Distributed Interference Pricing with MISO Channels

Changxin Shi, Randall A. Berry, and Michael L. Honig
Department of Electrical Engineering and Computer Science
Northwestern University, Evanston, Illinois 60208
Email: cshi@u.northwestern.edu, {rberry, mh}@ece.northwestern.edu

Abstract

We study a distributed algorithm for adapting transmit beamforming vectors in a multi-antenna peer-to-peer wireless network. The algorithm attempts to maximize a sum of per-user utility functions, where each user's utility is a function of his transmission rate, or equivalently the received signal-to-interference plus noise ratio (SINR). This is accomplished by exchanging *interference prices*, each of which represents the marginal cost of interference to a particular user. Given the interference prices, users update their beamforming vectors to maximize their utility minus the cost of interference. For a two-user system, we show that this algorithm converges for a suitable class of utility functions. Convergence of the algorithm with more than two users is illustrated numerically.

I. INTRODUCTION

Mitigating interference is critical for efficiently sharing wireless spectrum. When nodes are equipped with multiple antennas, the additional spatial degrees of freedom can be exploited to reduce interference. This paper considers such a setting, namely an interference channel consisting of multi-input, single-output (MISO) wireless links. The objective is to select the beamforming vector and transmission power level at each transmitter in an attempt to maximize the overall network performance. Performance is measured in terms of the sum of per user utilities, where each utility is a function of the transmission rate, or equivalently the signal-to-interference plus noise ratio (SINR). In a network without centralized control (e.g., an ad hoc network), maximizing the total utility is complicated due to the interference among users, which causes each user's beamforming vector to affect not only that user's utility but the utility of every other user. Solving this in a scalable manner requires a distributed algorithm with limited information exchange. Such an algorithm is the focus of this paper.

The algorithm we study is motivated by the work in [1] and [2], which consider algorithms for power allocation in single-antenna wireless networks. These *asynchronous distributed pricing* (ADP) algorithms are based on exchanging *interference prices* among users[1], where an interference price is a user's marginal change in utility per unit interference power. Given the interference prices from the neighboring users, each user then optimizes his own utility minus the interference cost to other users. By iteratively updating interference prices and powers, the algorithms in [1], [2] are shown to converge under suitable assumptions on the users' utility functions.

The ADP algorithm for a single-antenna network in [1] can be directly generalized to a MISO network, as shown in Section II. Such an algorithm was previously presented in [3] for the special case of rate utilities. Simulation results shown there with two users indicate that this algorithm converges to an allocation that is essentially optimal. However, there is no proof that the algorithm must converge, and it is not clear how it performs with other utility functions. Moreover, the convergence proof in

[1]We refer to each transmitter-receiver pair as a "user".

[1] for a single antenna network, which is based on relating this algorithm to best response updates in a supermodular game, does not directly generalize to this setting. Likewise, the convergence proof for the modified algorithm in [2] does not apply here. For a two user MISO interference channel we prove that this algorithm does converge under a similar condition as in [1]. Our proof is based on first showing several properties of the optimal beamforming vectors, which enable us to re-parameterize the original optimization problem. The steps in the ADP algorithm can also be written in terms of these re-parameterized values. After doing this, we can show that the algorithm again corresponds to best response updates in a supermodular game, under suitable choices of utility functions.

In terms of related work, in addition to [1], [2], the algorithms in [4] and [5] also exchange information similar to interference prices to facilitate distributed power control in single antenna interference channels. For MIMO interference channels, the most common approaches that have been studied for distributed optimization are based on iterative water-filling (e.g. [6]–[8]). In these approaches each user iteratively selects their transmit covariance matrix to maximize their rate given the current interference. Such approaches may not converge in general or may converge to a rate-pair that is not Pareto optimal. In [9], it is shown that for a two-user MISO interference channel, any Pareto optimal rate-pair can be achieved as a linear combination of a beamformer corresponding to the Nash equilibrium and the zero-forcing beamformer. This is closely related to the decomposition result we give in Section III. However, [9] does not address how to determine this combination in a distributed manner, which is our focus here.

In the next section, we give our system model and generalize the ADP algorithm in [1] to this setting. Section III contains our main analytical results. Simulation results are presented in Section IV, and conclusions are given in Section V.

II. System Model and Generalized Asynchronous Distributed Algorithm

We consider a MISO wireless network with a narrowband time-invariant wireless channel, in which there are M pairs of transmitters and receivers, each with N transmit antennas sharing a single flat fading channel. The received signal for user $i = 1, \ldots, M$ is given by

$$y_i = \mathbf{v_i}^\dagger \mathbf{h_{ii}} x_i + \sum_{j \neq i} \mathbf{v_j}^\dagger \mathbf{h_{ji}} x_j + n \tag{1}$$

where $\dagger$ denotes Hermitian transpose, $\mathbf{h_{ij}} = [h_{ij}^1, h_{ij}^2, \cdots, h_{ij}^N]^T$, is the channel vector from the i-th transmitter to the j-th receiver through the N antennas, x_i is the transmitted symbol of user i, n is additive complex Gaussian noise with variance n_0, and $\mathbf{v_i}$ is the beamforming vector for user i. Assuming the transmit symbol has unit variance for all users, the received SINR for each user i can be written as

$$\gamma_i = \frac{|\mathbf{v_i}^\dagger \mathbf{h_{ii}}|^2}{n_0 + \sum_{j \neq i} |\mathbf{v_j}^\dagger \mathbf{h_{ji}}|^2}. \tag{2}$$

The quality of service for each user i is measured via a utility function $u_i(\gamma_i)$, which is assumed to be a monotonically increasing and concave function of the received SINR[2]. Our objective is to maximize the total utility over all users' beamforming vectors $\{\mathbf{v_i}\}$, i.e.,

$$\max_{\mathbf{v_1}, \cdots, \mathbf{v_M}} \sum_{i=1}^M u_i \left(\frac{|\mathbf{v_i}^\dagger \mathbf{h_{ii}}|^2}{n_0 + \sum_{j \neq i} |\mathbf{v_j}^\dagger \mathbf{h_{ji}}|^2} \right) \tag{P_0}$$
$$\text{s.t. } |\mathbf{v_i}|^2 \leq P_i^{max} \quad \text{for all } i = 1, \ldots, M$$

where P_i^{max} denotes a power constraint for user i.

---

[2]Equivalently, one can view each user's utility as a function of his transmission rate given by the Shannon capacity with interference treated as noise.

Any locally optimal solution $\tilde{\mathbf{v}}_1, \ldots, \tilde{\mathbf{v}}_M$ of Problem P_0 should satisfy the Karush-Kuhn-Tucker (KKT) conditions [10]. In particular, there must exist unique Lagrange multipliers $\lambda_i \geq 0$ associated with each power constraint such that for all $i = 1, \ldots, M$:

$$\left[\frac{u_i'}{n_0 + \sum_{j \neq i} |\tilde{\mathbf{v}}_j^\dagger \mathbf{h}_{ji}|^2} \left(\mathbf{h}_{ii} \mathbf{h}_{ii}^\dagger \right) - \sum_{j \neq i} \frac{u_j' |\tilde{\mathbf{v}}_j^\dagger \mathbf{h}_{jj}|^2}{(n_0 + \sum_{k \neq j} |\tilde{\mathbf{v}}_k^\dagger \mathbf{h}_{kj}|^2)^2} \left(\mathbf{h}_{ij} \mathbf{h}_{ij}^\dagger \right) \right] \tilde{\mathbf{v}}_i = \lambda_i \tilde{\mathbf{v}}_i \tag{3}$$

where u_i' denotes the first order derivative of $u_i(\gamma_i)$ with respect to γ_i. Note that u_i' is a function of $\tilde{\mathbf{v}}_1, \ldots, \tilde{\mathbf{v}}_M$ but to simplify notation we do not explicitly denote this.

Following [1], for any given choice of beamforming vectors $\mathbf{v}_1, \ldots, \mathbf{v}_M$, let

$$\pi_i = -\frac{\partial u_i}{\partial I_i} = \frac{u_i' |\mathbf{v}_i^\dagger \mathbf{h}_{ii}|^2}{(n_0 + \sum_{j \neq i} |\mathbf{v}_j^\dagger \mathbf{h}_{ji}|^2)^2} \tag{4}$$

be the interference price for user i, where $I_i = \sum_{j \neq i} |\mathbf{v}_j^\dagger \mathbf{h}_{ji}|^2$ is the total interference power for user i, which depends on the beamforming vectors of all users other than i. Given fixed interference prices and beamforming vectors for every other user, suppose that user i then solves the following subproblem

$$\max_{\mathbf{v}_i} u_i \left(\frac{|\mathbf{v}_i^\dagger \mathbf{h}_{ii}|^2}{n_0 + \sum_{j \neq i} |\mathbf{v}_j^\dagger \mathbf{h}_{ji}|^2} \right) - \sum_{j \neq i} \pi_j |\mathbf{v}_i^\dagger \mathbf{h}_{ij}|^2 \tag{P_i}$$
$$\text{s.t. } |\mathbf{v}_i|^2 \leq P_i^{max}.$$

It can be seen that if the other users' interference prices and beamforming vectors are set at their optimal values, then the KKT condition of this subproblem matches the i-th KKT condition for Problem P_0. In other words, some locally optimal action for user i in this subproblem will also be the same as that user's action in the globally optimal solution of Problem P_0.

The previous observations lead to a natural generalization of the ADP algorithm in [1] to this setting. Namely, we allow each user to iteratively update their beamforming vectors by solving Problem P_i and then announce new interference prices. Formally, this algorithm is described as follows:

1) Each user i chooses an initial beamforming vector $\mathbf{v}_i$ satisfying the power constraint.
2) Using (4), each user i calculates the interference price π_i given the current beamforming vectors and announces this price to every other user.
3) Periodically, each user i solves Problem P_i and updates his beamforming vector, given all the other users' interference prices $\{\pi_j\}_{j \neq i}$.
4) Go to step 2 and repeat.

We refer to this as the *MISO-ADP* algorithm. Steps 2 and 3 of the algorithm may be performed asynchronously among the users. Note that each user only announces a single interference price. In addition to these prices, user i's calculations in this algorithm only require knowledge of that user's SINR, that user's received signal power $|\mathbf{v}_i^\dagger \mathbf{h}_{ii}|^2$, and the channel gains $\mathbf{h}_{ij}$ for all j. In particular, there is no need to know the other users' beamforming vectors or the channel gains $\mathbf{h}_{kj}$ for $k \neq i$.

The MISO-ADP algorithm gives a distributed approach for adapting each user's beamforming vector with limited information exchange. However, there are two key questions. First, how difficult is it to solve the optimization problem in step 3? Second, does this algorithm converge? From our previous arguments, it follows that if the algorithm converges, then the limit point will satisfy the KKT conditions of Problem P_0. Furthermore, if there is a unique solution to the KKT conditions, then this point will be globally optimal. Next, we provide answers to these questions for a two-user interference channel.

III. TWO-USER MISO INTERFERENCE CHANNEL

In this section we focus on the case of $M = 2$ users. To begin, we show a key structural property of the solution to Problem P_0, which enables us to re-parameterize this problem in terms of a single "angle variable" for each user. We then re-cast the MISO-ADP algorithm in terms of this parametrization and answer the two questions raised in the previous section.

A. Angle Parametrization

Consider Problem P_0 with $M = 2$ users, so that the optimization variables are $\mathbf{v_1}$ and $\mathbf{v_2}$. The KKT conditions (3) for this case can be re-written as:

$$\left[\frac{u'_i}{n_0 + |\tilde{\mathbf{v}}_\mathbf{j}^\dagger \mathbf{h_{ji}}|^2} \left(\mathbf{h_{ii}} \mathbf{h_{ii}}^\dagger \right) - \frac{u'_j |\tilde{\mathbf{v}}_\mathbf{j}^\dagger \mathbf{h_{jj}}|^2}{(n_0 + |\tilde{\mathbf{v}}_\mathbf{i}^\dagger \mathbf{h_{ij}}|^2)^2} \left(\mathbf{h_{ij}} \mathbf{h_{ij}}^\dagger \right) \right] \tilde{\mathbf{v}}_\mathbf{i} = \lambda_i \tilde{\mathbf{v}}_\mathbf{i}, \tag{5}$$

for $i = 1, 2$ and $j \neq i$.[3] This is a nonlinear equation in the beamforming vectors. However, given any fixed $\tilde{\mathbf{v}}_\mathbf{1}$ and $\tilde{\mathbf{v}}_\mathbf{2}$ that satisfy (5), for each i, $\tilde{\mathbf{v}}_\mathbf{i}$ must be an eigenvector of the matrix

$$A_i = \left[\frac{u'_i}{n_0 + |\tilde{\mathbf{v}}_\mathbf{j}^\dagger \mathbf{h_{ji}}|^2} \left(\mathbf{h_{ii}} \mathbf{h_{ii}}^\dagger \right) - \frac{u'_j |\tilde{\mathbf{v}}_\mathbf{j}^\dagger \mathbf{h_{jj}}|^2}{(n_0 + |\tilde{\mathbf{v}}_\mathbf{i}^\dagger \mathbf{h_{ij}}|^2)^2} \left(\mathbf{h_{ij}} \mathbf{h_{ij}}^\dagger \right) \right]$$

with eigenvalue λ_i. It follows that $\tilde{\mathbf{v}}_\mathbf{i}$ must lie in the span of $\{\mathbf{h_{ii}}, \mathbf{h_{ij}}\}$.

If $\mathbf{h_{ij}}$ is orthogonal to $\mathbf{h_{ii}}$, it is easy to see that the optimal beamforming vector $\tilde{\mathbf{v}}_\mathbf{i}$ will be aligned with $\mathbf{h_{ii}}$ and consume all power, i.e., $\tilde{\mathbf{v}}_\mathbf{i} = c\mathbf{h_{ij}}$, where c is some real constant. If not, since $\mathbf{h_{ij}}$ can be expressed as a linear combination of $\mathbf{h_{ii}}$ and $P_{\mathbf{h_{ij}}}^\perp \mathbf{h_{ii}}$, where $P_\mathbf{x}^\perp \mathbf{y}$ is the orthogonal projection of vector $\mathbf{y}$ onto vector $\mathbf{x}$, and the subspace spanned by $\mathbf{h_{ii}}$ and $\mathbf{h_{ij}}$ is equivalent to the subspace spanned by $\mathbf{h_{ii}}$ and $P_{\mathbf{h_{ij}}}^\perp \mathbf{h_{ii}}$, then the optimal $\tilde{\mathbf{v}}_\mathbf{i}$, $i = 1, 2$, can be written as

$$\tilde{\mathbf{v}}_\mathbf{1} = c_{11} \mathbf{h_{11}} + c_{12} P_{\mathbf{h_{12}}}^\perp \mathbf{h_{11}} \tag{6}$$

$$\tilde{\mathbf{v}}_\mathbf{2} = c_{22} \mathbf{h_{22}} + c_{21} P_{\mathbf{h_{21}}}^\perp \mathbf{h_{22}} \tag{7}$$

where the c's are the combining coefficients. Taking $c_{ii} = 1$ and $c_{ij} = 0$, $i \neq j$, is conventional (channel-matched) beamforming, whereas taking $c_{ii} = 0$, $c_{ij} = 1$ is a zero-forcing approach, which cause no interference to the other user. Note that in general neither is globally optimal.

From now on, we will focus on a general case, in which $\mathbf{h_{ij}}$ is not orthogonal to $\mathbf{h_{ii}}$, i.e., $\mathbf{h_{ii}} \neq P_{\mathbf{h_{ij}}}^\perp \mathbf{h_{ii}}$. Later on, it is easy to verify that our results of parametrization can still applied to the orthogonal case.

Without loss of generality, we can assume that c_{12} and c_{21} are real and nonnegative, but that c_{11} and c_{22} are complex (for the time being). Substituting for the optimal $\tilde{\mathbf{v}}_\mathbf{i}$ in the objective of the original problem gives

$$\text{Total Utility} = u_1 \left(\frac{\left| c_{11}^* |\mathbf{h_{11}}|^2 + c_{12} \left(|\mathbf{h_{11}}|^2 - \frac{|\mathbf{h_{12}}^\dagger \mathbf{h_{11}}|^2}{|\mathbf{h_{12}}|^2} \right) \right|^2}{n_0 + |c_{22}|^2 |\mathbf{h_{22}}^\dagger \mathbf{h_{21}}|^2} \right)$$
$$+ u_2 \left(\frac{\left| c_{22}^* |\mathbf{h_{22}}|^2 + c_{21} \left(|\mathbf{h_{22}}|^2 - \frac{|\mathbf{h_{21}}^\dagger \mathbf{h_{22}}|^2}{|\mathbf{h_{21}}|^2} \right) \right|^2}{n_0 + |c_{11}|^2 |\mathbf{h_{11}}^\dagger \mathbf{h_{12}}|^2} \right). \tag{8}$$

Proposition 1: For each user $i = 1, 2$, if $\mathbf{h_{ii}}$ is not aligned with $\mathbf{h_{ij}}$, i.e., $\mathbf{h_{ii}} \neq a\mathbf{h_{ij}}$, for any real coefficient a, then the corresponding power constraint is binding at optimality, i.e., $|\tilde{\mathbf{v}}_\mathbf{i}|^2 = P_i^{max}$.

The proof follows from the observation that user i can increase its power without increasing interference to user j by adjusting c_{ij}.

Proposition 2: There exists an optimal beamforming vector for which the coefficients c_{11}, c_{12}, c_{22} and c_{21} are all real-valued and nonnegative. In another word, $\tilde{\mathbf{v}}_\mathbf{i}$ is in the convex cone spanned by $\mathbf{h_{ii}}$ and $P_{\mathbf{h_{ij}}}^\perp \mathbf{h_{ii}}$, shown in Fig. 1.

The proof is based on checking the KKT conditions for the optimal beamformer. Namely, we first show that the optimal beamformer for user i, which maximizes his signal power given that the interference power to user j is fixed, can be achieved with real c_{ii} and c_{ij}. Then we show that these coefficients are

---

[3]In the remainder of this section, we will follow the convention that index j refers to the interfering user for user i (i.e. $j \neq i$).
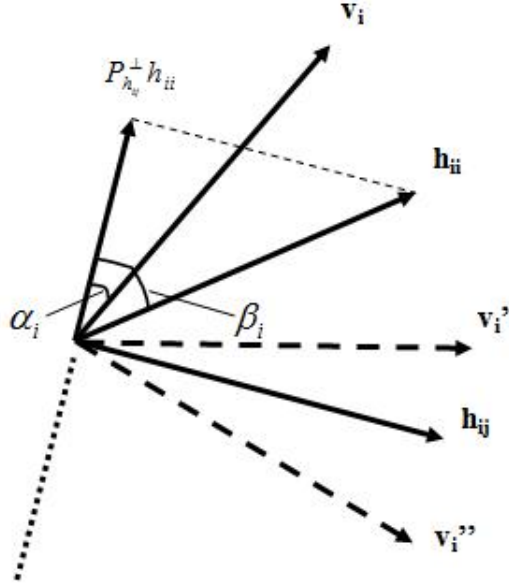
Fig. 1. Illustration of the vectors $\mathbf{h_{ii}}$, $\mathbf{h_{ij}}$, $P_{\mathbf{h_{ij}}}^{\perp}\mathbf{h_{ii}}$, $\mathbf{v_i}$, $\mathbf{v_i'}$, and $\mathbf{v_i}''$ and the angles α_i and β_i.

nonnegative by excluding the possibility that the optimal $\tilde{\mathbf{v}}_i$ is not in the convex cone spanned by $\mathbf{h_{ii}}$ and $P_{\mathbf{h_{ij}}}^{\perp}\mathbf{h_{ii}}$.

According to Proposition 2, the original optimization problem is equivalent to finding the optimal nonnegative and real coefficients c_{11}, c_{12}, c_{22} and c_{21} that maximize the total utility. Propositions 1 and 2 imply that we can characterize $\tilde{\mathbf{v}}_i$ with a single angle α_i. This angle is the same as that in real space, i.e., we view an N-dimensional complex vector as a $2N$-dimensional real vector. If we define the angle between two complex vectors $\mathbf{x}$ and $\mathbf{y}$ as α, denoted by $\alpha \equiv \mathbf{x} \wedge \mathbf{y}$, then we have

$$\cos \alpha = \frac{Re\{\mathbf{x}^{\dagger}\mathbf{y}\}}{|\mathbf{x}||\mathbf{y}|}. \tag{9}$$

Using this definition, we define α_i as the angle between the vector $\tilde{\mathbf{v}}_i$ and $P_{\mathbf{h_{ij}}}^{\perp}\mathbf{h_{ii}}$. Since the norm of $\tilde{\mathbf{v}}_i$ is fixed, and $\tilde{\mathbf{v}}_i$ is between $\mathbf{h_{ii}}$ and $P_{\mathbf{h_{ij}}}^{\perp}\mathbf{h_{ii}}$, we have $\alpha_i \in [0, \beta_i]$, where β_i is the angle between $\mathbf{h_{ii}}$ and $P_{\mathbf{h_{ij}}}^{\perp}\mathbf{h_{ii}}$. This is illustrated in Fig. 1.

We have

$$\begin{cases} \alpha_i &= \mathbf{v_i} \wedge P_{\mathbf{h_{ij}}}^{\perp}\mathbf{h_{ii}} \\ \beta_i - \alpha_i &= \mathbf{v_i} \wedge \mathbf{h_{ii}} \\ \beta_i &= \mathbf{h_{ii}} \wedge P_{\mathbf{h_{ij}}}^{\perp}\mathbf{h_{ii}} \end{cases}$$

and can therefore write

$$c_{ii} = \frac{1}{\sqrt{|\mathbf{h_{ii}}|^2 - |P_{\mathbf{h_{ij}}}^{\perp}\mathbf{h_{ii}}|^2}} \sin \alpha_i \tag{10}$$

$$c_{ij} = \frac{1}{|P_{\mathbf{h_{ij}}}^{\perp}\mathbf{h_{ii}}|} \cos \alpha_i - \frac{1}{\sqrt{|\mathbf{h_{ii}}|^2 - |P_{\mathbf{h_{ij}}}^{\perp}\mathbf{h_{ii}}|^2}} \sin \alpha_i. \tag{11}$$

Substituting for the coefficients c's in (8) gives[4]

$$\text{Total Utility} = u_1\left(\frac{|\mathbf{h_{11}}|^2 \cos^2(\beta_1 - \alpha_1)}{n_0 + |\mathbf{h_{21}}|^2 \sin^2 \alpha_2}\right) + u_2\left(\frac{|\mathbf{h_{22}}|^2 \cos^2(\beta_2 - \alpha_2)}{n_0 + |\mathbf{h_{12}}|^2 \sin^2 \alpha_1}\right) \tag{12}$$

and the maximization is over α_1 and α_2.

B. Parameterized Asynchronous Distributed Algorithm

With the previous parametrization, we restate the utility maximization problem as

$$\max_{\alpha_1,\alpha_2} \quad u_1\left(\frac{|\mathbf{h_{11}}|^2 \cos^2(\beta_1 - \alpha_1)}{n_0 + |\mathbf{h_{21}}|^2 \sin^2 \alpha_2}\right) + u_2\left(\frac{|\mathbf{h_{22}}|^2 \cos^2(\beta_2 - \alpha_2)}{n_0 + |\mathbf{h_{12}}|^2 \sin^2 \alpha_1}\right) \tag{P_a}$$
$$\text{s.t.} \quad \alpha_i \in [0, \beta_i] \quad i = 1, 2$$

where $\beta_i = \mathbf{h_{ii}} \wedge P_{\mathbf{h_{ij}}}^{\perp} \mathbf{h_{ii}}$. We can now apply the distributed algorithm proposed in Section I by replacing variables $\mathbf{v_i}$'s with α_i. The resulting subproblem for each user is

$$\max_{\alpha_i} \quad u_i\left(\frac{|\mathbf{h_{ii}}|^2 \cos^2(\beta_i - \alpha_i)}{n_0 + |\mathbf{h_{ji}}|^2 \sin^2 \alpha_j}\right) - \pi_j |\mathbf{h_{ij}}|^2 \sin^2 \alpha_i \tag{P_{ai}}$$
$$\text{s.t.} \quad \alpha_i \in [0, \beta_i].$$

The objective function can be interpreted as a payoff function in a non-cooperative game,

$$s_i(\alpha_i; \alpha_j, \pi_j) = u_i\left(\frac{|\mathbf{h_{ii}}|^2 \cos^2(\beta_i - \alpha_i)}{n_0 + |\mathbf{h_{ji}}|^2 \sin^2 \alpha_j}\right) - \pi_j |\mathbf{h_{ij}}|^2 \sin^2 \alpha_i, \tag{13}$$

where π_j is the interference price announced by user j given by

$$\pi_j = -\frac{\partial u_j}{\partial I_j} = u'_j \frac{|\mathbf{h_{jj}}|^2 \cos^2(\beta_j - \alpha_j)}{(n_0 + |\mathbf{h_{ij}}|^2 \sin^2 \alpha_i)^2}. \tag{14}$$

The parameterized asynchronous distributed pricing algorithm is described as follows:

1) Each user i (1 or 2) chooses an initial angle $\alpha_i^0 \in [0, \beta_i]$ and an initial interference price $\pi_i^0 \in [0, \max\{u'_i\}\frac{|\mathbf{h_{ii}}|^2}{n_0^2}]$.

2) At each time n, one user i is randomly selected to maximize its payoff function $s_i(\alpha_i)$ and update its angle α_i, given the other user's angle α_j and prices π_j, $j \neq i$, i.e.,

$$\alpha_i(n+1) = \underset{\alpha_i \in [0,\beta_i]}{\operatorname{argmax}} \ s_i(\alpha_i; \alpha_j(n), \pi_j(n)). \tag{15}$$

3) Each user i calculates the new interference price π_i given the current angles $\alpha_{1,2}$ and announces it to the other user.

4) Repeat from step 2.

As in [1], we can view the distributed algorithm as a noncooperative game, denoted by G, in which there are four players corresponding to $\{\alpha_1, \alpha_2, \pi_1, \pi_2\}$, and each of them maximizes its own payoff function. The payoff function for α_i is given in (13), while the payoff function for π_i is

$$s_{\pi_i} = -\left(\pi_i - u'_i \frac{|\mathbf{h_{ii}}|^2 \cos^2(\beta_i - \alpha_i)}{(n_0 + |\mathbf{h_{ji}}|^2 \sin^2 \alpha_j)^2}\right)^2 \tag{16}$$

in order that interference price is always updated according to (14).

Proposition 3: The game G can be viewed as a supermodular game with transformed strategies if $-\frac{u''_i \gamma_i}{u'_i} \in (0, 1]$, or $-\frac{u''_i \gamma_i}{u'_i} \in [1, 2]$, for both $i = 1, 2$, where γ_i is the received SINR of user i.

---

[4]Here, for simplicity, we assume unit power constraints. You can also view it as P_i^{max} is absorbed by channel vector $\mathbf{h}$.

Proof: In order to show a game is supermodular, we need to verify: a) the strategy space for each player is a nonempty and compact sublattice; b) each payoff function is continuous in all players' strategies; c) each player's payoff function is supermodular in his own strategy; and d) each player's payoff function has increasing differences between any component of his own strategy and any component of any other's strategy. For the game G, conditions a) - c) are trivial because each player's strategy space is one-dimensional. Condition d) does not hold with the original definition of strategies. However, it can be checked that if $-\frac{u_i'' \gamma_i}{u_i'} \in (0, 1]$, then G is supermodular in the transformed strategies $(\alpha_1, -\alpha_2, \pi_1, -\pi_2)$; if $-\frac{u_i'' \gamma_i}{u_i'} \in [1, 2]$, then G is supermodular in the transformed strategies $(\alpha_1, \alpha_2, -\pi_1, -\pi_2)$. $\blacksquare$

From Theorem 1 in [1], we conclude that the distributed pricing algorithm converges under the following conditions.

Corollary 4: For a two-user MISO network, if both users' utility functions satisfy $-\frac{u_i'' \gamma_i}{u_i'} \in (0, 1]$ (or $-\frac{u_i'' \gamma_i}{u_i'} \in [1, 2]$) and the distributed algorithm is initialized with the smallest (or largest) element of its transformed strategy space, then the strategies monotonically converge to the component-wise smallest (largest) Nash Equilibrium (NE), which corresponds to a solution to the KKT condition for the original problem.

Specifically, if $-\frac{u_i'' \gamma_i}{u_i'} \in (0, 1]$, the smallest element in the strategy space is $(\alpha_1 = 0, \alpha_2 = \beta_2, \pi_1 = 0, \pi_2 = \max\{u_2'\}\frac{|\mathbf{h_{22}}|^2}{n_0^2})$, while the largest element is $(\alpha_1 = \beta_1, \alpha_2 = 0, \pi_1 = \max\{u_1'\}\frac{|\mathbf{h_{11}}|^2}{n_0^2}, \pi_2 = 0)$. The initial strategies are similar when $-\frac{u_i'' \gamma_i}{u_i'} \in [1, 2]$. Since the theorem in [1] only shows that the strategies are eventually bounded component-wise by the smallest and largest NE, convergence is not guaranteed if each user starts from an arbitrary strategy (selection of beamformers). In other words, without excluding the possibility that there are multiple NE's, to guarantee convergence we must carefully select the initial strategy.

Now, we will show how to solve the subproblem P_{ai} efficiently, i.e., optimize the payoff function for α_i. Taking the derivative of s_i in (13) over α_i, we obtain

$$\frac{\partial s_i(\alpha_i)}{\partial \alpha_i} = u_i' \frac{2|\mathbf{h_{ii}}|^2 \cos(\beta_i - \alpha_i)\sin(\beta_i - \alpha_i)}{n_0 + |\mathbf{h_{ji}}|^2 \sin^2 \alpha_j} - \pi_j |\mathbf{h_{ij}}|^2 \sin 2\alpha_i$$
$$= 2u_i' \gamma_i \tan(\beta_i - \alpha_i) - \pi_j |\mathbf{h_{ij}}|^2 \sin 2\alpha_i$$

where u_i' and γ_i (SINR) are both functions of α_i. It is easy to check that the derivative is positive at $\alpha_i = 0$ and negative at $\alpha_i = \beta_i$. Therefore, the optimal solution is not binding. Therefore, we only need to solve the following equation

$$2u_i' \gamma_i \tan(\beta_i - \alpha_i) = \pi_j |\mathbf{h_{ij}}|^2 \sin 2\alpha_i \tag{17}$$

which is easier than solving subproblem P_i with an N-dimensional variable, although no closed-form solution is available.

Furthermore, if $-\frac{u_i'' \gamma_i}{u_i'} > 1$, i.e., $\frac{\partial u_i' \gamma_i}{\partial \gamma_i} = u_i'' \gamma_i + u_i' < 0$, then we know $u_i' \gamma_i$ is monotonically decreasing as γ_i increases, which implies that $u_i' \gamma_i$ decreases as α_i increases. Therefore, $2u_i' \gamma_i \tan(\beta_i - \alpha_i)$ is a monotonically decreasing function. Then, it is obvious that the solution to (17) is unique and the payoff function $s_i(\alpha_i)$ is quasi-concave. The utility function satisfying $-\frac{u_i'' \gamma_i}{u_i'} \in [1, 2]$ fits into this case.

For a two-user system, the problem can be further simplified if the utility function has the form of $\theta_i \log(\gamma_i)$. Then,

$$\text{Total Utility} = \log\left(\frac{[|\mathbf{h_{11}}|^2 \cos^2(\beta_1 - \alpha_1)]^{\theta_1}}{[n_0 + |\mathbf{h_{12}}|^2 \sin^2 \alpha_1]^{\theta_2}}\right) + \log\left(\frac{[|\mathbf{h_{22}}|^2 \cos^2(\beta_2 - \alpha_2)]^{\theta_2}}{[n_0 + |\mathbf{h_{21}}|^2 \sin^2 \alpha_2]^{\theta_1}}\right).$$

After reorganizing, the terms containing α_1 and α_2 separate. Therefore each user only needs to optimize the corresponding term regardless of the other's beamformer, which means no iteration is needed. The
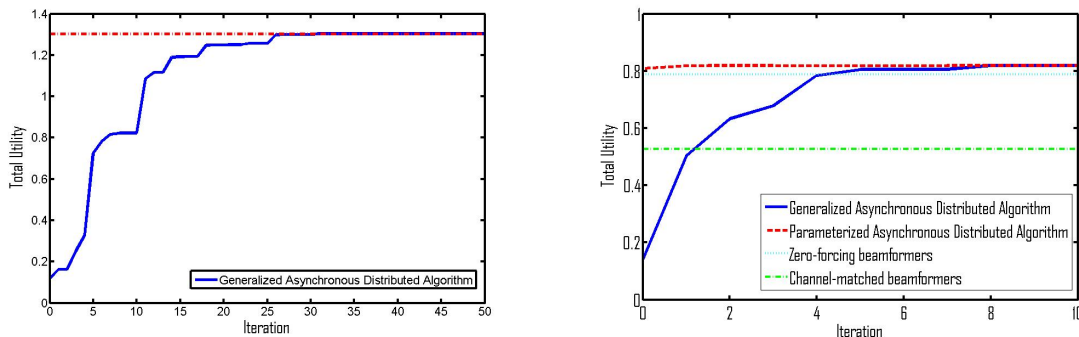
(a) The generalized asynchronous distributed algorithm. (b) The parameterized asynchronous distributed algorithm.

Fig. 2. Illustration of the performance of two distributed algorithms.

optimal beamformer for user i in this case is the solution to

$$2\theta_i \tan(\beta_i - \alpha_i) = \theta_j \frac{|\mathbf{h_{ij}}|^2 \sin 2\alpha_i}{n_0 + |\mathbf{h_{ij}}|^2 \sin^2 \alpha_i}. \tag{18}$$

The angle parametrization for the two-user MISO system therefore enables a distributed algorithm, which uniquely determines the optimal beamforming vector for each user. Furthermore, the computation in every iteration becomes relatively simple. In the following section, we compare the original (N-dimensional) and parameterized asynchronous distributed algorithms numerically in terms of convergence rate.

IV. SIMULATION RESULTS

In this section, we show some typical performance plots for both the generalized asynchronous distributed algorithm in Section II and the parameterized asynchronous distributed algorithm in Section III.

A. Generalized Asynchronous Distributed Algorithm

In the system there are 5 pairs of transmitters and receivers with 4 transmit antennas each. The users are randomly placed within a square of 1 km × 1 km. Each entry of the channel vector is assumed to be an *iid* complex Gaussian random variable, where the variance is determined by the distance attenuation. Specifically, the variance of both real and imaginary parts is $\sigma^2(d) = \sigma_0^2(\frac{d}{100})^{-4}$, where σ_0^2 is the reference variance at a distance of 100 m, which is the minimum separation between any transmitter and receiver, and d is the separation in meters. We also assume the normalized maximum transmit power for each user is one, $\sigma_0^2 = 10^{-7}$, and the variance of noise is 10^{-9} so that the average received SNR at 100 m is 200 (about 23 dB). For these results the utility function is $u(\gamma) = \log(1 + \gamma)$.

Although in Section II, we are not able to show the convergence of the generalized asynchronous distributed algorithm, it is observed to converge numerically in all cases simulated. Fig. 2(a) shows total utility versus number of iterations for a particular model realization. The algorithm starts from a random selection of beamformers, and converges to a stationary point, which is an NE of the corresponding noncooperative game. To check whether the limit point is indeed optimal, we used MATLAB to solve the global optimization problem starting from the limit point achieved by the algorithm. The total utility obtained in this way is indicated by the red dash-dot line in Fig. 2(a), which matches the limit point. Hence we conclude that the generalized asynchronous distributed algorithm does converge to a local optimum.

B. Parameterized Asynchronous Distributed Algorithm

Here, we consider a two-user MISO system with the same setting as in the previous subsection. Applying the parameterized asynchronous distributed algorithm proposed in Section III, we optimize the total utility iteratively. The numerical results show that the algorithm still converges even if we start from any arbitrary angle satisfying the constraints. As in Fig. 2(a), Fig. 2(b) also shows total utility versus number of iterations for a particular model realization, which illustrates the convergence of the algorithm. Furthermore, we compare it with the generalized asynchronous distributed algorithm. Both algorithms start from random beamformers or angles[5]. It is easy to see that the parameterized asynchronous distributed algorithm yields a higher initial total utility, and it converges more rapidly. The main reason is because we narrow the domain of potential beamformers by angle parametrization. The preceding observation is quite typical after trying many model realizations. When averaged over 100 model realizations, the results show that it takes 4.28 iterations for the parameterized asynchronous distributed algorithm to converge[6], while for the generalized asynchronous distributed algorithm, it takes 6.0 iterations for convergence. In Fig. 2(b), we also show the total utility with the (suboptimal) zero-forcing and channel-matched filters.

V. Conclusions

We have presented two distributed algorithms for selecting each user's beamforming vector to maximize the total utility. The generalized asynchronous distributed algorithm is an extension of the asynchronous distributed pricing algorithm in [1]. (See also [3].) In that algorithm, each user announces an interference price, which represents his current marginal cost per unit interference power. After exchanging interference prices, users update their beamforming vectors by maximizing a local payoff function. Although we have not proved that this algorithm converges, numerical results indicate that it performs quite well. For a two-user system we have presented an asynchronous distributed algorithm, based on an angle parametrization, which we have shown must converge, and which is computationally simpler than the first (generalized) algorithm.

Unfortunately, the angle parametrization we apply in Section III appears to be difficult to generalize to multi-user systems. An open issue, then, is how to prove the convergence of the generalized asynchronous distributed algorithm for multi-user systems. Furthermore, it still unknown if the optimum is unique. Extensions to MIMO channels are also interesting for future work.

References

[1] J. Huang, R. A. Berry, and M. L. Honig, "Distributed interference compensation for wireless networks," *IEEE Journal on Selected Areas in Communications*, vol.24, no. 5, pp.1074-1084, May 2006

[2] C. Shi, R. A. Berry, and M. L. Honig, "Distributed Interference Pricing for OFDM Wireless Networks with Non-Separable Utilities," in *Proc. Conference on Information Sciences and Systems*, March 2008.

[3] D. A. Schmidt, A. Grundinger, W. Utschick, and M. L. Honig, "Distributed precoder optimization for interfering MISO channels," in *International ITG Workshop on Smart Antennas*, Feb. 2008, pp. 285-290.

[4] M. Chiang, "Balancing Transport and Physical Layers in Wireless Multihop Networks: Jointly Optimal Congestion Control and Power Control," *IEEE J. Sel. Areas Comm*, vol. 23, no. 1, pp. 104-116, Jan. 2005.

[5] J. Yuan and W. Yu, "Distributed cross-layer optimization of wireless sensor networks: A game theoretic approach," in *Proc. IEEE Conf. Global Commun.*, Nov. 2006.

[6] M. F. Demirkol and M. A. Ingram, "Power-controlled capacity for interfering MIMO links," in *Proc. IEEE Vehicular Technology Conference,* vol. 1, Oct. 2001.

[7] S. Ye and R. S. Blum, "Optimized signalling for MIMO interference systems with feedback," *IEEE Trans. Signal Processing,* vol. 51, no. 11, 2003.

[8] G. Scutari, D. P. Palomar and S. Barbarossa, "Competitive Design of Multiuser MIMO Systems based on Game Theory: A Unified View," to appear in *IEEE Journal on Selected Areas in Communications: Special Issue on Game Theory*, Sept. 2008.

---

[5]Since an arbitrary beamforming vector $\mathbf{v_i}$ may not lie in the convex cone spanned by $\mathbf{h_{ii}}$ and $\mathbf{h_{ii}}$, we cannot ensure that the two algorithms start with the same beamformers.

[6]Numerically, we define convergence as the minimum number of iterations for which the difference in total utility between two consecutive iterations is no greater than some threshold (0.1% in our simulation).

[9] E. A. Jorswieck and E. G. Larsson, "The MISO interference channel from a game-theoretic perspective: A combination of selfishness and altruism achieves pareto optimality," in *Proc. ICASSP*, 2008, pp. 5364-5367.

[10] D. Bertsekas, *Nonlinear Programming*, 2nd ed. Belmont, MA: Athena Scientific, 1999.

APPENDIX A
PROOF OF PROPOSITION 1

Proof: Suppose that the optimal $|\tilde{\mathbf{v}}_i|^2 < P_i^{max}$ for some i. Then fixing the corresponding coefficient c_{ii}, which means the interference power to the other user will be fixed, we can always increase user i's signal power by adjusting c_{ij}. That is, if $Re\{c_{ii}^*|\mathbf{h}_{11}|^2 + c_{ij}\left(|\mathbf{h}_{ii}|^2 - \frac{|\mathbf{h}_{ij}^\dagger \mathbf{h}_{ii}|^2}{|\mathbf{h}_{ij}|^2}\right)\} \geq 0$, increase c_{ij} within the power constraint; otherwise , decrease c_{ij}. Note that the assumption that $\mathbf{h}_{ii}$ is not aligned with $\mathbf{h}_{ij}$ implies $|\mathbf{h}_{ii}|^2 - \frac{|\mathbf{h}_{ij}^\dagger \mathbf{h}_{ii}|^2}{|\mathbf{h}_{ij}|^2} > 0$, i.e., c_{ij} has a non-trivial form. ∎

APPENDIX B
PROOF OF PROPOSITION 2

Proof: Consider the optimal beamforming vector $\tilde{\mathbf{v}}_i$ for user i. The result is trivial if $\mathbf{h}_{ii}$ is aligned with $\mathbf{h}_{ij}$ or it is the case that either $c_{ii} = 0$ or $c_{ij} = 0$ at optimality. Then suppose that $|c_{ii}| > 0$ and $c_{ij} > 0$, and write $c_{ii} = c_{ii}' e^{j\theta_{ii}}$, where c_{ii}' is real and positive.

If we fix c_{ii}' and optimize over θ_{ii} and c_{ij}, then we claim that the optimal θ_{ii} is 0 or π. This is because when c_{ii}' is fixed, the interference power to user j is $|c_{ii}|^2|\mathbf{h}_{ii}^\dagger \mathbf{h}_{ij}|^2$, which is fixed, and θ_{ii} and c_{ij} can only influence the signal power of user i. Therefore this problem reduces to:

$$\max_{c_{ij},\, \theta_{ii}} |\mathbf{v}_i^\dagger \mathbf{h}_{ii}|^2$$
$$\text{s.t.} \qquad |\mathbf{v}_i|^2 = P_i^{max}$$

which can be further simplified as

$$\max_{c_{ij},\, \theta_{ii}} c_{ij}^2 \frac{|P_{\mathbf{h}_{ij}}^\perp \mathbf{h}_{ii}|^2}{|\mathbf{h}_{ii}|^2} + 2c_{ii}' c_{ij} \cos\theta_{ii}$$
$$\text{s.t.} \qquad c_{ij}^2 + 2c_{ii}' c_{ij} \cos\theta_{ii} = constant.$$

The KKT conditions are

$$\begin{cases} 2c_{ij}\frac{|P_{\mathbf{h}_{ij}}^\perp \mathbf{h}_{ii}|^2}{|\mathbf{h}_{ii}|^2} + 2c_{ii}'\cos\theta_{ii} + \lambda_i(2c_{ij} + 2c_{ii}'\cos\theta_{ii}) = 0 \\ -2c_{ii}'c_{ij}(1+\lambda_i)\sin\theta_{ii} = 0 \\ c_{ij}^2 + 2c_{ii}'c_{ij}\cos\theta_{ii} = constant \end{cases}$$

where λ_i is the Lagrange multiplier, and the unique solution is $\sin\theta_{ii} = 0$, i.e, $\theta_{ii} = 0$ or π. Therefore c_{ii} is real.

Now, we can assume all these coefficients are real. Then we want to exclude the possibility that the optimal $\tilde{\mathbf{v}}_i$ is not in the convex cone spanned by $\mathbf{h}_{ii}$ and $P_{\mathbf{h}_{ij}}^\perp \mathbf{h}_{ii}$. Without loss of generality, we can assume $c_{ii} > 0$ and only consider the right half-plane in Fig. 1. Since $|\mathbf{v}_i|^2 = P_i^{max}$, we know the angle between $\mathbf{v}_i$ and $\mathbf{h}_{ii}$ will uniquely decide its own signal power, while the angle between $\mathbf{v}_i$ and $\mathbf{h}_{ij}$ will decide the interference power to user j. It can be shown that if $\mathbf{v}_i$ is out of the cone (e.g., $\mathbf{v}_i'$ or $\mathbf{v}_i''$ in Fig. 1) it cannot be a candidate of the optimal $\tilde{\mathbf{v}}_i$ because we can always find another $\mathbf{v}_i$ which will yield either a higher signal power or a lower interference power. ∎