

Price-Based Joint Beamforming and Spectrum Management in Multi-Antenna Cognitive Radio Networks

Diep N. Nguyen and Marwan Krunz

Abstract—We consider the problem of maximizing the throughput of a multi-antenna cognitive radio (CR) network. With spatial multiplexing over each frequency band, a multi-antenna CR node controls its antenna radiation directions and allocates power for each data stream by appropriately adjusting its *precoding matrix*. Our objective is to design a set of precoding matrices (one per band) at each CR node so that power and spectrum are optimally allocated for the node and its interference is steered away from unintended receivers. The problem is non-convex, with the number of variables growing quadratically with the number of antenna elements. To tackle it, we translate it into a noncooperative game. We derive an optimal pricing policy for each node, which adapts to the node's neighboring conditions and drives the game to a Nash-Equilibrium (NE). The network throughput under this NE equals to that of a locally optimal solution of the non-convex centralized problem. To find the set of precoding matrices at each node (best response), we develop a low-complexity distributed algorithm by exploiting the strong duality of the convex per-user optimization problem. The number of variables in the distributed algorithm is independent of the number of antenna elements. A centralized (cooperative) algorithm is also developed. Simulations show that the network throughput under the distributed algorithm rapidly converges to that of the centralized one. Finally, we develop a MAC protocol that implements our resource allocation and beamforming scheme. Extensive simulations show that the proposed protocol dramatically improves the network throughput and reduces power consumption.

Index Terms—Noncooperative game, pricing, cognitive radio, MIMO, power allocation, frequency management, beamforming.

I. INTRODUCTION

RECENT years have witnessed a great interest in cognitive radio (CR) and multi-input multi-output (MIMO) technologies. Through spectrum sensing, CRs can opportunistically communicate on temporarily idle frequency bands while avoiding interference with licensed primary users (PUs). MIMO communications improve the link throughput by sending independent data streams simultaneously over different antennas (a.k.a. *spatial multiplexing*).

A crucial challenge in CR networks (CRNs) is how to effectively allocate transmission powers and spectrum among CRs (see Fig. 1(a)) so as to maximize the network throughput while protecting primary users (PUs) from CR interference.

Manuscript received 15 December 2011; revised 1 June 2012. This research was supported in part by NSF (under grants CNS-1016943, CNS-0904681, IIP-1231043, and IIP-0832238), Raytheon, and the Connection One center. Any opinions, findings, conclusions, or recommendations expressed in this paper are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

The authors are with Department of Electrical and Computer Engineering, University of Arizona (e-mail: {dnnguyen, krunz}@email.arizona.edu).

Digital Object Identifier 10.1109/JSAC.2012.121221.

Even for a single channel and single-antenna wireless devices, the problem is difficult due to the non-convexity of the network throughput function.

The incorporation of MIMO techniques into CR systems introduces two new control dimensions, besides power control and frequency management: power allocation over antennas (space dimension) and interference management. The latter comes from MIMO's *degrees of freedom* [1], which allow a MIMO node to suppress interference from others and beamform its antenna patterns to keep interference away from unintended receivers. MIMO's power allocation and interference management can be jointly controlled via *precoding matrices*, a spatial multiplexing technique [1]. Previous works (e.g., [2] [3] [4]) considered power allocation or *stream control* (see Fig. 1(b)), but did not take into account interference management via controlling the antenna beams. An optimal set of precoding matrices for each node allocates power over *both* space and frequency dimensions (Fig. 1(c)) and yields radiation patterns that induce minimum interference (Fig. 1(d)), so as to maximize network throughput. This problem is the focus of our work.

II. RELATED WORKS

Ignoring the need to protect PUs, the integration of MIMO into CRNs very much resembles multi-carrier (e.g., OFDM) MIMO (MC-MIMO) systems. In MC-MIMO, joint power and spectrum optimization is a non-convex problem, which was recently shown to be NP-hard [5], i.e., its complexity grows exponentially with the number of variables. Unfortunately, the number of variables in a MC-MIMO network can be very large. For instance, in a network of 10 links, 4 antennas per node, and 10 sub-carriers, the problem involves $4 \times 4 \times 10 \times 10 = 1600$ complex variables (or 3200 real variables).

Existing works on MIMO CR systems (e.g., [4] [6] [7] [8] [9]) generally overlook the optimization over the frequency dimension. Extending these works to multi-band MIMO CRNs is not trivial. First, scalar-value algorithms used for a single-band MIMO ad hoc network (e.g., bisection search in [9]) do not work when searching for optimal vectors in multi-band MIMO CRNs. Second, as shown in this paper, even without beamforming the price-based optimal power allocation over both frequency and spatial dimensions is not equivalent to a general water filling problem (with multiple water levels) [10]. Hence, existing algorithms for MIMO (e.g., [10] [11] [12] [13]) and SISO systems (e.g., [14]) are not applicable. Third, applying single-band MIMO techniques to each individual band of a multi-band MIMO CR system often leads to poor performance, as shown later in this paper.

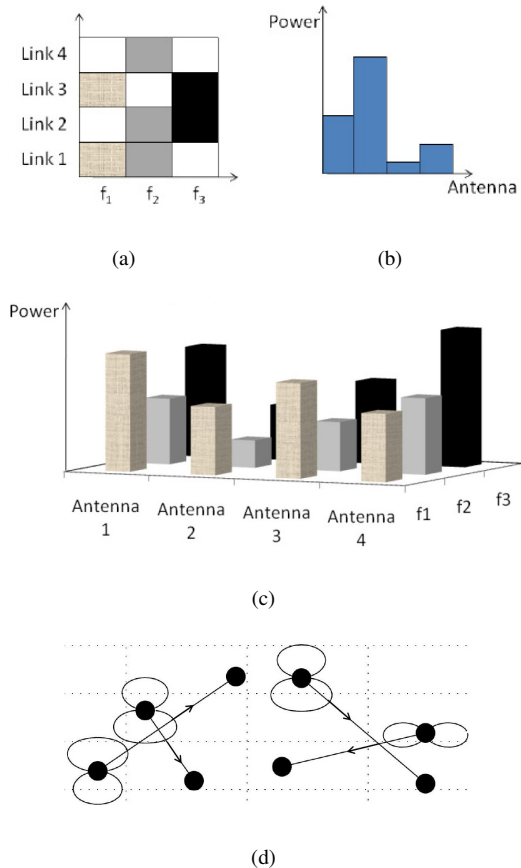


Fig. 1. Power allocation (a) in frequency, (b) in space, (c) in both dimensions, and (d) four transmit radiation patterns steering away from nearby unintended receivers.

In [4] [8], a single-frequency cognitive MIMO network was formulated as a noncooperative game. In these works, the spectrum management and beamforming capability of CR MIMO transmitters were not taken into account. Moreover, pricing [15] was not used in [4] [8], hence a node maximizes its throughput in a greedy manner. Simulations show that our pricing technique greatly outperforms greedy mechanisms in terms of network throughput and power efficiency. Relying on variational inequality theory, matrix projection, and fixed-point theorem, the authors in [4] [8] require the channel gain matrices among CR nodes to meet certain conditions so that the existence and convergence of the NE are guaranteed. In other words, the NE existence and the convergence property of their distributed algorithm depend on given channel conditions that are not always met.

Motivated by the above, this paper develops a low-complexity distributed algorithm that configures the transmit antenna radiation directions and allocates powers for all data streams so as to maximize the network throughput. Our main contributions are as follows. First, we formulate the joint power, spectrum allocation, and beamforming problem as a noncooperative game [16]. We prove that the game always admits at least one NE and we provide conditions for the uniqueness of the NE.

Second, to improve the NE efficiency, we derive user-dependent pricing policies that drive the game to a NE whose network throughput is the same as that of a locally

optimal point of the nonconvex network-wide problem. By capturing the interference from a transmitter to its unintended receivers, the pricing policy guides the MIMO transmitters to steer their beams away from nearby unintended receivers. Via simulations, we observe that the proposed pricing policies dramatically improve the network throughput over the greedy approach (which does not use pricing). Our approach is also more power-efficient than the greedy approach where users use all of their power budgets to greedily maximize their individual throughput.

Third, exploiting the strong duality in convex optimization, we design a low-complexity distributed algorithm that allows a node to compute its set of precoders (*best response*) in constant time w.r.t. the antenna array size. We also develop a centralized algorithm for the network optimization problem, where nodes are assumed to work in a cooperative way (cooperative game). Simulations show that the performance of the (noncooperative) distributed algorithm is almost the same as that of the (cooperative) centralized one.

Forth, we design full/generalized eigen MIMO precoding for multi-channel systems. This differs from a large body of works on MIMO precoder design (e.g., [6] [17] [18]), where only one data stream is sent from a MIMO transmitter on a single channel. In these works, precoders have a rank of one and reduce to vectors. In generalized eigencoding, there is no constraint on the rank of the precoding matrices [19], i.e., several data streams can be sent simultaneously. Inspired by the introduction of MIMO spatial multiplexing into existing networks (e.g., 802.11n allows up to four concurrent multiplexed streams), generalized eigencoding has recently attracted great interest.

In Section III, we present the network model and the problem formulation. The noncooperative game analysis, optimal pricing policy, convergence proof, the distributed algorithm, and a corresponding MAC protocol are given in Section IV. The centralized algorithm is developed in Section V. Numerical results are discussed in Section VI. Concluding remarks and future work are provided in Section VII. Throughout the paper, we use $(\cdot)^*$ to denote the conjugate matrix, $(\cdot)^H$ to denote the Hermitian matrix transpose, $\text{tr}(\cdot)$ for the trace of a matrix, $\det(\cdot)$ for the determinant, and $(\cdot)^T$ for matrix transpose. Matrices and vectors are bold-faced.

III. PROBLEM FORMULATION

We consider a CRN that coexists with several primary networks. The CRN consists of N CR links. Each CR node is equipped with M antennas. The available spectrum consists of K orthogonal frequency channels with central frequencies f_1, f_2, \dots, f_K . For simplicity, we also use f_k to refer to the k th channel. Let $\Phi_N \stackrel{\text{def}}{=} \{1, 2, \dots, N\}$ and $\Psi_K \stackrel{\text{def}}{=} \{1, 2, \dots, K\}$ denote the sets of CR links and channels, respectively. Each CR user i can simultaneously communicate over multiple channels. We impose a half-duplex constraint on all transmissions.

On a given channel, a CR transmitter can send up to M independent data streams on its M antennas. Formally, for channel f_k let $\mathbf{x}_i^{(k)}$ be a column vector of M information symbols, sent from node i to its destination node $d(i)$. Each element of $\mathbf{x}_i^{(k)}$ is from one data stream. Let $\tilde{\mathbf{T}}_i^{(k)} \in \mathbb{C}^{M \times M}$ denote the complex precoding matrix of node i on channel f_k . Then, the transmit vector is $\tilde{\mathbf{T}}_i^{(k)} \mathbf{x}_i^{(k)}$. We assume spectrum

sharing among different CRs. Specifically, for channel f_k , the received signal vector $\mathbf{y}_{d(i)}^{(k)}$ at receiver $d(i)$ of link $(i, d(i))$ is given by:

$$\mathbf{y}_{d(i)}^{(k)} = \mathbf{H}_{d(i),i}^{(k)} \tilde{\mathbf{T}}_i^{(k)} \mathbf{x}_i^{(k)} + \sum_{j \in \{\Phi_N \setminus i\}} \mathbf{H}_{d(i),j}^{(k)} \tilde{\mathbf{T}}_j^{(k)} \mathbf{x}_j^{(k)} + \mathbf{N}_k \quad (1)$$

where the first term in the RHS of (1) is the desired signal sent from transmitter i . $\mathbf{H}_{d(i),i}^{(k)}$ is an $M \times M$ channel gain matrix on channel f_k from the transmitter i to $d(i)$. Each element of $\mathbf{H}_{d(i),i}^{(k)}$ is a multiplication of a distance- and channel-dependent attenuation term and a random term that reflects multi-path fading (complex Gaussian variables with zero mean and unit variance). We assume a flat-fading channel. The second term in (1) represents interference from other CR links that link $(i, d(i))$ shares the channel f_k with. $\mathbf{N}_k \in \mathbb{C}^M$ is an $M \times 1$ complex Gaussian noise vector with identity covariance matrix \mathbf{I} , representing the floor noise plus normalized (and whitened) interference from PUs on channel k .

The Shannon rate of link $(i, d(i))$ on channel f_k is [1]:

$$R_i^{(k)} = \log \det(\mathbf{I} + \tilde{\mathbf{T}}_i^{(k)H} \mathbf{H}_{d(i),i}^{(k)} \mathbf{C}_{d(i)}^{-1} \mathbf{H}_{d(i),i}^{(k)} \tilde{\mathbf{T}}_i^{(k)}) \quad (2)$$

where $\mathbf{C}_{d(i)}^{(k)}$ is the noise-plus-interference covariance matrix at $d(i)$ over channel f_k , given by:

$$\mathbf{C}_{d(i)}^{(k)} = \mathbf{I} + \sum_{j \in \{\Phi_N \setminus i\}} \mathbf{H}_{d(i),j}^{(k)} \tilde{\mathbf{T}}_j^{(k)} \tilde{\mathbf{T}}_j^{(k)H} \mathbf{H}_{d(i),j}^{(k)H}$$

The total channel rate over all frequency bands of link i is:

$$R_i = \sum_{k \in \Psi_K} R_i^{(k)}. \quad (3)$$

We use $P_{s,k}^{(i)}$ to denote the power allocated on channel k (frequency dimension) at antenna s (space dimension) of CR user i . For user i , the total power allocated on all frequency bands and all antennas should not exceed a maximum power budget P_{\max} . Consequently,

$$\sum_{k \in \Psi_K} \sum_{s=1}^M P_{s,k}^{(i)} = \sum_{k \in \Psi_K} \text{tr}(\tilde{\mathbf{T}}_i^{(k)} \tilde{\mathbf{T}}_i^{(k)H}) \leq P_{\max}. \quad (4)$$

PU protection is provided in the form of database-authorized access [20] and frequency-dependent power masks on CR transmit powers. Note that the FCC [20] recently imposed power masks even for idle channels, if such channels are adjacent to PU-active channels. Let $\mathbf{P}_{\text{mask}} \stackrel{\text{def}}{=} (P_{\text{mask}}(f_1), P_{\text{mask}}(f_2), \dots, P_{\text{mask}}(f_K))$ denote the power mask on all channels, we require:

$$\sum_{s=1}^M P_{s,k}^{(i)} = \text{tr}(\tilde{\mathbf{T}}_i^{(k)} \tilde{\mathbf{T}}_i^{(k)H}) \leq P_{\text{mask}}(f_k). \quad (5)$$

Mathematically, the network throughput maximization problem can be stated as follows:

$$\begin{aligned} & \text{maximize} && \sum_{i \in \Phi_N} R_i \\ & \text{s.t. C1:} && \sum_{k \in \Psi_K} \text{tr}(\tilde{\mathbf{T}}_i^{(k)} \tilde{\mathbf{T}}_i^{(k)H}) \leq P_{\max}, \quad \forall i \in \Phi_N \\ & && \text{C2: } \text{tr}(\tilde{\mathbf{T}}_i^{(k)} \tilde{\mathbf{T}}_i^{(k)H}) \leq P_{\text{mask}}(f_k), \quad \forall k \in \Psi_K, \forall i \in \Phi_N. \end{aligned} \quad (6)$$

IV. GAME THEORETIC DESIGN

The optimization problem in (6) is not convex. Thus, even a centralized computation of the globally optimal solution is prohibitively expensive. To develop a distributed algorithm, we reformulate (6) as a noncooperative game and derive a pricing function for each CR link that guarantees a locally optimal solution.

A. Game Formulation

A noncooperative game is characterized by a set of players, their action/strategy spaces, and corresponding utility/payoff functions. For the above CRN, the set of CR links Φ_N represents the set of players. The action space is the union of the action spaces of various players, subject to constraints C1 and C2 in (6). The action/strategy space for each player is the set of all possible precoding matrices for the K frequency channels in Ψ_K . Formally, an action from the action space of link i is denoted by $\tilde{\mathbf{T}}_i \stackrel{\text{def}}{=} (\tilde{\mathbf{T}}_i^{\{1\}}, \tilde{\mathbf{T}}_i^{\{2\}}, \dots, \tilde{\mathbf{T}}_i^{(k)})$, which can be viewed as an $M \times KM$ block matrix, comprised of K $M \times M$ matrices. Let $\tilde{\mathbf{T}}_{-i} \stackrel{\text{def}}{=} (\tilde{\mathbf{T}}_1, \tilde{\mathbf{T}}_2, \dots, \tilde{\mathbf{T}}_{i-1}, \tilde{\mathbf{T}}_{i+1}, \dots, \tilde{\mathbf{T}}_N)$ be the set of actions from all links, except link i . The utility or payoff of player i for its action $\tilde{\mathbf{T}}_i$ is mapped to link i 's Shannon rate, which also depends on the selection of precoding matrices from other CR links $\tilde{\mathbf{T}}_{-i}$:

$$U_i(\tilde{\mathbf{T}}_i, \tilde{\mathbf{T}}_{-i}) \stackrel{\text{def}}{=} R_i. \quad (7)$$

Due to the noncooperative nature of the game, the transmitter of each link allocates its transmission power over both space and frequency dimensions, and configures its radiation pattern to maximize its own return. Formally, each CR user i solves the following problem for its set of precoding matrices $\tilde{\mathbf{T}}_i$:

$$\begin{aligned} & \text{maximize} && U_i(\tilde{\mathbf{T}}_i, \tilde{\mathbf{T}}_{-i}) \\ & \text{s.t. C1':} && \sum_{k \in \Psi_K} \text{tr}(\tilde{\mathbf{T}}_i^{(k)} \tilde{\mathbf{T}}_i^{(k)H}) \leq P_{\max} \\ & && \text{C2': } \text{tr}(\tilde{\mathbf{T}}_i^{(k)} \tilde{\mathbf{T}}_i^{(k)H}) \leq P_{\text{mask}}(f_k), \quad \forall k \in \Psi_K. \end{aligned} \quad (8)$$

By solving the above problem, CR users implicitly interact with each other through their choices of the precoding matrices. Under some conditions, the game reaches a NE where no user has an incentive to unilaterally deviate from. However, as each CR user behaves selfishly, the resulting NE is often far from the Pareto optimum, and the network throughput can be low. The efficiency of the NE can be improved by using appropriate pricing policies [15]. The utility function with price is defined as:

$$U'_i(\tilde{\mathbf{T}}_i, \tilde{\mathbf{T}}_{-i}) \stackrel{\text{def}}{=} U_i(\tilde{\mathbf{T}}_i, \tilde{\mathbf{T}}_{-i}) - F(\tilde{\mathbf{T}}_i) \quad (9)$$

where $F(\tilde{\mathbf{T}}_i)$ is the pricing function for link i . Consequently, we come up with the following noncooperative game with pricing in which each player $i \in \Phi_N$ solves the following problem:

$$\begin{aligned} & \text{maximize} && U'_i(\tilde{\mathbf{T}}_i, \tilde{\mathbf{T}}_{-i}) \\ & \text{s.t. C1' and C2' as in problem (8).} \end{aligned} \quad (10)$$

B. Pricing Policy

In economics, the pricing function can take various forms to account for various marketing and pricing policies, e.g.,

volume discount, coupon discount, etc. In the context of network resource allocation, both linear (e.g., [14] [21]) and nonlinear [22] pricing functions have been proposed to achieve one of two purposes. First, pricing has been used to impose desirable constraints by adjusting the cost/price of violation. These constraints can be, for examples, transmission rate demands [23] or interference constraints on CR transmissions [24]. Second, pricing has been used to improve the efficiency of a NE of a noncooperative game [14]. Pricing discourages players from behaving selfishly and incentivizes them to work in a cooperative way [15]. In this case, actions are not free or equally expensive. Players have to pay different taxes or prices for different actions, based on the level that these actions adversely affect the social welfare. In this work, pricing serves the second purpose.

We define the pricing function $F_u(\tilde{\mathbf{T}}_i)$ as follows:

$$F(\tilde{\mathbf{T}}_i) = \text{tr} \left[\tilde{\mathbf{T}}_i^H \mathbf{A}_i \tilde{\mathbf{T}}_i \right] \quad (11)$$

where

$$\mathbf{A}_i = \begin{bmatrix} \mathbf{A}_i^{(1)} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_i^{(2)} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{A}_i^{(K)} \end{bmatrix} \quad (12)$$

is a $KM \times KM$ block diagonal matrix, consisting of K blocks along its diagonal. The k th block $\mathbf{A}_i^{(k)}$ is an $M \times M$ positive-semidefinite matrix. \mathbf{A}_i is referred to as the pricing-factor matrix of CR link i and $\mathbf{A}_i^{(k)}$ is referred to as the pricing-factor submatrix at channel k of link i . The following theorem guarantees the existence of a NE of the game (10).

Theorem 1: There exists at least one NE for the noncooperative game in (10).

Proof: See [25]. \square

To guarantee a lower bound on the efficiency of the achieved NE, we next derive a user-dependent pricing function. The proposed pricing policy ensures that at the resulting NE, the CRN throughput is at least as good as that of a locally optimal solution to the network optimization problem (6).

Theorem 2: Let the k th matrix $\mathbf{A}_i^{(k)}$ of the block diagonal pricing-factor matrix \mathbf{A}_i in (12) be set to:

$$\mathbf{A}_i^{(k)} = \sum_{j \in \{\Phi_N \setminus i\}} \mathbf{H}_{d(j),i}^{(k)H} \mathbf{C}_{d(j)}^{(k)-1} \mathbf{H}_{d(j),j}^{(k)} [(\tilde{\mathbf{T}}_j^{(k)} \tilde{\mathbf{T}}_j^{(k)H})^{-1} + \mathbf{H}_{d(j),j}^{(k)H} \mathbf{C}_{d(j)}^{(k)-1} \mathbf{H}_{d(j),i}^{(k)}]^{-1} \mathbf{H}_{d(j),j}^{(k)H} \mathbf{C}_{d(j)}^{(k)-1} \mathbf{H}_{d(j),i}^{(k)}. \quad (13)$$

Then, the CRN's throughput at a NE of the game (10) equals to that of a locally optimal solution of the network-wide problem (6).

Proof: See Appendix A. \square

The rationale behind the choice of pricing function in (11) is to facilitate the derivation of the pricing factor matrix in Theorem 2. To give a physical interpretation of the pricing function, consider a special case where a CR node i uses omnidirectional transmission and equally allocates power on all frequency bands. In such a case, the precoders are diagonal matrices with identical diagonal elements. Hence, the pricing function in (11) is a weighted function of the powers $P_{s,k}^{(i)}$ allocated on streams (s, k) . The weights are exactly the diagonal elements (s, k) of the pricing factor matrix $\mathbf{A}_i^{(k)}$ that captures the per-unit price of possible interference on that spatial-

spectrum direction. Hence, the pricing function captures the interference that a transmitter induces on unintended receivers for a given set of precoders. It is worth noting that the per-unit price of interference depends on the set of unintended receivers. In other words, the interference price varies from one "market" (user) to another.

To compute the pricing-factor matrix \mathbf{A}_i in (13), a CR transmitter i needs to obtain feedback regarding the interference-plus-noise covariance, and the precoding and channel matrices from other links. In practice, if the channel gain matrix from i to $d(j)$ is weak, i.e., $\mathbf{H}_{d(j),i}^{(k)} \approx \mathbf{0}$, there is no need for $d(j)$ to send its feedback to i . Hence, i only gets feedback from receivers $d(j)$ that are within i 's vicinity. It is also worth noting that the feedback information is locally available at a receiver $d(j)$ as a byproduct of its decoding process (e.g., successive interference cancelation (SIC) receivers [1]). The k th block $\mathbf{A}_i^{(k)}$ of the pricing factor matrix in (13) is similar to that in [9] for a single-band MIMO ad hoc network using first-order Taylor series approximation.

C. Best Response: Optimal Antenna Radiation Directions and Power Allocation

We now solve the individual utility optimization problem (10), from which a CR user finds its best response given others' actions. Because problem (10) is convex, it can be solved by standard methods, e.g., interior point [26], requiring polynomial time w.r.t. to the number of variables. In [6], the authors solved a similar problem using semidefinite programming. However, the number of variables in (10) grows quadratically with the number of antennas, and can be very large. In this section, we develop an efficient algorithm whose complexity is independent of the number of antennas.

Recalling the convexity of (10) and that the Slater's conditions can easily be shown to hold [26], strong duality holds for problem (10), i.e., an optimal solution $\tilde{\mathbf{T}}_i$ to (10) should also solve the following dual problem:

$$\text{DP : } \underset{\{\alpha_i^{(k)}, \gamma_i \geq 0, \forall k \in \Psi_K\}}{\text{minimize}} \quad D(\alpha_i^{(k)}, \gamma_i) \quad (14)$$

where $D(\alpha_i^{(k)}, \gamma_i)$ is the dual function, defined as:

$$D(\alpha_i^{(k)}, \gamma_i) = \max_{\{\tilde{\mathbf{T}}_i^{(k)}, \forall k \in \Psi_K\}} L_i(\tilde{\mathbf{T}}_i, \alpha_i^{(k)}, \gamma_i). \quad (15)$$

In the above, $L_i(\tilde{\mathbf{T}}_i, \alpha_i^{(k)}, \gamma_i)$ is the Lagrangian function of the utility maximization problem at user i , written as (16), where $\alpha_i^{(k)}$ and γ_i are nonnegative Lagrangian multipliers.

Theorem 3: The $M \times KM$ block matrix $\tilde{\mathbf{T}}_i$ that solves (10) (for the user's best response) must have its k th block, $\tilde{\mathbf{T}}_i^{(k)}$ in a form of a generalized eigen matrix of the matrices $\mathbf{H}_{d(i),i}^{(k)H} \mathbf{C}_{d(i)}^{(k)-1} \mathbf{H}_{d(i),i}^{(k)}$ and $\mathbf{A}_i^{(k)} + (\alpha_i^{(k)} + \gamma_i) \mathbf{I}$, where $\alpha_i^{(k)}$ and γ_i are the optimal Lagrange multipliers of (10). In other words, the following equations must hold $\forall k \in \Psi_K$ for a $M \times M$ diagonal matrix $\Lambda_i^{(k)}$:

$$\mathbf{H}_{d(i),i}^{(k)H} \mathbf{C}_{d(i)}^{(k)-1} \mathbf{H}_{d(i),i}^{(k)} \tilde{\mathbf{T}}_i^{(k)} = [\mathbf{A}_i^{(k)} + (\alpha_i^{(k)} + \gamma_i) \mathbf{I}] \tilde{\mathbf{T}}_i^{(k)} \Lambda_i^{(k)}. \quad (20)$$

Proof: See Appendix B. \square

As discussed before, the precoding matrix $\tilde{\mathbf{T}}_i^{(k)}$ determines both the directions of the antenna radiation as well as how

$$\begin{aligned}
 L_i(\tilde{\mathbf{T}}_i, \alpha_i^{(k)}, \gamma_i) &= U'_i(\tilde{\mathbf{T}}_i, \tilde{\mathbf{T}}_{-i}) - \sum_{k \in \Psi_K} \alpha_i^{(k)} [\text{tr}(\tilde{\mathbf{T}}_i^{(k)} \tilde{\mathbf{T}}_i^{(k)H}) - P_{\text{mask}}(f_k)] - \gamma_i [\sum_{k \in \Psi_K} \text{tr}(\tilde{\mathbf{T}}_i^{(k)} \tilde{\mathbf{T}}_i^{(k)H}) - P_{\text{max}}] \\
 &= \sum_{k \in \Psi_K} \{R_i^{(k)} - \text{tr}(\tilde{\mathbf{T}}_i^{(k)H} \mathbf{A}_i^{(k)} \tilde{\mathbf{T}}_i^{(k)})\} - \sum_{k \in \Psi_K} \alpha_i^{(k)} [\text{tr}(\tilde{\mathbf{T}}_i^{(k)} \tilde{\mathbf{T}}_i^{(k)H}) - P_{\text{mask}}(f_k)] - \gamma_i [\sum_{k \in \Psi_K} \text{tr}(\tilde{\mathbf{T}}_i^{(k)} \tilde{\mathbf{T}}_i^{(k)H}) - P_{\text{max}}]
 \end{aligned} \tag{16}$$

$$L_i(\tilde{\mathbf{T}}_i, \alpha_i^{(k)}, \gamma_i) = \sum_{k \in \Psi_K} \left\{ \sum_{s=1}^M \left\{ \log(1 + P_{s,k}^{(i)} \text{diag}_s(\mathbf{\Pi}_i^{(k)})) - P_{s,k}^{(i)} \text{diag}_s(\mathbf{\Omega}_i^{(k)}) \right\} + \alpha_i^{(k)} P_{\text{mask}}(f_k) + \frac{\gamma_i}{K} P_{\text{max}} \right\} \tag{17}$$

$$D(\alpha_i^{(k)}, \gamma_i) = \sum_{k \in \Psi_K} \left\{ \sum_{s=1}^M \left\{ \log \frac{\text{diag}_s(\mathbf{\Pi}_i^{(k)})}{\text{diag}_s(\mathbf{\Omega}_i^{(k)})} - 1 + \frac{\text{diag}_s(\mathbf{\Omega}_i^{(k)})}{\text{diag}_s(\mathbf{\Pi}_i^{(k)})} \right\} + \alpha_i^{(k)} P_{\text{mask}}(f_k) + \frac{\gamma_i}{K} P_{\text{max}} \right\} \tag{18}$$

$\forall s, k$ such that $\text{diag}_s(\mathbf{\Pi}_i^{(k)}) > \text{diag}_s(\mathbf{\Omega}_i^{(k)}) > 0$.

$$L(\alpha_i^{(k)}, \gamma_i, p, \boldsymbol{\lambda}) = D(\alpha_i^{(k)}, \gamma_i) + \frac{p}{2} \{(\max\{0, \lambda_1 - p\gamma_i\})^2 - (\lambda_1)^2\} + \frac{p}{2} \sum_{k \in \Psi_K} \{(\max\{0, \lambda_{k+1} - p\alpha_i^{(k)}\})^2 - (\lambda_{k+1})^2\} \tag{19}$$

node i allocates its transmission power on different antennas over channel k . Theorem 3 provides a class of matrices that the solutions of (10) must belong to. This class gives the directions that user i should point its antenna radiation to.

The next step is to find the optimal power allocation $P_{s,k}^{(i)}$ over the set of KM data streams. To ensure that $\tilde{\mathbf{T}}_i^{(k)}$ belongs to the class of matrices specified by Theorem 3, let:

$$\tilde{\mathbf{T}}_i^{(k)} \stackrel{\text{def}}{=} \mathbf{T}_i^{(k)} \mathbf{P}_k^{(i)1/2} \tag{21}$$

where $\mathbf{T}_i^{(k)}$ is an $M \times M$ matrix with unit-norm column vectors that satisfies (20). This matrix can be found by normalizing the generalized eigen matrix $\tilde{\mathbf{T}}_i^{(k)}$. $\mathbf{P}_k^{(i)1/2}$ is a square root matrix of the $M \times M$ diagonal matrix $\mathbf{P}_k^{(i)}$ whose diagonal entry (s, s) is the power allocated for sub-channel (s, k) , $P_{s,k}^{(i)}$. We can verify that the expression of $\tilde{\mathbf{T}}_i^{(k)}$ in (21) satisfies (20).

As $\tilde{\mathbf{T}}_i^{(k)}$ is a generalized eigen matrix of matrices $\mathbf{H}_{d(i),i}^{(k)H} \mathbf{C}_{d(i)}^{(k)-1} \mathbf{H}_{d(i),i}^{(k)}$ and $\mathbf{A}_i^{(k)} + (\alpha_i^{(k)} + \gamma_i) \mathbf{I}$, $\mathbf{T}_i^{(k)}$ should diagonalize each of the two matrices [27]:

$$\begin{aligned}
 \mathbf{T}_i^{(k)H} [\mathbf{H}_{d(i),i}^{(k)H} \mathbf{C}_{d(i)}^{(k)-1} \mathbf{H}_{d(i),i}^{(k)}] \mathbf{T}_i^{(k)} &= \mathbf{\Pi}_i^{(k)} \quad \text{and} \\
 \mathbf{T}_i^{(k)H} [\mathbf{A}_i^{(k)} + (\alpha_i^{(k)} + \gamma_i) \mathbf{I}] \mathbf{T}_i^{(k)} &= \mathbf{\Omega}_i^{(k)}
 \end{aligned} \tag{22}$$

where $\mathbf{\Pi}_i^{(k)}$ and $\mathbf{\Omega}_i^{(k)}$ are $M \times M$ diagonal matrices.

Note that although its columns have unit-norm, $\mathbf{T}_i^{(k)}$ is not an orthonormal matrix, as $\mathbf{A}_i^{(k)}$ is generally not similar to \mathbf{I} . Hence, $\mathbf{T}_i^{(k)}$ (and $\tilde{\mathbf{T}}_i^{(k)}$) does not necessarily diagonalize $\mathbf{A}_i^{(k)}$. This observation is twofold. First, this points out that the derivation in [12] does not hold in general. Second, although optimal power allocation for KM data streams seems very similar to a general water filling problem [10] with multiple water levels (one water level per each frequency band), this allocation cannot be determined by the algorithms in [10] [11]. This is because $\tilde{\mathbf{T}}_i^{(k)}$ does not diagonalize $\mathbf{A}_i^{(k)}$ in (10), and hence we cannot convert (10) to a general water filling problem.

Plugging (22) into the Lagrangian function (16), we have (17). The optimal power allocation $P_{s,k}^{(i)}$ is obtained by equat-

ing the derivative of (17) w.r.t $P_{s,k}^{(i)}$ to zero:

$$\frac{\partial L_i(\tilde{\mathbf{T}}_i, \alpha_i^{(k)}, \gamma_i)}{\partial P_{s,k}^{(i)}} = \frac{\text{diag}_s(\mathbf{\Pi}_i^{(k)})}{1 + P_{s,k}^{(i)} \text{diag}_s(\mathbf{\Pi}_i^{(k)})} - \text{diag}_s(\mathbf{\Omega}_i^{(k)}) = 0 \tag{23}$$

Thus,

$$P_{s,k}^{(i)} = \max \left(0, \frac{\text{diag}_s(\mathbf{\Pi}_i^{(k)}) - \text{diag}_s(\mathbf{\Omega}_i^{(k)})}{\text{diag}_s(\mathbf{\Pi}_i^{(k)}) \text{diag}_s(\mathbf{\Omega}_i^{(k)})} \right). \tag{24}$$

Plugging (24) into (17), we obtain the dual function $D(\alpha_i^{(k)}, \gamma_i)$ in (18). To solve the DP (14) for $\alpha_i^{(k)}$, $k = 1, \dots, K$, and γ_i ($K+1$ variables), we note that the problem is convex. Hence, any stationary point is a globally optimal solution. Moreover, as the objective function and constraints of the primal problem (10) are continuous w.r.t. every entry of $\tilde{\mathbf{T}}_i$, the dual function $D(\alpha_i^{(k)}, \gamma_i)$ is differentiable w.r.t. $\alpha_i^{(k)}$ and γ_i [26]. Hence, a gradient algorithm can be used to obtain the optimal Lagrangian multipliers $\alpha_i^{(k)}$ and γ_i by searching for a stationary point of the augmented Lagrangian of the DP in (19), where p is a positive penalty parameter (for violating the constraints) and $\boldsymbol{\lambda} \stackrel{\text{def}}{=} \{\lambda_1, \dots, \lambda_{K+1}\}$ are nonnegative multipliers.

Our gradient algorithm uses Armijo step in the steepest descent direction. This search mechanism together with the above analysis are summarized in Algorithm 1. By exploiting the strong duality, Algorithm 1 needs to deal with only $K+1$ variables, instead of $2KM^2$ variables for the primal problem (10). Before developing a centralized algorithm that serves as a performance benchmark, we briefly discuss how a MAC protocol can implement Algorithm 1.

D. MAC Protocol

Using either a predefined or frequency-hopping-based control channel, we can design a MAC protocol that executes Algorithm 1. This protocol consists of three windows: *Access window*, *training window*, and *data window*. The access window allows CR nodes to contend for channels. These nodes first exchange RTS and CTS packets. Unlike IEEE 802.11, our MAC design does not use RTS/CTS packets to silence nearby nodes and reserve the transmission floor for the upcoming transmissions. After this phase, CR pairs who have just sent and received RTS/CTS packets are admitted to the

training window. The training window is used by nodes to exchange/negotiate their transmit strategies (precoding matrices). The signalling packets in either access or transmit windows can also be used to embed training sequences to obtain CSI matrices. The data window then follows with multiple data packets, sent using negotiated transmission strategies. This approach is referred to as a flow-based approach in [14].

Algorithm 1 : Distributed Algorithm for Power Allocation and Spectrum Management

```

1: Input:
 $\tilde{\mathbf{T}}_{-i} = [\tilde{\mathbf{T}}_1(t+1), \dots, \tilde{\mathbf{T}}_{i-1}(t+1), \tilde{\mathbf{T}}_{i+1}(t), \dots, \tilde{\mathbf{T}}_N(t)]$ 
with Gauss-Seidel iteration
 $\tilde{\mathbf{T}}_{-i} = [\tilde{\mathbf{T}}_1(t), \dots, \tilde{\mathbf{T}}_{i-1}(t), \tilde{\mathbf{T}}_{i+1}(t), \dots, \tilde{\mathbf{T}}_N(t)]$ 
with Jacobi iteration
2: Initialize
 $\tilde{\mathbf{T}}_i^{(k)}(t+1) \leftarrow \tilde{\mathbf{T}}_i^{(k)}(t), \gamma_i \leftarrow 0; \alpha_i^{(k)} \leftarrow 0, \forall k \in \Psi_K$ 
3: while true do
4:    $\beta \leftarrow .7, \sigma \leftarrow .1$  (%used in Armijo search)
5:    $\lambda_k \leftarrow .1 \forall k = 1 \dots (K+1)$ 
6:    $p \leftarrow 1$ 
7:   while  $\partial L(\alpha_i^{(k)}, \gamma_i, p, \lambda) \neq 0$  do
8:      $step \leftarrow 0.1$ 
9:      $D \leftarrow \partial L(\alpha_i^{(k)}, \gamma_i, p, \lambda)$ 
10:     $d \leftarrow -step \times D; m \leftarrow 0$ 
11:    (%find Armijo step size)
12:    while  $L(\alpha_i^{(k)}, \gamma_i, p, \lambda) - L(\alpha_i^{(k)} + d, \gamma_i + d, p, \lambda) \geq$ 
 $-\sigma \beta^m step \partial LD$  do
13:       $step \leftarrow step \times \beta; m \leftarrow m + 1$ 
14:       $d \leftarrow -step \times D$ 
15:    end while
16:     $\alpha_i^{(k)} \leftarrow \alpha_i^{(k)} + d, \gamma_i \leftarrow \gamma_i + d$ 
17:  end while
18:  if  $\min(\alpha_i^{(k)}, \gamma_i, \forall k \in \Psi_K) \geq 0$  break
19:   $\forall k \in \Psi_K :$ 
20:     $\lambda_k \leftarrow \lambda_k - p\alpha_i^{(k)}$  if  $\lambda_k - p\alpha_i^{(k)} > 0$  else  $\lambda_k = 0$ 
21:     $\lambda_1 \leftarrow \lambda_1 - p\gamma_i$  if  $\lambda_1 - p\gamma_i \geq 0$  else  $\lambda_1 = 0$ 
22:     $p \leftarrow p \times \mu$  (% $\mu \geq 1$ , increase cost of violation)
23: end while
24: Plug  $\gamma_i, \alpha_i^{(k)}$  into (20) (Theorem 3) to find  $\mathbf{T}_i^{(k)}$ 
 $\mathbf{P}_k^{(i)}$  is found from (22) and (24).  $\tilde{\mathbf{T}}_i^{(k)}$  is found from (21).
25: RETURN  $\tilde{\mathbf{T}}_i^{(k)}(t+1), \forall k \in \Psi_K$  at time  $(t+1)$ 

```

To reduce the overhead of the training window, one may relax the time scale of recalculating the pricing-factor matrix, trading off throughput for less frequent updates. One can even omit the training window altogether by having nodes embed updated information into every data and ACK packet. Then, upon receiving an ACK for each data packet, a transmitter recomputes its pricing-factor matrix. This method is referred to as packet-based [28] [14].

An important issue for protocol designers is how to set the size of the training window. This depends on the convergence speed of the updating process. To ensure that the training window is not too long, the updating and negotiation process must converge. During the training window, a node can use either Gauss-Seidel (sequential) or Jacobi (parallel) iterations (see Algorithm 1) to update its precoding matrices. Although we cannot prove the convergence under the Jacobi iteration, simulations show that the distributed algorithm converges even faster with Jacobi iterations than with Gauss-Seidel. Convergence under the Gauss-Seidel iteration is claimed in the following theorem.

Theorem 4: Under the sequential updating procedure (Gauss-Seidel), Algorithm 1 drives the game (10) to its NE.

Proof: See [25]. \square

In [8] [11], the authors map the throughput maximization game (of a single-frequency MIMO CRN) into a *variational inequality* problem (see [29] and therein references for a tutorial on variational inequality theory) for the purpose of providing the convergence and uniqueness conditions of the NE. Intuitively, the game converges to a unique NE if there is not too much interference at a receiver and all channel matrices are full column-rank. The former condition requires transmitters to adjust their transmission parameters and the latter is not always guaranteed. However, using our proposed pricing policy and the Gauss-Seidel procedure, the game (10) always converges to a NE without requiring transmitters and channel matrices to meet any additional requirements.

Both Gauss-Seidel and Jacobi updating procedures are *synchronous*, requiring CR nodes to be in sync. This may not always be possible. By contrast, using *asynchronous* update, some players, at some iterations, may not update other players with their strategies (e.g., due to packet collisions). Using variational inequality theory, the mean-value theorem, and following the routine in [8], we can provide conditions under which the game (10) converges to a unique NE under either synchronous or asynchronous updates (we omit the detailed manipulation due to space limitation).

Theorem 5: The game (10) always converges to a unique NE if all channel matrices are full column-rank and the spectrum radius of the matrix $J^{-1}\Gamma$ is less than 1, where J is an $N \times N$ diagonal matrix with diagonal elements σ_i and Γ is a $N \times N$ matrix of elements $\kappa_{i,j}$ with the equation at the top of the following page.

The operators $\text{eig}_{\min}(\cdot)$ and $\text{eig}_{\max}(\cdot)$ give the smallest and largest eigenvalues of a matrix, respectively.

V. CENTRALIZED ALGORITHM

A centralized algorithm can be obtained by formulating the problem as a cooperative game, where a network operator controls the behaviors of all players in order to maximize the network throughput. In this section, we use the augmented Lagrangian multiplier method to derive such a centralized algorithm. We rewrite the network-wide problem (6) as (25).

The augmented Lagrangian [26] of (25) is given in (26), where p is a positive penalty parameter (for violating the constraints), and $\alpha_i^{(k)}$ and γ_i are nonnegative Lagrangian multipliers. At a local optimal solution, (27) holds. The first term in (27) is computed in (33). Its second term is given by (30). Since c_i and $c_{k,i}$ are continuously differentiable w.r.t. every entry of $\tilde{\mathbf{T}}$, the third and fourth terms in (27) are also continuously differentiable [26]. Their derivatives are as follows:

$$\frac{\partial \{(\max\{0, \gamma_i + pc_i\})^2\}}{\partial \tilde{\mathbf{T}}_i^{(k)*}} = \begin{cases} 0 & \text{if } \gamma_i + pc_i \leq 0 \\ 2p(\gamma_i + pc_i)\tilde{\mathbf{T}}_i^{(k)} & \end{cases}$$

$$\frac{\partial \{(\max\{0, \alpha_i^{(k)} + pc_{k,i}\})^2\}}{\partial \tilde{\mathbf{T}}_i^{(k)*}} = \begin{cases} 0 & \text{if } \alpha_i^{(k)} + pc_{k,i} \leq 0 \\ 2p(\alpha_i^{(k)} + pc_{k,i})\tilde{\mathbf{T}}_i^{(k)} & \end{cases}$$

We use the gradient search algorithm with Armijo step size [26] to find $(\tilde{\mathbf{T}}, \alpha_i^{(k)}, \gamma_i, p)$ such that (27) holds for all bands k and all users i . The details of the centralized algorithm is presented in Algorithm 2. We emphasize that the network throughput may vary from a locally optimal point to another. Hence, to account for such phenomenon, one can run the simulations several times with different initializations and take

$$\begin{aligned} \sigma_i &\stackrel{\text{def}}{=} \min_{k \in \Psi_K} \left\{ \text{eig}_{\min}^2 \left(\mathbf{H}_{d(i),i}^{(k)H} \left(\mathbf{I} + \sum_{j \in \Phi_N} P_{\text{mask}}(f_k) \mathbf{H}_{d(j),i}^{(k)} \mathbf{H}_{d(j),i}^{(k)H} \right) \mathbf{H}_{d(i),i}^{(k)} \right)^{-1} \right\} \\ \kappa_{j,i} &\stackrel{\text{def}}{=} \max_{k \in \Psi_K} \left(\text{eig}_{\max} \left(\mathbf{H}_{d(j),i}^{(k)} \mathbf{H}_{d(j),i}^{(k)H} \right) \text{eig}_{\max} \left(\mathbf{H}_{d(i),i}^{(k)H} \mathbf{H}_{d(i),i}^{(k)} \right) \right) \\ &\quad \forall i \neq j \in \Psi_N \quad \text{and } \kappa_{j,i} = 0 \text{ if } i = j \end{aligned}$$

$$\begin{aligned} &\text{minimize} \quad - \sum_{i \in \Phi_N} R_i \\ &\text{s.t.} \quad c_i = \sum_{k \in \Psi_K} \text{tr}(\tilde{\mathbf{T}}_i^{(k)} \tilde{\mathbf{T}}_i^{(k)H}) - P_{\text{max}} \leq 0, \forall i \in \Phi_N \\ &\quad c_{i,k} = \text{tr}(\tilde{\mathbf{T}}_i^{(k)} \tilde{\mathbf{T}}_i^{(k)H}) - P_{\text{mask}}(f_k) \leq 0, \forall k \in \Psi_K, \forall i \in \Phi_N \end{aligned} \quad (25)$$

$$L(\tilde{\mathbf{T}}, \alpha_i^{(k)}, \gamma_i, p) = - \sum_{i \in \Phi_N} R_i + \frac{p}{2} \sum_{i \in \Phi_N} \{(\max\{0, \gamma_i + pc_i\})^2 - (\gamma_i)^2\} + \frac{p}{2} \sum_{i \in \Phi_N} \sum_{k \in \Psi_K} \{(\max\{0, \alpha_i^{(k)} + pc_{k,i}\})^2 - (\alpha_i^{(k)})^2\} \quad (26)$$

$$0 = \frac{\partial L(\tilde{\mathbf{T}}, \alpha_i^{(k)}, \gamma_i, p)}{\partial \tilde{\mathbf{T}}_i^{(k)*}} = - \sum_{j \in \Phi_N \setminus i} \frac{\partial R_j^{(k)}}{\partial \tilde{\mathbf{T}}_i^{(k)*}} - \frac{\partial R_i^{(k)}}{\partial \tilde{\mathbf{T}}_i^{(k)*}} + \frac{p}{2} \left\{ \frac{\partial \{(\max\{0, \gamma_i + pc_i\})^2\}}{\partial \tilde{\mathbf{T}}_i^{(k)*}} + \frac{\partial \{(\max\{0, \alpha_i^{(k)} + pc_{k,i}\})^2\}}{\partial \tilde{\mathbf{T}}_i^{(k)*}} \right\} \quad (27)$$

$$\mathbf{x}_i = \left[\Re[\text{vec}(\tilde{\mathbf{T}}_i)]^T, \Im[\text{vec}(\tilde{\mathbf{T}}_i)]^T \right]^T = \left[\Re[\text{vec}(\tilde{\mathbf{T}}_i^{(1)})]^T, \dots, \Re[\text{vec}(\tilde{\mathbf{T}}_i^{(K)})]^T, \Im[\text{vec}(\tilde{\mathbf{T}}_i^{(1)})]^T, \dots, \Im[\text{vec}(\tilde{\mathbf{T}}_i^{(K)})]^T \right]^T \quad (28)$$

$$\nabla_{\mathbf{x}} L = 2 \left[\Re[\text{vec}(\frac{\partial L}{\partial \tilde{\mathbf{T}}_1^{(1)*}})]^T, \dots, \Re[\text{vec}(\frac{\partial L}{\partial \tilde{\mathbf{T}}_N^{(K)*}})]^T, \Im[\text{vec}(\frac{\partial L}{\partial \tilde{\mathbf{T}}_1^{(1)*}})]^T, \dots, \Im[\text{vec}(\frac{\partial L}{\partial \tilde{\mathbf{T}}_N^{(K)*}})]^T \right]^T \quad (29)$$

$$\frac{\partial R_i^{(k)}}{\partial \tilde{\mathbf{T}}_i^{(k)*}} = \mathbf{H}_{d(i),i}^{(k)H} (\mathbf{C}_{d(i)}^{(k)} + \mathbf{H}_{d(i),i}^{(k)} \tilde{\mathbf{T}}_i^{(k)} \tilde{\mathbf{T}}_i^{(k)H} \mathbf{H}_{d(i),i}^{(k)H})^{-1} \mathbf{H}_{d(i),i}^{(k)} \tilde{\mathbf{T}}_i^{(k)} \quad (30)$$

the average throughput. The running time for Algorithm 2 can be high as it involves NKM^2 complex variables (or $2NKM^2$ real ones). To implement Algorithm 2, we use the following isomorphism mapping from a complex matrix to a vector of real variables. The vector of variables $\mathbf{x} = [(\mathbf{x}_i^T)_{i=1}^N]^T$, with \mathbf{x}_i in (28) and the corresponding Lagrangian gradient in (29).

VI. NUMERICAL RESULTS

In this section, we evaluate the performance of the distributed algorithm using MATLAB simulations. We compare the network throughput of the distributed algorithm with the centralized one and the greedy algorithm, where nodes act selfishly to maximize their own rates. The greedy algorithm is exactly the same as the distributed one except that the pricing-factor matrix \mathbf{A}_i is a null matrix. Another algorithm called *uniform* is obtained by uniformly dividing a node's total transmit power over all available channels and then applying the single-band approach in [9] for each channel. We emphasize that the uniform algorithm neither meets the optimality conditions (38) of the network problem (6) nor solves the per-user problem (10).

Since the number of variables in the centralized algorithm is quite high ($2NKM^2$), its running time can be very long. To

compare the four algorithms, we consider a CRN of 10 links, 3 channels ($f_1 = 2.4$ GHz, $f_2 = 2.7$ GHz, and $f_3 = 3$ GHz with identical channel bandwidth of 1 MHz), and 4 antennas per node. The results are averaged over 30 runs. In each run, links are randomly placed in a square of length 100 meters. We take $P_{\text{max}} = 2$ W and $P_{\text{mask}} = 0.8$ W for all channels. The channel fading is flat with attenuation factor of 2. The spreading angles of the signal at received antennas vary from $-\pi/5$ to $\pi/5$. For the lowest frequency, we assume that the received power at a reference distance of 100 meters reduces by 10 dB compared with the transmit power. To account for the frequency-dependent attenuation factor, we assume that the received power at the reference distance decreases 2 dB more if the frequency increases by 300 MHz. The PU interference is treated as floor noise that together with the thermal noise is normalized to a unit variance.

A snapshot of the network topology and antenna radiation patterns at steady state over channel f_2 is shown in Fig. 2. We can visually notice that the transmitters under the distributed and centralized algorithms often steer their beams away from neighboring receivers (three representative pairs of transmitter and a nearby unintended receiver are highlighted in ovals). This results from attempting to minimize the price function (11). It can also be seen that the antenna patterns

Algorithm 2 : Centralized Algorithm

```

1: Initialize
    $\mathbf{T}_i^{(k)} \leftarrow \mathbf{I}, \gamma_i \leftarrow 0; \alpha_i^{(k)} \leftarrow 0, \forall k \in \Psi_K, \forall i \in \Phi_N$ 
2: while true do
3:    $\beta \leftarrow .7, \sigma \leftarrow .1$  (%used in Armijo search)
4:    $\gamma_i \leftarrow 0, \alpha_i^{(k)} \leftarrow 0, \forall k \in \Psi_K, \forall i \in \Phi_N$ 
5:    $p \leftarrow 1$ 
6:   while  $\partial L(\tilde{\mathbf{T}}, \alpha_i^{(k)}, \gamma_i, p) \neq 0$  do
7:     step  $\leftarrow 0.1$ 
8:      $D \leftarrow \partial L(\tilde{\mathbf{T}}, \alpha_i^{(k)}, \gamma_i, p)$ 
9:      $d \leftarrow -\text{step} \times D; m \leftarrow 0$ 
10:    (%find Armijo step size)
11:    while  $L(\tilde{\mathbf{T}}, \alpha_i^{(k)}, \gamma_i, p) - L(\tilde{\mathbf{T}} + d, \alpha_i^{(k)}, \gamma_i, p) \leq$ 
        $-\sigma \beta^m \text{step} \partial L D$  do
12:      step  $\leftarrow \text{step} \times \beta; m \leftarrow m + 1$ 
13:       $d \leftarrow -\text{step} \times D$ 
14:    end while
15:     $\mathbf{T} \leftarrow \mathbf{T} + d$ 
16:  end while
17:  if  $\max(c_i, c_{i,k}, \forall k \in \Psi_K, \forall i \in \Phi_N) \leq 0$  break
18:   $\forall k \in \Psi_K, \forall i \in \Phi_N$  :
19:     $\gamma_i = \gamma_i + p c_i$  if  $\gamma_i + p c_i \geq 0$  else  $\gamma_i = 0$ 
20:     $\alpha_i^{(k)} = \alpha_i^{(k)} + p c_{k,i}$  if  $\alpha_i^{(k)} + p c_{k,i} \geq 0$  else  $\alpha_i^{(k)} = 0$ 
21:     $p \leftarrow p \times \mu$  (% $\mu \geq 1$ , increase cost of violation)
22:  end while
23: Return  $\tilde{\mathbf{T}}_i^{(k)}, \forall k \in \Psi_K, \forall i \in \Phi_N$ 

```

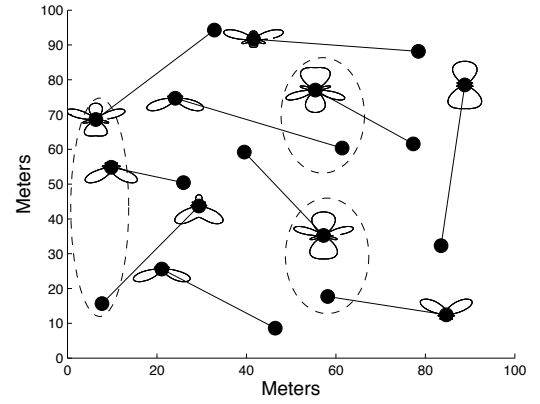
of the distributed and centralized algorithms are very similar, suggesting the two algorithms may converge to the same point.

Fig. 3(a) depicts the network throughput under four algorithms (distributed, centralized, greedy, and uniform) versus the number of iterations. Although the network performance at the converged points of the distributed and centralized algorithms change with their starting points, after averaging over multiple runs with different initializations, the throughput of the distributed algorithm is almost the same as that of the centralized one. We also notice that by using the proposed pricing policy to regulate interference, the distributed algorithm has almost twice the throughput of the greedy one. The uniform algorithm also improves network throughput over the greedy one but it remains inferior to our distributed algorithm. This is because the uniform algorithm evenly allocates power over all channels and does not optimize over the frequency dimension, while the distributed algorithm attempts to optimize the antenna radiation patterns and power allocation over both space and frequency.

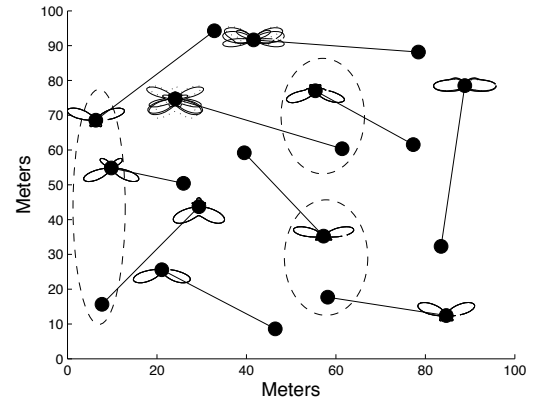
We say that an algorithm converges if the normalized difference in throughput between two consecutive iterations is less than a given threshold (i.e., 3%). The convergence speed of the distributed algorithm versus the number of links is shown in Fig. 3(b).

Fig. 3(c) depicts the network throughput under the distributed and greedy algorithms versus the number of links. The distributed algorithm consistently provides higher throughput than the greedy and uniform algorithms. The improvement becomes more significant with more links. That is because as node density increases (higher number of links), network interference increases, so interference management becomes more critical in improving the throughput.

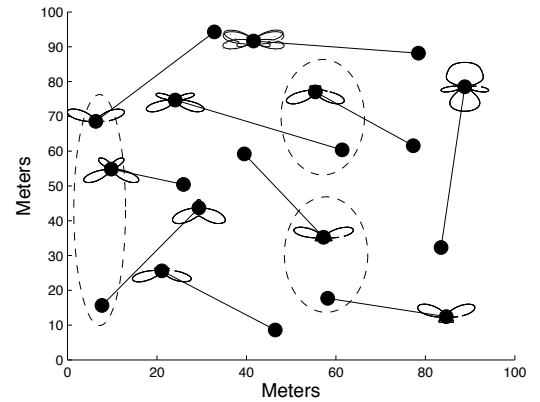
To evaluate the energy efficiency of the four algorithms, we record in Table I the average power consumption and power allocation per node for various algorithms. Without regulating interference, nodes under the greedy algorithm selfishly compete for their own throughput by always using their maximum



(a) Greedy algorithm



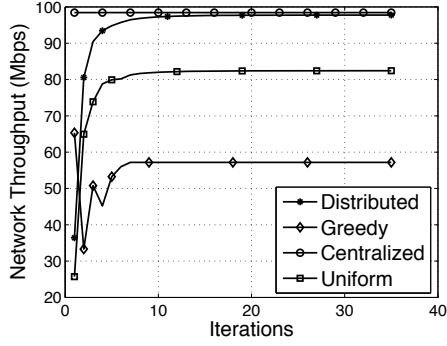
(b) Distributed algorithm



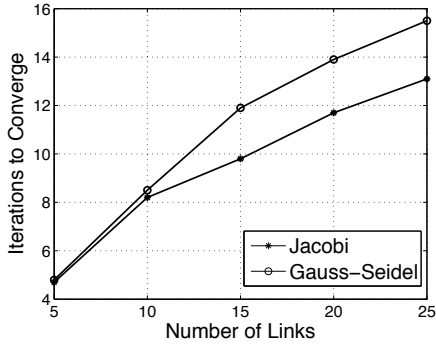
(c) Centralized algorithm

Fig. 2. Antenna radiation patterns on channel 2 under the greedy, distributed, and centralized algorithms.

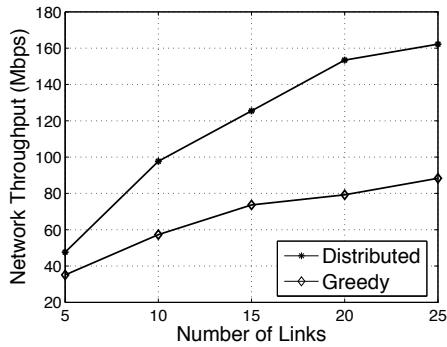
power (2 W), leading to the highest power consumption among the four algorithms. The power consumption of the distributed algorithm is comparable with that of the centralized and uniform ones, and 10% less than that of the greedy one. Power allocation over both space and frequency at a representative node under the distributed algorithm is shown Table II. From



(a)



(b)



(c)

Fig. 3. (a) Network throughput vs. iterations, (b) convergence speed of the distributed algorithm, (c) network throughput vs. the number of links.

Tables I and II, we notice that the inequality constraints in problems (10) and (6) are not active at their solutions. That is because transmitting at high power may be expensive due to the proposed pricing method.

VII. CONCLUSIONS AND FUTURE WORK

In this work, we investigated the spectrum sharing problem in a multi-antenna CRN. By adjusting the precoding matrices, we optimized the allocation of power over both the frequency and space dimensions while managing the antennas' radiation beams to reduce network interference, aiming at maximizing

TABLE I
AVERAGE POWER CONSUMPTION (IN WATTS) ALLOCATED OVER DIFFERENT CHANNELS.

Channels	Centralized	Greedy	Distributed	Uniform
f_1	0.768	0.71	0.76	0.658
f_2	0.643	0.66	0.61	0.556
f_3	0.422	0.63	0.44	0.627
Total	1.823	2.00	1.81	1.831

TABLE II
POWER ALLOCATION (IN WATTS) AT A NODE OVER SPACE AND FREQUENCY DIMENSIONS UNDER THE DISTRIBUTED ALGORITHM.

Antennas	f_1	f_2	f_3
1	0.135	0.085	$0.15e - 10$
2	0.209	0.386	0.02
3	0.550	0.314	$0.06e - 10$
4	0.194	0.035	0.305
Total=1.913	0.788	0.8	0.325

network throughput. Using game theory and the strong duality in convex optimization, we designed a low-complexity distributed algorithm that achieves the same throughput as a locally optimal point of the non-convex centralized network problem. The key idea behind the algorithm is the introduction of a diagonal block pricing-factor matrix for each CR. This matrix regulates network interference by encouraging CRs to work in a cooperative manner. Simulations show that the proposed algorithm dramatically improves network throughput and achieves higher energy efficiency, compared with existing solutions. As a future work, one can extend the proposed pricing policy to coordinated multi-cell systems and also heterogeneous spectrum sharing networks in which the sets of available frequencies at nodes are different. Moreover, because CSI is vulnerable to estimation errors, one may wish to design a robust game model to deal with partial CSI.

APPENDIX A PROOF OF THEOREM 2

The achieved NE is characterized by the solutions of all N per-user optimization problems (10). Since the individual utility optimization problem is convex, a locally optimal solution is globally optimal. The optimal solution can be found by solving its K.K.T. conditions [26], given by:

$$\begin{aligned} \frac{\partial L_i(\tilde{\mathbf{T}}_i, \alpha_i^{(k)}, \gamma_i)}{\partial \tilde{\mathbf{T}}_i^{(k)*}} &= \frac{\partial R_i^{(k)}}{\partial \tilde{\mathbf{T}}_i^{(k)*}} - \mathbf{A}_i^{(k)} \tilde{\mathbf{T}}_i^{(k)} - (\alpha_i^{(k)} + \gamma_i) \tilde{\mathbf{T}}_i^{(k)} = 0, \forall k \\ \text{tr}(\tilde{\mathbf{T}}_i^{(k)} \tilde{\mathbf{T}}_i^{(k)H}) - P_{\text{mask}}(f_k) &\leq 0, \forall k \in \Psi_K \\ \alpha_i^{(k)} [\text{tr}(\tilde{\mathbf{T}}_i^{(k)} \tilde{\mathbf{T}}_i^{(k)H}) - P_{\text{mask}}] &= 0, \forall k \in \Psi_K \\ \sum_{k \in \Psi_K} \text{tr}(\tilde{\mathbf{T}}_i^{(k)} \tilde{\mathbf{T}}_i^{(k)H}) - P_{\text{max}} &\leq 0 \\ \gamma_i [\sum_{k \in \Psi_K} \text{tr}(\tilde{\mathbf{T}}_i^{(k)} \tilde{\mathbf{T}}_i^{(k)H}) - P_{\text{max}}] &= 0 \end{aligned} \quad (31)$$

The Lagrangian function of the network optimization problem (6) is in (32), where $\tilde{\mathbf{T}} \stackrel{\text{def}}{=} \bigcup_i \tilde{\mathbf{T}}_i$, the set of precoding matrices over all users and frequency bands.

All the stationary (or locally optimal) points of the network

$$L(\tilde{\mathbf{T}}, \alpha_i^{(k)}, \gamma_i) = \sum_{i \in \Phi_N} \sum_{k \in \Psi_K} R_i^{(k)} - \sum_{i \in \Phi_N} \sum_{k \in \Psi_K} \alpha_i^{(k)} [\text{tr}(\tilde{\mathbf{T}}_i^{(k)} \tilde{\mathbf{T}}_i^{(k)H}) - P_{\text{mask}}(f_k)] - \sum_{i \in \Phi_N} \gamma_i [\sum_{k \in \Psi_K} \text{tr}(\tilde{\mathbf{T}}_i^{(k)} \tilde{\mathbf{T}}_i^{(k)H}) - P_{\text{max}}] \quad (32)$$

$$\frac{\partial R_j^{(k)}}{\partial \tilde{\mathbf{T}}_i^{(k)*}} = -\mathbf{H}_{d(j),i}^{(k)H} \mathbf{C}_{d(j)}^{(k)-1} \mathbf{H}_{d(j),i}^{(k)} \tilde{\mathbf{T}}_i^{(k)} + \mathbf{H}_{d(j),i}^{(k)H} (\mathbf{C}_{d(j)}^{(k)} + \mathbf{H}_{d(j),j}^{(k)} \tilde{\mathbf{T}}_j^{(k)} \tilde{\mathbf{T}}_j^{(k)H} \mathbf{H}_{d(j),j}^{(k)})^{-1} \mathbf{H}_{d(j),i}^{(k)} \tilde{\mathbf{T}}_i^{(k)} \quad (33)$$

$$= -\mathbf{H}_{d(j),i}^{(k)H} \mathbf{C}_{d(j)}^{(k)-1} \mathbf{H}_{d(j),j}^{(k)} [(\tilde{\mathbf{T}}_j^{(k)} \tilde{\mathbf{T}}_j^{(k)H})^{-1} + \mathbf{H}_{d(j),j}^{(k)H} \mathbf{C}_{d(j)}^{(k)-1} \mathbf{H}_{d(j),j}^{(k)}]^{-1} \mathbf{H}_{d(j),j}^{(k)H} \mathbf{C}_{d(j)}^{(k)-1} \mathbf{H}_{d(j),i}^{(k)} \tilde{\mathbf{T}}_i^{(k)}$$

$$L_i(\tilde{\mathbf{T}}_i, \alpha_i^{(k)}, \gamma_i) = \sum_{k \in \Psi_K} \{ \log \det(\mathbf{I} + \tilde{\mathbf{T}}_i^{(k)H} \mathbf{H}_{d(i),i}^{(k)H} \mathbf{C}_{d(i)}^{(k)-1} \mathbf{H}_{d(i),i}^{(k)} \tilde{\mathbf{T}}_i^{(k)}) - \text{tr}(\tilde{\mathbf{T}}_i^{(k)H} [\mathbf{A}_i^{(k)} + (\alpha_i^{(k)} + \gamma_i) \mathbf{I}] \tilde{\mathbf{T}}_i^{(k)}) + \alpha_i^{(k)} P_{\text{mask}}(f_k) + \frac{\gamma_i}{K} P_{\text{max}} \} \quad (34)$$

$$L_i(\tilde{\mathbf{T}}_i, \alpha_i^{(k)}, \gamma_i) = \sum_{k \in \Psi_K} \{ \alpha_i^{(k)} P_{\text{mask}}(f_k) + \frac{\gamma_i}{K} P_{\text{max}} - \text{tr}(\tilde{\mathbf{T}}_i^{(k)H} \tilde{\mathbf{T}}_i^{(k)}) + \log \det(\mathbf{I} + \tilde{\mathbf{T}}_i^{(k)H} \mathbf{E}_i^{(k)-1} \mathbf{H}_{d(i),i}^{(k)H} \mathbf{C}_{d(i)}^{(k)-1} \mathbf{H}_{d(i),i}^{(k)} \mathbf{E}_i^{(k)H-1} \tilde{\mathbf{T}}_i^{(k)}) \} \quad (35)$$

$$L'_i = \sum_{k \in \Psi_K} \{ -\text{tr}(\tilde{\mathbf{T}}_i^{(k)H} \tilde{\mathbf{T}}_i^{(k)}) + \log \det(\mathbf{I} + \tilde{\mathbf{T}}_i^{(k)H} \mathbf{E}_i^{(k)-1} \mathbf{H}_{d(i),i}^{(k)H} \mathbf{C}_{d(i)}^{(k)-1} \mathbf{H}_{d(i),i}^{(k)} \mathbf{E}_i^{(k)H-1} \tilde{\mathbf{T}}_i^{(k)}) \} \quad (36)$$

$$L'_i \leq \sum_{k \in \Psi_K} \{ -\text{tr}(\tilde{\mathbf{T}}_i^{(k)H} \tilde{\mathbf{T}}_i^{(k)}) + \sum_{s=1}^M \log(1 + \text{diag}_s \{ \tilde{\mathbf{T}}_i^{(k)H} \mathbf{E}_i^{(k)-1} \mathbf{H}_{d(i),i}^{(k)H} \mathbf{C}_{d(i)}^{(k)-1} \mathbf{H}_{d(i),i}^{(k)} \mathbf{E}_i^{(k)H-1} \tilde{\mathbf{T}}_i^{(k)} \}) \} \quad (37)$$

problem must satisfy its K.K.T. conditions:

$$\begin{aligned} \frac{\partial L(\tilde{\mathbf{T}}, \alpha_i^{(k)}, \gamma_i)}{\partial \tilde{\mathbf{T}}_i^{(k)*}} &= \frac{\partial R_i^{(k)}}{\partial \tilde{\mathbf{T}}_i^{(k)*}} + \sum_{j \in \{\Phi_N \setminus i\}} \frac{\partial R_j^{(k)}}{\partial \tilde{\mathbf{T}}_i^{(k)*}} (\alpha_i^{(k)} + \gamma_i) \tilde{\mathbf{T}}_i^{(k)} = 0 \\ \text{tr}(\tilde{\mathbf{T}}_i^{(k)} \tilde{\mathbf{T}}_i^{(k)H}) - P_{\text{mask}}(f_k) &\leq 0, \forall k \in \Psi_K, \forall i \in \Phi_N \\ \alpha_i^{(k)} [\text{tr}(\tilde{\mathbf{T}}_i^{(k)} \tilde{\mathbf{T}}_i^{(k)H}) - P_{\text{mask}}] &= 0, \forall k \in \Psi_K, \forall i \in \Phi_N \\ \sum_{k \in \Psi_K} \text{tr}(\tilde{\mathbf{T}}_i^{(k)} \tilde{\mathbf{T}}_i^{(k)H}) - P_{\text{max}} &\leq 0, \forall i \in \Phi_N \\ \gamma_i [\sum_{k \in \Psi_K} \text{tr}(\tilde{\mathbf{T}}_i^{(k)} \tilde{\mathbf{T}}_i^{(k)H}) - P_{\text{max}}] &= 0, \forall i \in \Phi_N \end{aligned} \quad (38)$$

To guarantee that the game (10) with the price function defined in (11) converges to a NE at which the CRN's throughput is the same as that of a locally optimal solution to problem (6), the NE of the game (10) must be a stationary point of problem (6). In other words, the K.K.T. conditions of (6) have to hold at the stationary point of (10). For that to happen, the following equality must hold (through comparing conditions (31) and (38)):

$$-\mathbf{A}_i^{(k)} \tilde{\mathbf{T}}_i^{(k)} = \sum_{j \in \{\Phi_N \setminus i\}} \frac{\partial R_j^{(k)}}{\partial \tilde{\mathbf{T}}_i^{(k)*}}. \quad (39)$$

To compute $\frac{\partial R_j^{(k)}}{\partial \tilde{\mathbf{T}}_i^{(k)*}}$ in (33), recall (2) and note that:

$$\begin{aligned} \mathbf{C}_{d(j)}^{(k)} &= \mathbf{I} + \mathbf{H}_{d(j),i}^{(k)} \tilde{\mathbf{T}}_i^{(k)} \tilde{\mathbf{T}}_i^{(k)H} \mathbf{H}_{d(j),i}^{(k)H} \\ &+ \sum_{v \in \Phi_N \setminus \{i,j\}} \mathbf{H}_{d(j),v}^{(k)} \tilde{\mathbf{T}}_v^{(k)} \tilde{\mathbf{T}}_v^{(k)H} \mathbf{H}_{d(j),v}^{(k)H}. \end{aligned}$$

The last equality in (33) follows by applying the Woodbury identity [30] to $(\mathbf{C}_{d(j)}^{(k)} + \mathbf{H}_{d(j),j}^{(k)} \tilde{\mathbf{T}}_j^{(k)} \tilde{\mathbf{T}}_j^{(k)H} \mathbf{H}_{d(j),j}^{(k)H})^{-1}$. Plugging (33) into (39), we get (13). One can also easily verify that the derived $\mathbf{A}_i^{(k)}$ matrix is positive-semidefinite. Additionally, if the pricing-factor has the form (13), the achieved NE meets the K.K.T. conditions of the network-wide problem (6). \square

APPENDIX B

PROOF OF THEOREM 3

Let's rewrite the Lagrangian function of (10) as in (34), then using the Cholesky decomposition $[\mathbf{A}_i^{(k)} + (\alpha_i^{(k)} + \gamma_i) \mathbf{I}] = \mathbf{E}_i^{(k)} \mathbf{E}_i^{(k)H}$. We have (35), where $\tilde{\mathbf{T}}_i^{(k)H} = \tilde{\mathbf{T}}_i^{(k)H} \mathbf{E}_i^{(k)}$.

As P_{max} and $P_{\text{mask}}(f_k)$ are predetermined values and from (35), L_i is maximized if we maximize L'_i in (36). Following the routine of using Hadamard inequality (e.g., [9], [19]), applying the Hadamard's inequality [31] to the second term of (36), we have (37), where $\text{diag}_s(\cdot)$ is the (s, s) diagonal element of a matrix (\cdot) .

The equality happens when $\tilde{\mathbf{T}}_i^{(k)H} \mathbf{E}_i^{(k)-1} \mathbf{H}_{d(i),i}^{(k)H} \mathbf{C}_{d(i)}^{(k)-1} \mathbf{H}_{d(i),i}^{(k)} \mathbf{E}_i^{(k)H-1} \tilde{\mathbf{T}}_i^{(k)}$ is a diagonal matrix. This is the case if there exists an orthonormal matrix $\tilde{\mathbf{T}}_i^{(k)}$ that diagonalizes the matrix $\mathbf{E}_i^{(k)-1} \mathbf{H}_{d(i),i}^{(k)H} \mathbf{C}_{d(i)}^{(k)-1} \mathbf{H}_{d(i),i}^{(k)} \mathbf{E}_i^{(k)H-1}$. Hence, we should have:

$$\begin{aligned} \tilde{\mathbf{T}}_i^{(k)H} \tilde{\mathbf{T}}_i^{(k)} &= \mathbf{I} \\ \tilde{\mathbf{T}}_i^{(k)-1} \mathbf{E}_i^{(k)-1} \mathbf{H}_{d(i),i}^{(k)H} \mathbf{C}_{d(i)}^{(k)-1} \mathbf{H}_{d(i),i}^{(k)} \mathbf{E}_i^{(k)H-1} \tilde{\mathbf{T}}_i^{(k)} &= \mathbf{\Lambda}_i^{(k)} \end{aligned} \quad (40)$$

where $\mathbf{\Lambda}_i^{(k)}$ is a $M \times M$ diagonal matrix.

Multiplying both sides of (40) by $\tilde{\mathbf{T}}_i^{(k)}$, then $\mathbf{E}_i^{(k)}$, we have:

$$\mathbf{H}_{d(i),i}^{(k)H} \mathbf{C}_{d(i)}^{(k)-1} \mathbf{H}_{d(i),i}^{(k)} \mathbf{E}_i^{(k)H-1} \tilde{\mathbf{T}}_i^{(k)} = \mathbf{E}_i^{(k)} \tilde{\mathbf{T}}_i^{(k)} \mathbf{\Lambda}_i^{(k)}.$$

Recall that $\tilde{\mathbf{T}}_i^{(k)H} = \tilde{\mathbf{T}}_i^{(k)H} \mathbf{E}_i^{(k)}$ and the above Cholesky decomposition we have:

$$\begin{aligned} \mathbf{H}_{d(i),i}^{(k)H} \mathbf{C}_{d(i)}^{(k)-1} \mathbf{H}_{d(i),i}^{(k)} \tilde{\mathbf{T}}_i^{(k)} &= \mathbf{E}_i^{(k)} \mathbf{E}_i^{(k)H} \tilde{\mathbf{T}}_i^{(k)} \mathbf{\Lambda}_i^{(k)} \\ &= [\mathbf{A}_i^{(k)} + (\alpha_i^{(k)} + \gamma_i) \mathbf{I}] \tilde{\mathbf{T}}_i^{(k)} \mathbf{\Lambda}_i^{(k)}. \end{aligned}$$

This concludes the proof. \square

REFERENCES

- [1] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, May 2005.

- [2] K. Sundaresan, R. Sivakumar, M. Ingram, and T.-Y. Chang, "A fair medium access control protocol for ad-hoc networks with MIMO links," in *Proceedings of the INFOCOM Conference*, vol. 4, March 2004, pp. 2559–2570.
- [3] S. Chu and X. Wang, "Opportunistic and cooperative spatial multiplexing in MIMO ad hoc networks," in *Proceedings of the Mobihoc Conference*, 2008, pp. 63–72.
- [4] G. Scutari and D. Palomar, "MIMO cognitive radio: A game theoretical approach," *IEEE Transactions on Signal Processing*, vol. 58, no. 2, pp. 761–780, Feb. 2010.
- [5] Z.-Q. Luo and S. Zhang, "Dynamic spectrum management: Complexity and duality," *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, no. 1, pp. 57–73, Feb. 2008.
- [6] Y. J. Zhang and A. So, "Optimal spectrum sharing in MIMO cognitive radio networks via semidefinite programming," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 2, pp. 362–373, Feb. 2011.
- [7] R. Zhang and Y.-C. Liang, "Exploiting multi-antennas for opportunistic spectrum sharing in cognitive radio networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, no. 1, pp. 88–102, Feb. 2008.
- [8] J. Wang, G. Scutari, and D. Palomar, "Robust MIMO cognitive radio via game theory," *IEEE Transactions on Signal Processing*, vol. 59, no. 3, pp. 1183–1201, March 2011.
- [9] S.-J. Kim and G. B. Giannakis, "Optimal resource allocation for MIMO ad hoc cognitive radio networks," *IEEE Transactions on Information Theory*, vol. 57, no. 5, pp. 3117–3131, May 2011.
- [10] D. Palomar and J. Fonollosa, "Practical algorithms for a family of waterfilling solutions," *IEEE Transactions on Signal Processing*, vol. 53, no. 2, pp. 686–695, Feb. 2005.
- [11] G. Scutari, D. Palomar, and S. Barbarossa, "Asynchronous iterative water-filling for gaussian frequency-selective interference channels," *IEEE Transactions on Information Theory*, vol. 54, no. 7, pp. 2868–2878, July 2008.
- [12] X. Dong, Y. Rong, and Y. Hua, "Cooperative power scheduling for a network of MIMO links," *IEEE Transactions on Wireless Communications*, vol. 9, no. 3, pp. 939–944, March 2010.
- [13] G. Scutari, D. Palomar, and S. Barbarossa, "Competitive design of multiuser MIMO systems based on game theory: A unified view," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 7, pp. 1089–1103, Sept. 2008.
- [14] F. Wang, M. Krunz, and S. Cui, "Price-based spectrum management in cognitive radio networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, no. 1, pp. 74–87, 2008.
- [15] J. Hirshleifer, A. Glazer, and D. Hirshleifer, *Price Theory and Applications Decisions, Markets, and Information*. Cambridge University Press, 2005.
- [16] M. J. Osborne, *An Introduction to Game Theory*. Oxford University Press, 2004.
- [17] R. Iltis, S.-J. Kim, and D. Hoang, "Noncooperative iterative MMSE beamforming algorithms for ad hoc networks," *IEEE Transactions on Communications*, vol. 54, no. 4, pp. 748–759, 2006.
- [18] C. Shi, R. Berry, and M. Honig, "Local interference pricing for distributed beamforming in MIMO networks," in *Proceedings of the IEEE MILCOM Conference*, 2009, pp. 1–6.
- [19] D. Hoang and R. Iltis, "Noncooperative eigencoding for MIMO ad hoc networks," *IEEE Transactions on Signal Processing*, vol. 56, no. 2, pp. 865–869, 2008.
- [20] FCC, "Spectrum policy task force report," *ET Docket No. 02-380 and No. 04-186*, Sep. 2010.
- [21] W. Yu, "Competition and cooperation in multi-user communication environments," *PhD Thesis, Stanford University, Stanford, CA*, 2002.
- [22] W. Wang, Y. Cui, T. Peng, and W. Wang, "Noncooperative power control game with exponential pricing for cognitive radio network," in *Proceedings of the IEEE 65th Vehicular Technology Conference*, April 2007.
- [23] G. Arslan, M. Demirkol, and S. Yuksel, "Power games in MIMO interference systems," in *Proceedings of the International Conference on Game Theory for Networks*, May 2009, pp. 52–59.
- [24] J.-S. Pang, G. Scutari, D. Palomar, and F. Facchinei, "Design of cognitive radio systems under temperature-interference constraints: A variational inequality approach," *IEEE Transactions on Signal Processing*, vol. 58, no. 6, pp. 3251–3271, June 2010.
- [25] D. Nguyen and M. Krunz, "Spectrum management and power allocation in MIMO cognitive networks," University of Arizona, Tech. Rep. TR-UA-ECE-2011-2, August 2011. [Online]. Available: <http://www.ece.arizona.edu/~krunz>
- [26] D. P. Bertsekas, *Nonlinear Programming*. Athena Scientific, 1995.
- [27] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 1990.
- [28] R. Maheswaran and T. Basar, "Decentralized network resource allocation as a repeated noncooperative market game," in *Proceedings of the 40th IEEE Conference on Decision and Control*, vol. 5, 2001, pp. 4565–4570.
- [29] G. Scutari, D. Palomar, F. Facchinei, and J.-S. Pang, "Convex optimization, game theory, and variational inequality theory," *IEEE Signal Processing Magazine*, pp. 35–47, May 2010.
- [30] W. W. Hager, "Updating the inverse of a matrix," *SIAM Review. A Publication of the Society for Industrial and Applied Mathematics*, vol. 31, no. 2, pp. 221–239, 1989.
- [31] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*. Academic Press, 1979.



Diep N. Nguyen received the B.S. degree in telecommunications and electronics from Posts and Telecommunications Institute of Technology (PTIT), Vietnam, in 2004, and the M.E. degree in electrical and computer engineering from the University of California, San Diego in 2008. From 2004 to 2006, he was a lecturer at the Department of Telecommunications, PTIT, Vietnam. He was a member of technical staff at Broadcom Corporation from 2008 to 2009. He is currently working towards the Ph.D. degree at the Department of Electrical and Computer Engineering, The University of Arizona. His research interests include cognitive radios, ad hoc, and sensor networks, with emphasis on resource allocation and network management for multi-antenna systems.



Marwan Krunz received the PhD degree in electrical engineering from Michigan State University in 1995. He has been with the University of Arizona since 1997, where he is a professor in the Electrical and Computer Engineering and Computer Science Departments. He is the UA site director for Connection One. He has held several visiting research positions and was a Chair of Excellence at the University of Carlos III, Madrid, Spain. His research interests include computer networking and wireless communications with a focus on distributed radio resource management in wireless and sensor networks, protocol design, and secure communications. He has published more than 170 journal articles and refereed conference papers and is a coinventor on three US patents. He was a recipient of the US National Science Foundation (NSF) CAREER Award in 1998. He has served on the editorial boards for the IEEE Transactions on Network and Service Management, IEEE/ACM Transactions on Networking, IEEE Transactions on Mobile Computing, and Computer Communications Journal, and as a TPC chair for several international conferences. He is a fellow of the IEEE.