}_1, \quad (28)$$

where to obtain (27), we have applied the *Cauchy-Schwarz* inequality  $[\mathbb{E}(A_1 A_2)]^2 \leq \mathbb{E}(A_1^2) \mathbb{E}(A_2^2)$  to the numerator of (26), and to obtain (28), we have applied *Proposition 1* and utilized the fact that  $Y_{\min} \leq Y_{p_1} \leq Y_{\max}$  and  $W_{\min} \leq W_{p_2} \leq W_{\max}$ . Note that (26)-(28) apply to all  $p_1, p_2 \in [0, (N-1)/N]$ . Thus, the first part in (16) is proved. ■

## REFERENCES

- [1] A. Goldsmith, S. A. Jafar, I. Maric, and S. Srinivasa, "Breaking spectrum gridlock with cognitive radios: an information theoretic perspective," *Proc. IEEE*, vol. 97, no. 5, pp. 894-914, 2009.
- [2] T. S. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Trans. Inf. Theory*, vol. IT-27, no. 1, pp. 49-60, Jan. 1981.
- [3] N. Devroye, P. Mitran, and V. Tarokh, "Achievable rates in cognitive radio channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 1813-1827, May 2006.
- [4] T. Gou and S. A. Jafar, "Degrees of freedom of the  $K$ -user  $M \times N$  MIMO interference channel," *IEEE Trans. Inf. Theory*, vol. 56, no. 12, pp. 6040-6057, Dec. 2010.
- [5] G. Scutari, D. P. Palomar, and S. Barbarossa, "Optimal linear precoding strategies for wideband noncooperative systems based on game theory - Part I: Nash equilibria," *IEEE Trans. Signal Process.*, vol. 56, no. 3, pp. 1230-1249, Mar. 2008.
- [6] W. Yu, G. Ginis, and J. M. Cioffi, "Distributed multiuser power control for digital subscriber lines," *IEEE JSAC*, vol. 20, no. 5, pp. 1105-1115, Jun. 2002.
- [7] E. G. Larsson and E. Jorswieck, "Competition versus cooperation on the MISO interference channel," *IEEE JSAC*, vol. 26, no. 7, pp. 1059-1069, Sep. 2008.
- [8] E. A. Jorswieck, E. G. Larsson, and D. Danev, "Complete characterization of the Pareto boundary for the MISO interference channel," *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 5292-5296, Oct. 2008.
- [9] J. Wang, G. Scutari, and D. P. Palomar, "Robust MIMO cognitive radio via game theory," *IEEE Trans. Signal Process.*, vol. 59, no. 3, pp. 1183-1201, Mar. 2011.
- [10] E. Visotsky and U. Madhow, "Space-time transmit precoding with imperfect feedback," *IEEE Trans. Inf. Theory*, vol. 47, no. 6, pp. 2632-2639, Sep. 2001.
- [11] S. A. Jafar and A. Goldsmith, "Transmitter optimization and optimality of beamforming for multiple antenna systems with imperfect feedback," *IEEE Trans. Wireless Commun.*, vol. 3, no. 4, pp. 1165-1175, Jul. 2004.
- [12] J. Lindblom, E. G. Larsson, and E. A. Jorswieck, "Parameterization of the MISO IFC rate region: the case of partial channel state information," *IEEE Trans. Wireless Commun.*, vol. 9, no. 2, pp. 500-504, Feb. 2010.
- [13] M. Ding and Q. T. Zhang, "Stochastic precoding for MISO interference channels with channel mean feedback," *IEEE Trans. Commun.*, vol. 60, no. 4, pp. 1082-1090, Apr. 2012.
- [14] M. Shaked and J. Shanthikumar, *Stochastic Orders*, Springer, 2007.
- [15] J. Huang, R. A. Berry, and M. L. Honig, "Distributed interference compensation for wireless networks," *IEEE JSAC*, vol. 24, no. 5, pp. 1074-1084, May 2006.
- [16] C. Shi, R. A. Berry, and M. L. Honig, "Monotonic convergence of distributed interference pricing in wireless networks," in *Proc. IEEE ISIT*, Jun. 2009, pp. 1619-1623.
- [17] E. Chong and S. Zak, *An Introduction to Optimization*, Wiley, 2008.
- [18] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and distributed computation: numerical methods*, Athena Scientific, 1997.
- [19] E. A. Jorswieck and E. G. Larsson, "Monotonic optimization framework for the two-user MISO interference channel," *IEEE Trans. Commun.*, vol. 58, no. 7, pp. 2159-2168, Jul. 2010.
- [20] A. Shapiro, D. Dentcheva, and A. Ruszczyński, *Lectures on Stochastic Programming: Modeling and Theory*, SIAM, 2009.
- [21] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.
- [22] J. G. Proakis, *Digital Communications*, McGraw-Hill, 2000.