Energy Efficient Transmissions in Cognitive MIMO Systems With Multiple Data Streams

Liqun Fu, *Member, IEEE*, Mikael Johansson, *Member, IEEE*, and Mats Bengtsson, *Senior Member, IEEE*

***Abstract*—We investigate energy-efficient communications for time-division multiple access (TDMA) multiple-input multiple-output (MIMO) cognitive radio (CR) networks operating in underlay mode. In particular, we consider the joint optimization over both the time resource and the transmit precoding matrices to minimize the overall energy consumption of a single cell secondary network with multiple secondary users (SUs), while ensuring their quality of service (QoS). The corresponding mathematical formulations turn out to be non-convex, and thus of high complexity to solve in general. We give a comprehensive treatment of this problem, considering both the cases of perfect channel state information (CSI) and statistical CSI of the channels from the SUs to the primary receiver. We tackle the non-convexity by applying a proper optimization decomposition that allows the overall problem to be efficiently solved. In particular, we show that when the SUs only have statistical CSI, the optimal solution can be found in polynomial time. Moreover, if we consider additional integer constraints on the time variable which is usually a requirement in practical wireless system, the overall problem becomes a mixed-integer non-convex optimization which is more complicated. By exploring the special structure of this particular problem, we show that the optimal integer time solution can be obtained in polynomial time with a simple greedy algorithm. When the SUs have perfect CSI, the decomposition based algorithm is guaranteed to find the optimal solution when the secondary system is *under-utilized*. Simulation results show that the energy-optimal transmission scheme adapts to the traffic load of the secondary system to create a win-win situation where the SUs are able to decrease the energy consumption and the PUs experience less interference from the secondary system. The effect is particularly pronounced when the secondary system is under-utilized.**

***Index Terms*—Cognitive radio networks, energy consumption, resource allocation, MIMO, precoding.**

I. Introduction

FUTURE wireless systems are evolving to support the exponentially increasing traffic demands, which, in most cases, is achieved at the expense of a higher energy consumption and a considerable impact on the environment. Energy-efficient transmission is of critical importance to reduce the carbon footprint and to prolong the battery lifetime of wireless devices [1]–[8]. In this paper, we are interested in energy-efficient transmissions in a MIMO cognitive radio (CR) network. The cognitive radio is to allow the secondary users (SUs) to opportunistically use the spectrum that is assigned to the licensed primary users (PUs) in order to alleviate spectrum scarcity [9]. The MIMO techniques open a new dimension of "space" to enable the SUs to coexist with the PUs without causing harmful interference to the primary links [10]–[12]. Apart from the above advantage, performing energy-efficient transmissions among the SUs could also alleviate the interference to the primary system.

The key idea that enables the underlay mode in a CR network is that with multiple antennae, the SU can carefully design its precoding matrix so as to suppress the interference at the primary receivers. Such a technique normally requires channel state information (CSI), which can be assumed to be available in the traditional MIMO network setup. However, in a CR network, the PUs are usually not aware of the existence of the SUs. This has two effects: 1) the primary receivers may not feed back the CSI to the SUs; 2) the primary receiver will not perform interference cancellation at the receiver side. In other words, the SUs are only able to do pre-interference suppression, and may have to do this with only access to statistical CSI. This is the main challenge when designing energy-efficient transmission strategies for MIMO CR networks.

Studies of energy-efficient transmissions of MIMO networks fall into two main classes: traditional MIMO networks [1]–[5] and MIMO CR networks [7], [8]. For example, in the former class, [1] investigated the energy consumption of a single MIMO/SISO link under different link distances. In [3], the authors considered power minimization through downlink transmit beamforming and solved it with semidefinite programming (SDP) approach. As discussed above, perfect CSI is typically assumed in traditional MIMO networks. In the latter class, some recent papers, [7], [8], considered the energy/power minimization problem in MIMO CR networks. In particular, multicast precoding where a secondary transmitter communicates with multiple secondary receivers under statistical CSI is considered. The problem was shown to be non-convex [13], and an SDP relaxation was proposed. Since there is only one secondary transmitter, it is sufficient to optimize the SU's precoding matrix only at the physical layer in [7], [8].

A closely related problem to energy minimization is rate maximization. Rate maximization in MIMO CR networks has been considered in [14]–[19]. In [14], Zhang *et al.* showed that

Manuscript received September 16, 2014; revised February 8, 2015; accepted April 30, 2015. Date of publication May 18, 2015; date of current version September 7, 2015. This work was supported in part by Swedish Research Council under Contracts 621-2012-4073, 621-2012-4134 and 621-2012-5325, a postdocoral fellowship from the ACCESS Linnaeus Center, the National Natural Science Foundation of China under Project No. 61401380, and the Macao Science and Technology Development Fund under Grant No. 096/2013/A3. The associate editor coordinating the review of this paper and approving it for publication was K. Huang.

The authors are with the ACCESS Linnaeus Centre, KTH Royal Institute of Technology, Stockholm, 100 44, Sweden (e-mail: liqun@kth.se; mikaelj@kth.se; mats.bengtsson@ee.kth.se).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TWC.2015.2434372

rate maximization for a single secondary link under perfect CSI and no interference from the primary system to the secondary system is a convex optimization problem. Practical algorithms based on the singular-value decomposition of the SU's MIMO channel matrix were proposed. In [17], weighted sum-rate maximization of the multiple-access channel in a MIMO CR network was investigated and a capped multi-level water-filling algorithm was proposed. In [19], the authors considered maximizing the rate of a single secondary link under different levels of CSI availability of the channels from the SUs to the primary receivers. A unified homogeneous QCQP formulation was proposed, and an SDP relaxation was shown to produce the optimal solution in some special cases.

This paper considers the joint use of physical and MAC layer techniques to minimize the overall energy consumption of the secondary network with multiple SUs. With the aid of MIMO techniques, it is possible that multiple SUs send uplink traffic to the secondary base station (BS) simultaneously with spatial multiplexing. However, in a CR network, the interference power at each secondary receiver does not only come from the primary users but also from other SUs. The MIMO interference channel is well known to result in NP-hard problems [20]. Therefore, in order to avoid excessive interference among the SUs, we focus on the case that the SUs send their traffic via TDMA. Unlike existing works on instantaneous power minimization [7], [8], we aim to jointly optimize the time allocation and the transmit precoding matrices while satisfying the SUs' rate requirements. As will be elaborated in Section III, the problem formulations are non-convex, and thus can be expected to be hard to solve in general. We give a comprehensive investigation on this problem, under different assumptions of CSI knowledge. Quite surprisingly, we are able to develop efficient algorithms that are guaranteed to find the optimal solutions in many scenarios.

When the secondary system has statistical CSI, we adopt a probabilistic interference constraint, i.e., ensuring that the interference at the primary receiver is below the threshold with high probability [21]. Such probabilistic constraints are practical, as many wireless applications (such as video streaming, voice over IP, etc.) can tolerate occasional outages without affecting user QoS. To tackle the non-convexity, we apply decomposition method that subdivides the overall problem into two separate problems that can be solved efficiently. As a result, the optimal time allocation and the transmit precoding matrices can be found in polynomial time. In particular, the optimal time allocation can be found by solving a special convex optimization problem which is not second-order differentiable. In this case, we apply a first-order method with fast convergence rate. Given the optimal time allocation, the optimal precoding matrix can be found by "water-filling". Moreover, in practical wireless systems, a slot is the smallest unit during time allocation, which implies that there is additional integer constraint on the time variable. The overall problem formulation then becomes a mixed-integer non-convex optimization problem, which is typically even more complicated. By exploring the special structure of the time optimization problem, we show that the optimal integer time allocation can be found in polynomial time with a simple greedy algorithm.

When the SUs have perfect CSI, we propose different methods which depend on the traffic load to the secondary system. When the secondary system is *under-utilized*, the optimal solution to the non-convex problem can still be found through the decomposition method. It is only in the case where the secondary system is *heavily-utilized* and the SUs have perfect CSI that we need to resort to a heuristic algorithm. In this case, we show that the proposed iterative algorithm is guaranteed to converge to a feasible solution, albeit not necessarily optimal.

This paper is related to our previous work, [22], which considered the special case where each SU transmits only one data stream on all its antennae. Such transmission strategies do not fully utilize the multiplexing gain of MIMO systems and, are obviously not optimal in general. When multiple data stream transmission is considered, the number of data streams is discrete by nature. This adds significant complexity in both the problem formulation and the solution techniques.

The remainder of this paper is organized as follows. In Section II, we describe the system model. Section III lays out the problem formulations in both scenarios with perfect and statistical CSI. Section IV is devoted to the optimization decomposition and the solution method under the statistical CSI scenario. In Section V, we focus on the solution methods under perfect CSI scenario. In Section VI, we present the simulation results. Section VII concludes this paper.

II. System Model

We consider a CR network with K SUs and J PUs. The primary links could potentially always be active, and thus need to be protected at all times. The primary network is composed of J pairs of transmitters and receivers. The secondary system is a single cell network, where the SUs send uplink traffic to the same secondary BS via TDMA. The uplink transmissions are synchronized by the secondary BS so that they are allocated different time slots for their transmissions and thus do not cause interference to each other.

We use S_k to denote the kth SU. Let M_{S_k} denote the number of transmit antennas of S_k and N_{BS} denote the number of receive antennas at the secondary BS. Let $\mathbf{H}_{BS,S_k} \in \mathbb{C}^{N_{BS}\times M_{S_k}}$ denote the channel matrix from S_k to the secondary BS. We use P_j to denote the jth primary transmitter-receiver pair. Let M_{P_j} and N_{P_j} denote the number of transmit antennas and the number of receive antennas of P_j, respectively. Since the SUs coexist with the PUs, their signals may interfere with each other. Let $\mathbf{H}_{P_j,S_k} \in \mathbb{C}^{N_{P_j}\times M_{S_k}}$ and $\mathbf{H}_{BS,P_j} \in \mathbb{C}^{N_{BS}\times M_{P_j}}$ denote the channel matrix from S_k to the receiver of P_j and the channel matrix from the transmitter of P_j to the secondary BS, respectively. We assume a frequency flat fading channel so that the channel is the same for the considered bandwidth. Furthermore, we assume block fading channels, so that the channel matrices do not change during a TDMA frame,[1] and the channel realizations in different frames are uncorrelated. Since the secondary system is centralized, the secondary BS can estimate $\mathbf{H}_{BS,S_k}$ and feed back it to each S_k with a separate control channel. Thus, it is

[1]This assumption is valid if the mobile user does not move very fast. In this case, the coherence time is long enough to cover the whole TDMA frame.

reasonable to assume that $\mathbf{H}_{BS,S_k}$ is known to both S_k and the secondary BS.

Both the primary and secondary users can transmit multiple data streams. Let D_{S_k} and D_{P_j} denote the number of data streams of S_k and P_j, respectively. Let $\mathbf{x}_{S_k} \in \mathbb{C}^{M_{S_k}\times 1}$ and $\mathbf{x}_{P_j} \in \mathbb{C}^{M_{P_j}\times 1}$ denote the actual transmitted vectors of S_k and P_j, respectively. The covariance matrices of $\mathbf{x}_{S_k}$ and $\mathbf{x}_{P_j}$ are denoted by $\mathbf{Q}_{S_k}$ and $\mathbf{Q}_{P_j}$, which are Hermitian positive semidefinite matrices.

The received vector of S_k at the secondary BS is

$$\mathbf{y}_{BS_k} = \mathbf{H}_{BS,S_k}\mathbf{x}_{S_k} + \sum_{j=1}^{J}\mathbf{H}_{BS,P_j}\mathbf{x}_{P_j} + \mathbf{n}_{BS},\ k = 1, \cdots, K.$$

The vector $\mathbf{n}_{BS} \in \mathbb{C}^{N_{BS}\times 1}$ is a circular complex additive Gaussian noise vector with a noise power of $N_0 w$ at the secondary BS, where $N_0/2$ is the noise power spectral density and w is the bandwidth used in the secondary system. We assume that the secondary BS treats the interference from the primary transmitters as noise, and that there is no successive interference cancellation at the secondary BS. The interference-plus-noise covariance matrix at the secondary BS when S_k transmits is then

$$\mathbf{C}_{S_k} = \sum_{j=1}^{J}\mathbf{H}_{BS,P_j}\mathbf{Q}_{P_j}\mathbf{H}_{BS,P_j}^{H} + N_0 w\mathbf{I}_{N_{BS}},$$

which is an $N_{BS} \times N_{BS}$ Hermitian positive semidefinite matrix.

According to Shannon's capacity formula for a MIMO link [10], [23], the achievable transmission rate of S_k is

$$r_{S_k} = w\log\left[\det\left(\mathbf{I} + \mathbf{H}_{BS,S_k}\mathbf{Q}_{S_k}\mathbf{H}_{BS,S_k}^{H}\mathbf{C}_{S_k}^{-1}\right)\right], \quad k = 1, \cdots, K.$$

Here r_{S_k} is the instantaneous transmission rate (in nats/second) when S_k is active. The total transmit power of S_k on all its transmit antennas is $p_{S_k} = \mathrm{tr}(\mathbf{Q}_{S_k})$, and S_k causes a total interference power to the jth primary receiver at the level of

$$q_{P_j,S_k} = \mathrm{tr}\left(\mathbf{H}_{P_j,S_k}\mathbf{Q}_{S_k}\mathbf{H}_{P_j,S_k}^{H}\right),\ k = 1, \cdots, K,\ j = 1, \cdots, J.$$

III. Problem Formulations

The system target is to choose the proper time allocation and the transmit precoding matrix for each SU to minimize the total energy consumption of all the SUs while protecting the PUs and ensuring a minimum QoS for each SU. Specifically, the interference from each SU to each of the PUs need to be below a certain threshold, and each SU has a rate requirement R_{S_k} to be satisfied. Here R_{S_k} (with the unit of nats/frame) is the number of nats that S_k needs to transmit in each time frame. Without loss of generality, the TDMA frame length of the secondary system is normalized to be 1. Each S_k is allocated a time fraction t_{S_k} $(0 \le t_{S_k} \le 1)$ to transmit its data. The instantaneous transmit power of S_k is limited by a maximum power of $p_{S_k,\max}$. This problem can be mathematically formulated as follows:

$$\begin{aligned}
\min_{t_{S_k},\mathbf{Q}_{S_k}} \quad & \sum_{k=1}^{K} t_{S_k}\mathrm{tr}(\mathbf{Q}_{S_k}) \\
\text{s.t.} \quad & t_{S_k} w\log\left[\det\left(\mathbf{I} + \mathbf{H}_{BS,S_k}\mathbf{Q}_{S_k}\mathbf{H}_{BS,S_k}^{H}\mathbf{C}_{S_k}^{-1}\right)\right] \\
& \qquad \ge R_{S_k}, \qquad \forall k, & \text{(1a)} \\
& \sum_{k=1}^{K} t_{S_k} \le 1, & \text{(1b)} \\
& \mathrm{tr}\left(\mathbf{H}_{P_j,S_k}\mathbf{Q}_{S_k}\mathbf{H}_{P_j,S_k}^{H}\right) \le \phi_{P_j}, \qquad \forall k, \qquad \forall j, & \text{(1c)} \\
& \mathrm{tr}\left(\mathbf{Q}_{S_k}\right) \le p_{S_k,\max}, \qquad \forall k, & \text{(1d)} \\
& t_{S_k} \ge 0, \qquad \forall k, \\
& \mathbf{Q}_{S_k} \succeq \mathbf{0}, \qquad \forall k.
\end{aligned}$$

The objective function is the total energy consumption of all SUs. Constraint (1a) guarantees the rate requirement for each SU. Constraint (1b) ensures that the total time allocated to all the SUs is no larger than the TDMA frame length. Since the secondary network is a TDMA network, the SUs do not transmit simultaneously. The interference constraint to the primary network ensures that the interference from each secondary transmitter to each primary receiver is no larger than the threshold ϕ_{P_j}, as shown in (1c). Note that since the secondary system does not know the receive beamforming at each primary receiver, the interference power in (1c) is the interference power at the antennas of each primary receiver. Constraint (1d) states that each SU has limited transmit power. The last two constraints state that each t_{S_k} is non-negative, and each $\mathbf{Q}_{S_k}$ is positive semidefinite. Problem (1) is non-convex due to both the objective function and Constraint (1a), and is thus in general difficult to solve.

It can be shown that the rank of the optimal covariance matrix to Problem (1) is never higher than the corresponding channel, i.e., $\mathrm{rank}(\mathbf{Q}_{S_k}^{*}) \le \mathrm{rank}(\mathbf{H}_{BS,S_k}^{H}\mathbf{C}_{S_k}^{-1}\mathbf{H}_{BS,S_k}) \le \mathrm{rank}(\mathbf{H}_{BS,S_k})$. Suppose there is one optimal $\mathbf{Q}_{S_k}^{*}$ with a higher rank than $\mathrm{rank}(\mathbf{H}_{BS,S_k})$. We can obtain a new solution of the transmit covariance matrix by projecting $\mathbf{Q}_{S_k}^{*}$ to the row space of $\mathbf{H}_{BS,S_k}^{H}\mathbf{C}_{S_k}^{-1}\mathbf{H}_{BS,S_k}$. The new solution obtained by the projection satisfies Constraints (1a) and (1c). Further, it reduces the LHS of Constraint (1d) and the objective function. This contradicts with that $\mathbf{Q}_{S_k}^{*}$ with a higher rank than $\mathrm{rank}(\mathbf{H}_{BS,S_k})$ is the optimal solution. Therefore, we do not need to impose rank constraint on $\mathbf{Q}_{S_k}$ in Problem (1).

In a CR network, some parameters in (1) may be available and some may not. As discussed earlier, the channel matrix $\mathbf{H}_{BS,S_k}$ is available to both S_k and the secondary BS. Furthermore, as the secondary system usually is aware of the existence of the primary system, we can assume that the secondary BS can overhear the transmissions on the primary links. Therefore, it is able to estimate the interference-plus-noise covariance matrix $\mathbf{C}_{S_k}$ that comes from all the PUs. However, the secondary system is usually transparent to the primary system and the primary system may not deliberately provide the CSI to the secondary system. Therefore, the secondary system may not be able to obtain the channel matrix $\mathbf{H}_{P_j,S_k}$ in Constraint (1c). To this end, we consider two scenarios in this paper:

- Statistical CSI: the secondary system knows the statistics of $\mathbf{H}_{P_j,S_k}$ (e.g., the type of distribution and $\mathbb{E}[\mathbf{H}_{P_j,S_k}\mathbf{H}_{P_j,S_k}^H]$). However, it does not know the precise realization of $\mathbf{H}_{P_j,S_k}$.
- Perfect CSI: the secondary system has perfect knowledge of $\mathbf{H}_{P_j,S_k}$.

The problem formulation for the perfect CSI scenario is given in Problem (1), while the problem formulation for statistical CSI scenario is formalized in the next section.

A. *Formulation Recast for the Statistical CSI Scenario*

In the statistical CSI scenario, the secondary system is not able to know the realization of $\mathbf{H}_{P_j,S_k}$. The requirement of satisfying Constraint (1c) using a fixed guess of the channel matrix would easily lead to suboptimal or even infeasible solutions. Interestingly, many wireless applications (such as video streaming, voice over IP, etc.) can tolerate occasional outages without affecting the QoS. Thus, we consider a more realistic requirement, which is to satisfy the interference constraints with a high probability. In other words, the CR network allows the interference from each secondary transmitter to each primary receiver to exceed the power threshold ϕ_{P_j} with a small outage probability δ_{P_j}. Constraint (1c) is then replaced by

$$\Pr_{\mathbf{H}_{P_j,S_k}}\left\{\mathrm{tr}\left(\mathbf{H}_{P_j,S_k}\mathbf{Q}_{S_k}\mathbf{H}_{P_j,S_k}^H\right)\le\phi_{P_j}\right\}\ge 1-\delta_{P_j},\quad\forall k,\quad\forall j,\quad(2)$$

where the probability is taken over $\mathbf{H}_{P_j,S_k}$.

In particular, we consider Rayleigh fading channels and a rich scattering environment under the statistical CSI scenario, so that the entries of $\mathbf{H}_{P_j,S_k}$ are independently and identically distributed (i.i.d.) complex Gaussian random variables with zero mean and a variance of β_{P_j,S_k} [24], where β_{P_j,S_k} denotes the path loss from S_k to the jth primary receiver. We assume that β_{P_j,S_k} is known to S_k. Given the Rayleigh distribution of the channels, we can obtain that $\mathrm{tr}(\mathbf{H}_{P_j,S_k}\mathbf{Q}_{S_k}\mathbf{H}_{P_j,S_k}^H)$ follows an exponential distribution with the parameter $\frac{1}{\beta_{P_j,S_k}\mathrm{tr}(\mathbf{Q}_{S_k})}$ [19]:

$$\Pr_{\mathbf{H}_{P_j,S_k}}\left\{\mathrm{tr}\left(\mathbf{H}_{P_j,S_k}\mathbf{Q}_{S_k}\mathbf{H}_{P_j,S_k}^H\right)\le\phi_{P_j}\right\} = 1-\exp\left(-\frac{\phi_{P_j}}{\beta_{P_j,S_k}\mathrm{tr}\left(\mathbf{Q}_{S_k}\right)}\right).$$

Thus, the outage probability constraint (2) is equivalent to

$$\mathrm{tr}\left(\mathbf{Q}_{S_k}\right)\le\frac{-\phi_{P_j}}{\beta_{P_j,S_k}\log\delta_{P_j}},\qquad\forall k,\ \forall j.\qquad(3)$$

Furthermore, after converting the outage probability constraint to (3), we notice that it has the same form as the transmit power constraints, and can be combined with Constraint (1d). Let $\rho_{S_k}=\min\left\{\frac{-\phi_{P_1}}{\beta_{P_1,S_k}\log\delta_{P_1}},\cdots,\frac{-\phi_{P_J}}{\beta_{P_J,S_k}\log\delta_{P_J}},p_{S_k,\max}\right\}$. Constraints (3) and (1d) are equivalent to

$$\mathrm{tr}\left(\mathbf{Q}_{S_k}\right)\le\rho_{S_k},\qquad\forall k.$$

Therefore, in the statistical CSI scenario, the problem formulation can be recast as follows:

$$\begin{aligned}
\min_{t_{S_k},\mathbf{Q}_{S_k}}\quad & \sum_{k=1}^{K}t_{S_k}\mathrm{tr}\left(\mathbf{Q}_{S_k}\right)\\
\text{s.t.}\quad & t_{S_k}w\log\left[\det\left(\mathbf{I}+\mathbf{H}_{BS,S_k}\mathbf{Q}_{S_k}\mathbf{H}_{BS,S_k}^H\mathbf{C}_{S_k}^{-1}\right)\right]\\
& \quad\ge R_{S_k},\qquad\forall k, && (4a)\\
& \sum_{k=1}^{K}t_{S_k}\le 1, && (4b)\\
& \mathrm{tr}\left(\mathbf{Q}_{S_k}\right)\le\rho_{S_k},\qquad\forall k, && (4c)\\
& t_{S_k}\ge 0,\qquad\forall k,\\
& \mathbf{Q}_{S_k}\succeq\mathbf{0},\qquad\forall k.
\end{aligned}$$

Note that similar to Problem (1), we do not need to add rank constraint on $\mathbf{Q}_{S_k}$ in Problem (4). Furthermore, Problem (4) is also a non-convex optimization problem. It is challenging to solve the non-convex Problems (1) and (4) directly. As can be seen in the subsequent sections, we will tackle this difficulty by finding a closed-form solution for $\mathbf{Q}_{S_k}$ and thereby reducing (4) to a convex problem in t_{S_k} only. As a result, in the statistical CSI scenario, we can find optimal solutions to Problem (4); in the perfect CSI scenario, we can find optimal solutions to Problem (1) when the secondary system is under-utilized.

B. *Feasibility*

The feasible set in Problem (1) (or (4)) may not always be non-empty. For each S_k, its maximum feasible instantaneous transmission rate $r_{S_k,\max}$, with the unit of nats/second, depends on its maximum transmit power and the interference constraints at the primary receivers. In the statistical CSI scenario, the maximum link rate for S_k can be obtained by solving

$$\begin{aligned}
\max_{\mathbf{Q}_{S_k}}\quad & w\log\left[\det\left(\mathbf{I}+\mathbf{H}_{BS,S_k}\mathbf{Q}_{S_k}\mathbf{H}_{BS,S_k}^H\mathbf{C}_{S_k}^{-1}\right)\right]\\
\text{s.t.}\quad & \mathrm{tr}\left(\mathbf{Q}_{S_k}\right)\le\rho_{S_k},\\
& \mathbf{Q}_{S_k}\succeq\mathbf{0}. && (5)
\end{aligned}$$

Problem (5) can be solved with standard "water-filling" [10], [23]. In the perfect CSI scenario, the maximum link rate for S_k can be obtained by solving the following problem

$$\begin{aligned}
\max_{\mathbf{Q}_{S_k}}\quad & w\log\left[\det\left(\mathbf{I}+\mathbf{H}_{BS,S_k}\mathbf{Q}_{S_k}\mathbf{H}_{BS,S_k}^H\mathbf{C}_{S_k}^{-1}\right)\right]\\
\text{s.t.}\quad & \mathrm{tr}\left(\mathbf{H}_{P_j,S_k}\mathbf{Q}_{S_k}\mathbf{H}_{P_j,S_k}^H\right)\le\phi_{P_j},\qquad\forall j,\\
& \mathrm{tr}\left(\mathbf{Q}_{S_k}\right)\le p_{S_k,\max},\\
& \mathbf{Q}_{S_k}\succeq\mathbf{0}. && (6)
\end{aligned}$$

The objective function in (6) is a concave function of $\mathbf{Q}_{S_k}$, and the constraint set is a convex set. Thus, Problem (6) is a convex optimization problem, which can be solved in polynomial time with standard interior-point methods [25].

The minimum time resource $t_{S_k,\min}$ that each S_k needs to satisfy its rate requirement is

$$t_{S_k,\min} = \frac{R_{S_k}}{r_{S_k,\max}}.$$

Problem (1) (or (4)) is feasible when the traffic load in the secondary system does not exceed its capacity, i.e.,

$$\sum_{k=1}^{K} t_{S_k,\min} \leq 1.$$

IV. Statistical CSI Scenario

In the statistical CSI scenario, the optimal solutions to the non-convex Problem (4) can be found in polynomial time by first optimizing the time fractions t_{S_k} and then the transmit covariance matrices $\mathbf{Q}_{S_k}$.

A. Optimal Solution for Continuous Time Allocation

Given any feasible time allocation $(t_{S_1}, \cdots, t_{S_K})$, Problem (4) reduces to K separate transmit covariance matrix optimization problems, one for each S_k:

$$\min_{\mathbf{Q}_{S_k}} \quad \mathrm{tr}\left(\mathbf{Q}_{S_k}\right)$$

$$\text{s.t.} \quad w \log\left[\det\left(\mathbf{I} + \mathbf{H}_{BS,S_k}\mathbf{Q}_{S_k}\mathbf{H}_{BS,S_k}^H\mathbf{C}_{S_k}^{-1}\right)\right] \geq \frac{R_{S_k}}{t_{S_k}}, \quad \text{(7a)}$$

$$\mathrm{tr}\left(\mathbf{Q}_{S_k}\right) \leq \rho_{S_k}, \quad \text{(7b)}$$

$$\mathbf{Q}_{S_k} \succeq \mathbf{0}.$$

The optimal solution to Problem (7) can be computed by standard "water-filling" [10], [23]. Let $\mathbf{A}_{S_k}$ denote $\mathbf{H}_{BS,S_k}^H\mathbf{C}_{S_k}^{-1}\mathbf{H}_{BS,S_k}$, which is an $M_{S_k} \times M_{S_k}$ Hermitian positive semidefinite matrix. Let $W_{S_k} = \mathrm{rank}(\mathbf{A}_{S_k})$ and $\lambda_{S_k,1} \geq \lambda_{S_k,2} \geq \cdots \geq \lambda_{S_k,W_{S_k}}$ denote all the non-negative eigenvalues of matrix $\mathbf{A}_{S_k}$. Let $\mathbf{v}_{S_k,i}$ ($\|\mathbf{v}_{S_k,i}\|_2^2 = 1$) denote the normalized eigenvector of $\mathbf{A}_{S_k}$ associated with eigenvalue $\lambda_{S_k,i}$, ($1 \leq i \leq W_{S_k}$), and $\mathbf{V}_{S_k}$ be a matrix whose ith column is $\mathbf{v}_{S_k,i}$. The optimal solution to Problem (7) is given in the following lemma.

Lemma 1: The necessary and sufficient condition for Problem (7) to be feasible is

$$t_{S_k} \geq t_{S_k,\min}. \quad (8)$$

Furthermore, let $\tilde{\mathbf{Q}}_{S_k}^*$ be a diagonal matrix with entries

$$\tilde{Q}_{S_k,ii}^* = \left(\mu_{S_k} - \frac{1}{\lambda_{S_k,i}}\right)^+, \quad (9)$$

where $(x)^+ = \max\{x, 0\}$, and the value of μ_{S_k} is chosen to satisfy

$$\prod_{i=1}^{W_{S_k}} \left(\lambda_{S_k,i}\mu_{S_k}\right)^+ = \exp\left(\frac{R_{S_k}}{w t_{S_k}}\right). \quad (10)$$

Then, when Condition (8) is satisfied, the optimal solution to Problem (7) is

$$\mathbf{Q}_{S_k}^* = \mathbf{V}_{S_k}\tilde{\mathbf{Q}}_{S_k}^*\mathbf{V}_{S_k}^H. \quad (11)$$

Proof: If we do not consider Constraint (7b) when solving Problem (7), the optimal solution is given by standard "water-filling". Constraint (7b) only states that the minimum objective value should be no greater than ρ_{S_k}. Since the MIMO link rate obtained from the water-filling is an increasing function in $\mathrm{tr}(\mathbf{Q}_{S_k})$ [23], we can find that Constraint (7b) is satisfied if t_{S_k} satisfies Condition (8). □

The optimal transmit covariance matrices are functions of the time allocation $(t_{S_1}, \cdots, t_{S_K})$. Thus, the optimal number of data streams and the optimal energy consumption of each S_k are dependent on its time resource allocation. We will now explore this dependence in more detail. To simplify the notation, let

$$\tau_{S_k}(m_{S_k}) = \frac{R_{S_k}}{w\left(\left(\sum_{i=1}^{m_{S_k}} \log \lambda_{S_k,i}\right) - m_{S_k} \log \lambda_{S_k,\left(m_{S_k}+1\right)}\right)},$$

where $m_{S_k} \in \{1, \cdots, (W_{S_k} - 1)\}$. Observe that τ_{S_k} is a decreasing function of m_{S_k}. The optimal number of data streams $D_{S_k}^*$ is a step-wise function of t_{S_k}, given by

$$D_{S_k}^*(t_{S_k}) = \begin{cases} W_{S_k}, & 0 < t_{S_k} < \tau_{S_k}\left(W_{S_k} - 1\right), \\ m_{S_k}, & \tau_{S_k}(m_{S_k}) \leq t_{S_k} < \tau_{S_k}(m_{S_k} - 1), \\ 1, & t_{S_k} \geq \tau_{S_k}(1). \end{cases} \quad (12)$$

Furthermore, the optimal "water level" $\mu_{S_k}^*$ is a function of t_{S_k}, given by

$$\mu_{S_k}^*\left(t_{S_k}\right) = \left(\frac{\exp\left(\frac{R_{S_k}}{w t_{S_k}}\right)}{\prod_{i=1}^{D_{S_k}^*\left(t_{S_k}\right)} \lambda_{S_k,i}}\right)^{\frac{1}{D_{S_k}^*\left(t_{S_k}\right)}}.$$

Therefore, the optimal energy consumption of S_k is a function of t_{S_k}, which can be computed according to

$$E_{S_k}(t_{S_k}) = t_{S_k}\left(D_{S_k}^*\left(t_{S_k}\right)\left(\frac{\exp\left(\frac{R_{S_k}}{w t_{S_k}}\right)}{\prod_{i=1}^{D_{S_k}^*\left(t_{S_k}\right)} \lambda_{S_k,i}}\right)^{\frac{1}{D_{S_k}^*\left(t_{S_k}\right)}} - \sum_{i=1}^{D_{S_k}^*\left(t_{S_k}\right)} \frac{1}{\lambda_{S_k,i}}\right), \quad (13)$$

Note that because the optimal number of data streams $D_{S_k}^*(t_{S_k})$ is a step-wise function of t_{S_k}, $E_{S_k}(t_{S_k})$ is a piece-wise defined function[2] of t_{S_k}. The following proposition establishes some

[2]Here the piece-wise defined function means that $E_{S_k}(t_{S_k})$ takes on different forms for different intervals of t_{S_k}. As shown in Proposition 1, $E_{S_k}(t_{S_k})$ is a continuous function in t_{S_k}.

key properties of the optimal energy consumtion $E_{S_k}(t_{S_k})$, which will be useful later when solving the time allocation problem among the SUs.

Proposition 1: The optimal energy consumption $E_{S_k}(t_{S_k})$ is a strictly convex, continuous, first-order differentiable, and monotonically decreasing function in t_{S_k}.

Proof: The proof is in Appendix A. □

Substituting (8), (11), and (13) into Problem (4), it then becomes an optimization problem in the time fraction variables t_{S_k} only. Furthermore, as shown in Proposition 1, the energy consumption (13) is a monotonically decreasing function in t_{S_k}. Thus, the optimal solution could be achieved only when Constraint (4b) is satisfied with equality. Therefore, the time resource optimization problem among the SUs is given by

$$\begin{aligned} \min_{t_{S_k}} \quad & \sum_{k=1}^{K} E_{S_k}(t_{S_k}) \\ \text{s.t.} \quad & \sum_{k=1}^{K} t_{S_k} = 1, \\ & t_{S_k} \geq t_{S_k,\min}, \qquad \forall k. \end{aligned} \tag{14}$$

The objective function $E_{S_k}(t_{S_k})$ is convex as shown in Proposition 1, and the constraints in (14) are linear. Therefore, Problem (14) is a convex optimization problem.

Now we are ready to solve Problem (4) through the decomposition method.

Theorem 1: In the statistical CSI scenario, the optimal time allocation and the optimal transmit covariance matrices can be found separately. In particular, the optimal time allocation is the optimal solution to the convex optimization problem (14). After obtaining the optimal $t^*_{S_k}$, the optimal transmit covariance matrix of each SU can then be computed by "water-filling", given in (11).

Proof: From Lemma 1, we know that given any feasible time allocation, the optimal transmit covariance matrix is computed by "water-filling". In Proposition 1, we further show that the energy consumption, based on the optimal covariance matrix, is a strictly convex function of t_{S_k}. Thus, the optimal solution to Problem (14) is the global optimal time allocation. Given this optimal $t^*_{S_k}$, the optimal transmit covariance matrix is then computed by "water-filling". □

Note that although the objective function in (14) is a piecewise defined function in t_{S_k}, we have shown that it is continuous, and its first-order derivative is also continuous. However, it is not second-order differentiable. Thus, only first-order methods for convex optimization problems (such as the gradient methods) can be applied to solve Problem (14) [25], [26]. Here we adopt the Spectral Projected Gradient (SPG) method [27], which is a variation of the projected gradient method. The SPG method makes two modifications. First, it incorporates the nonmonotone line search scheme proposed in [28]. Second, the stepsize is chosen to be the one introduced in [29]. With these two modifications, the number of iterations can be significantly reduced. Thus, the SPG method has been shown to converge to the global optimal solution with a competitive convergence rate for convex optimization problems [27], [30].

The time complexity for the SPG method to solve (14) is $O(K^2)$ in terms of iterations [27]. In each iteration, the gradient can be easily updated from an explicit function. The time complexity for each SU to obtain the optimal transmit covariance matrix by "water-filling" is $O(W^3_{S_k})$, where $W_{S_k} = \min\{M_{S_k}, N_{BS}\}$ [31]. Therefore, the overall time complexity to obtain the optimal solution in the statistical CSI scenario is $O(K^2) + O(KW^3_{S_k})$.

B. Optimal Solution for Discrete Time Allocation

In many wireless systems, the time frame is divided into a number of time slots. A slot is the smallest unit in the time allocation process. The time resource allocated to each S_k should be an integer indicating the number of time slots instead of a real number. With the additional integer constraint on the variable t_{S_k} in Problem (4), it then becomes a mixed-integer non-convex optimization problem, which is generally very difficult to solve. Fortunately, by exploring the special structure of Problem (4), optimal solutions can be obtained within polynomial time.

Without loss of generality, suppose each normalized frame has a total number of T time slots. the time resource optimization problem among the SUs is given by

$$\begin{aligned} \min_{t_{S_k}} \quad & \sum_{k=1}^{K} E_{S_k}\left(t_{S_k}\right) \\ \text{s.t.} \quad & \sum_{k=1}^{K} t_{S_k} = T, \\ & t_{S_k} \in \left\{ t^{(I)}_{S_k,\min}, \left(t^{(I)}_{S_k,\min}+1\right), \cdots, T \right\}, \qquad \forall k, \end{aligned} \tag{15}$$

where $t^{(I)}_{S_k,\min} = \lceil t_{S_k,\min} T \rceil$ is the minimum number of time slots that each S_k needs.

Problem (15) is an integer optimization problem, which is feasible if and only if

$$\sum_{k=1}^{K} t^{(I)}_{S_k,\min} \leq T.$$

The objective function of (15) is a separable sum of convex functions. With this special property, it is not difficult to find the optimal solution to (15) with a greedy algorithm [32]. Let $\Delta_k(t_{S_k})$ denote the change in the energy consumption when the number of time slots allocated to S_k is increased from $t_{S_k}-1$ to t_{S_k},

$$\Delta_k\left(t_{S_k}\right) = E_{S_k}(t_{S_k}) - E_{S_k}\left(t_{S_k}-1\right), \qquad t_{S_k} \in \left\{ \left(t^{(I)}_{S_k,\min}+1\right), \cdots, T \right\}.$$

The value of $\Delta_k(t^{(I)}_{S_k,\min})$ can be defined as $-\infty$.

As shown in Proposition 1, the energy consumption $E_{S_k}(t_{S_k})$ is strictly convex and monotonically decreasing in t_{S_k}. We then have

$$\Delta_k\left(t^{(I)}_{S_k,\min}\right) \leq \Delta_k\left(t^{(I)}_{S_k,\min}+1\right) \leq, \cdots, \leq \Delta_k(T), \qquad \forall k.$$

In the greedy algorithm, starting from the minimum slots allocation $\mathbf{t}_{\min}^{(I)} = \left(t_{S_1,\min}^{(I)}, \cdots, t_{S_K,\min}^{(I)}\right)$, one time slot is allocated at a time. A time slot is added to the S_k which has the minimum $\Delta_k(t_{S_k})$ among all the SUs. The algorithm stops when all the T time slots are allocated. The optimal solution to Problem (15) can be characterized by

$$\begin{cases} \Delta_k\left(t_{S_k}^*\right) \leq \varphi^*, & \forall k, \\ \Delta_k\left(t_{S_k}^*+1\right) \geq \varphi^*, & \forall k, \\ \sum_{k=1}^{K} t_{S_k}^* = T, & \\ t_{S_k}^* \in \left\{t_{S_k,\min}^{(I)}, \left(t_{S_k,\min}^{(I)}+1\right), \cdots, T\right\}, & \forall k. \end{cases}$$

In the allocation of one slot, the energy consumption can be easily computed from an explicit function. The time complexity for finding the minimum $\Delta_k(t_{S_k})$ among all the SUs is $O(K)$. Therefore, the overall time complexity to obtain the optimal discrete time allocation with the greedy algorithm is $O(KT)$.

V. Perfect CSI Scenario: Optimal Condition and Solution

The problem formulation in the perfect CSI scenario is shown in (1). We will show that it can also be solved to optimality through the decomposition method under the condition that the secondary system is *under-utilized* so that each SU has ample time resource for its transmission.

A. Decomposition Condition and Optimal Solution

Consider a relaxation of Problem (1) where Constraint (1c), which is the interference constraint to the primary system, has been removed:

$$\begin{aligned} \min_{t_{S_k},\mathbf{Q}_{S_k}} \quad & \sum_{k=1}^{K} t_{S_k} \mathrm{tr}\left(\mathbf{Q}_{S_k}\right) \\ \text{s.t.} \quad & t_{S_k} w \log\left[\det\left(\mathbf{I}+\mathbf{H}_{BS,S_k}\mathbf{Q}_{S_k}\mathbf{H}_{BS,S_k}^H\mathbf{C}_{S_k}^{-1}\right)\right] \\ & \quad \geq R_{S_k}, \qquad \forall k, \\ & \sum_{k=1}^{K} t_{S_k} \leq 1, \\ & \mathrm{tr}\left(\mathbf{Q}_{S_k}\right) \leq p_{S_k,\max}, \qquad \forall k, \\ & t_{S_k} \geq 0, \qquad \forall k, \\ & \mathbf{Q}_{S_k} \succeq \mathbf{0}, \qquad \forall k. \end{aligned} \tag{16}$$

Problem (16) is of the same form as the problem formulation in the statistical CSI scenario, with the only difference that ρ_{S_k} in Constraint (4c) has been replaced by $p_{S_k,\max}$. Therefore, Problem (16) can be solved to optimality with the same decomposition method as in Section IV. However, the relaxed Problem (16) is not equivalent to Problem (1). We next show a sufficient condition so that there is no gap between the optimal objective values of Problems (16) and (1). To simplify notation, let $\mathbf{B}_{P_j,S_k} = \mathbf{V}_{S_k}^H\mathbf{H}_{P_j,S_k}^H\mathbf{H}_{P_j,S_k}\mathbf{V}_{S_k}$, with b_{P_j,S_k}^{ii} denoting the entry in the ith row and ith column. Let $\eta_{S_k} = \min\left\{\frac{\phi_{P_1}+\sum_{i=1}^{D_{S_k}^*} b_{P_1,S_k}^{ii}/\lambda_{S_k,i}}{\sum_{i=1}^{D_{S_k}^*} b_{P_1,S_k}^{ii}}, \cdots, \frac{\phi_{P_J}+\sum_{i=1}^{D_{S_k}^*} b_{P_J,S_k}^{ii}/\lambda_{S_k,i}}{\sum_{i=1}^{D_{S_k}^*} b_{P_J,S_k}^{ii}}\right\}$, where $D_{S_k}^*$ is the optimal number of data streams for each S_k to Problem (16).

Proposition 2: In the perfect CSI scenario, if the optimal time allocation to Problem (16), $\mathbf{t}^* = (t_{S_1}^*, \cdots, t_{S_K}^*)$, satisfies

$$t_{S_k}^* \geq \frac{R_{S_k}}{w\left(\sum_{i=1}^{D_{S_k}^*} \log\lambda_{S_k,i} + D_{S_k}^*\log\eta_{S_k}\right)}, \qquad \forall k, \tag{17}$$

then the optimal solution to Problem (16) is also the optimal solution to Problem (1). In this case, the decomposition method can be applied to find the optimal solution to Problem (1).

Proof: The proof is in Appendix B. □

The physical meaning of the RHS of (17) is the time needed for each S_k to satisfy the interference constraint to the primary system. If for all the SUs, the energy-optimal time allocation satisfies (17), then we call the secondary system *under-utilized*. It is related to the values of K, R_{S_k}, and the channel conditions. When the secondary system is under-utilized, in the energy-optimal time allocation, the SUs will use up the entire frame for transmissions. Thus, the time resource for each SU is ample enough so that the instantaneous transmission rate is so low and thus does not violate the interference constraint to the primary system.

B. Heuristic Solution for Heavily-Utilized Secondary System

When the optimal time allocation to Problem (16) does not satisfy Condition (17), i.e., the secondary system is heavily-utilized but does not exceed its capacity, the decomposition method cannot be applied. In this case, we propose a heuristic algorithm using alternating optimization, to provide near-optimal solution to Problem (1). In particular, we first fix the time allocation, and optimize the covariance matrix for each SU. Given the solutions of the covariance matrices of all the SUs, we then optimize the time allocation among the SUs. The algorithm iterates between these two steps until convergence.

We choose an initial feasible time allocation to be the minimum time resource required in the perfect CSI scenario, i.e., $\mathbf{t_0} = (t_{S_1,\min}, \cdots, t_{S_K,\min})$. In the first step, given the fixed time fraction allocation $\mathbf{t} = (t_{S_1}, \cdots, t_{S_K})$, Problem (1) becomes K separate transmit covariance matrix optimization problems, one for each S_k:

$$\begin{aligned} \min_{\mathbf{Q}_{S_k}} \quad & p_{S_k} = \mathrm{tr}\left(\mathbf{Q}_{S_k}\right) \\ \text{s.t.} \quad & w\log\left[\det\left(\mathbf{I}+\mathbf{H}_{BS,S_k}\mathbf{Q}_{S_k}\mathbf{H}_{BS,S_k}^H\mathbf{C}_{S_k}^{-1}\right)\right] \geq \frac{R_{S_k}}{t_{S_k}}, && \text{(18a)} \\ & \mathrm{tr}\left(\mathbf{H}_{P_j,S_k}\mathbf{Q}_{S_k}\mathbf{H}_{P_j,S_k}^H\right) \leq \phi_{P_j}, \qquad \forall j, && \text{(18b)} \\ & \mathrm{tr}\left(\mathbf{Q}_{S_k}\right) \leq p_{S_k,\max}, && \text{(18c)} \\ & \mathbf{Q}_{S_k} \succeq \mathbf{0}. \end{aligned}$$

Problem (18) is a convex optimization problem. Thus, the optimal covariance matrix can be found in polynomial time using interior point methods [25], [33]. With a similar method as in Problem (1), it can be shown that the optimal covariance matrix to Problem (18) satisfies $\mathrm{rank}(\mathbf{Q}_{S_k}^*) \leq \mathrm{rank}(\mathbf{H}_{BS,S_k}^H \mathbf{C}_{S_k}^{-1} \mathbf{H}_{BS,S_k}) \leq \mathrm{rank}(\mathbf{H}_{BS,S_k})$.

In the second step, we fix the covariance matrices of all the SUs to be the values obtained in step one. Then we optimize the time allocation with the SPG method. The SPG method requires the gradient of the energy consumption at each given $\mathbf{t}$, which can be obtained based on the local sensitivity analysis [25]. Let $p_{S_k}^*$ and $\xi_{S_k}^*$ denote the optimal objective value of Problem (18) and the corresponding Lagrange multiplier associated with Constraint (18a), respectively. According to the local sensitivity analysis, $\xi_{S_k}^*$ is related to the gradient of $p_{S_k}^*$ at given t_{S_k}:

$$\frac{\partial p_{S_k}^*}{\partial \left(R_{S_k}/t_{S_k}\right)} = \xi_{S_k}^*.$$

We then have

$$\frac{\partial p_{S_k}^*}{\partial t_{S_k}} = -\xi_{S_k}^* \frac{R_{S_k}}{t_{S_k}^2}.$$

Thus, the gradient of the energy consumption E_{S_k} at given t_{S_k} is

$$\frac{\partial E_{S_k}}{\partial t_{S_k}} = \frac{\partial \left(p_{S_k}^* t_{S_k}\right)}{\partial t_{S_k}} = p_{S_k}^* + t_{S_k} \frac{\partial p_{S_k}^*}{\partial t_{S_k}} = p_{S_k}^* - \xi_{S_k}^* \frac{R_{S_k}}{t_{S_k}}.$$

In the time allocation problem, the objective function is $E(\mathbf{t}) = \sum_{k=1}^{K} E_{S_k}(t_{S_k})$. The feasible set for $\mathbf{t} = (t_{S_1}, \cdots, t_{S_K})$ is $\mathcal{T} = \{\mathbf{t} \mid \sum_{k=1}^{K} t_{S_k} \leq 1, t_{S_k} \geq t_{S_k,\min}, k = 1, \cdots, K\}$. The new time allocation $\mathbf{t}^{(n+1)}$ is updated according to

$$\mathbf{t}^{(n+1)} = \mathcal{P}_{\mathcal{T}}\left(\mathbf{t}^{(n)} - \alpha^{(n)} \nabla E\left(\mathbf{t}^{(n)}\right)\right),$$

where $\mathcal{P}_{\mathcal{T}}[\cdot]$ denotes projection on the feasible set $\mathcal{T}$. The step-size $\alpha^{(n)}$ is chosen to be the one introduced in [29]. The SPG method has been shown to experimentally speed up the convergence compared with the traditional projected gradient method [27].

VI. Simulation Results

We evaluate the performance of the proposed algorithms in terms of energy consumption, interference to primary system, convergence, and optimality. In the simulation setup, there are two primary links with 10m distance between the transmitter and corresponding receiver. The SUs are uniformly distributed in a square area of 200 m × 200 m, with a minimum separation of 35 m from the primary receivers. An example CR network with 35 SUs is shown in Fig. 1. We assume that each node in the network has 4 antennae. For each network setting, we perform 1000 independent simulation runs.

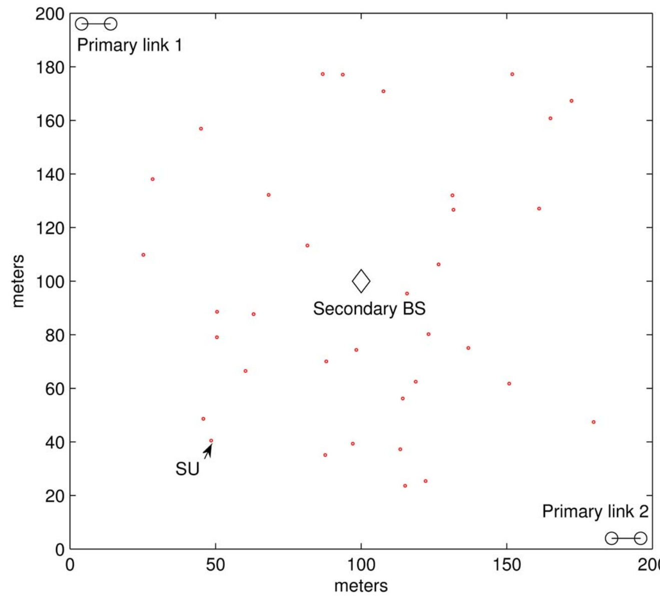


Fig. 1. A sample random CR network with 2 primary links and 35 secondary users.

The frame length of the secondary system is 20 ms, and each SU has a rate requirement of 32 kbps. The carrier frequency is 1 GHz and the network bandwidth is 20 MHz. We adopt the i.i.d. Rayleigh fading channel model with path loss exponent equal to 4. The PUs transmit at the maximum power of 20 dBm, while the maximum transmit power for each SU is 27.5 dBm. The noise power density is −174 dBm/Hz. The interference power threshold ϕ_{P_j} is set such that $\frac{\phi_{P_j}}{N_0 w}$ is 25 dB. The outage probability δ_{P_j} in the statistical CSI scenario is 1%.

A. Energy Consumption

The energy consumption per bit of the SUs in the statistical CSI scenario, named “optimal time & multiple streams”, is shown in Fig. 2. As discussed in Section IV, the optimal time allocation and transmit covariance matrices can be found in polynomial time. For comparison, we investigate the energy consumption of the “max-rate” and “optimal time & single stream” transmission schemes. In the “max-rate” scheme, each SU uses the minimum time resource and the transmit covariance matrix of the solution to Problem (5). While in the “optimal time & single stream” scheme proposed in [22], each SU optimizes its time allocation and the transmit beamforming, but with the restriction of transmitting one data stream on all its antennae.

Compared with the “max-rate” scheme, the energy consumption of both the “optimal time & single stream” and “optimal time & multiple streams” is reduced significantly. The energy consumption per bit roughly remains constant in the “max-rate” scheme. In the “optimal time & single stream” and “optimal time & multiple streams” schemes, we could see a tradeoff between the energy consumption and the system traffic load (i.e., the number of SUs). Taking the “max-rate” scheme as

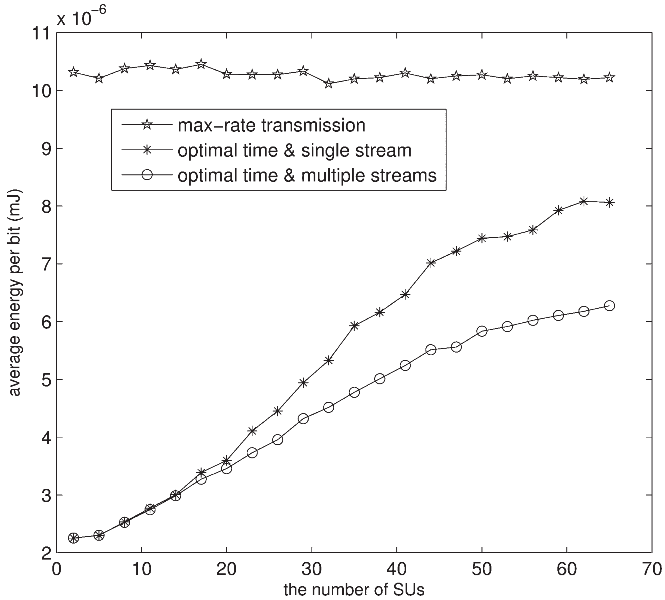


Fig. 2. Average energy consumption per bit of the secondary system in the statistical CSI scenario.

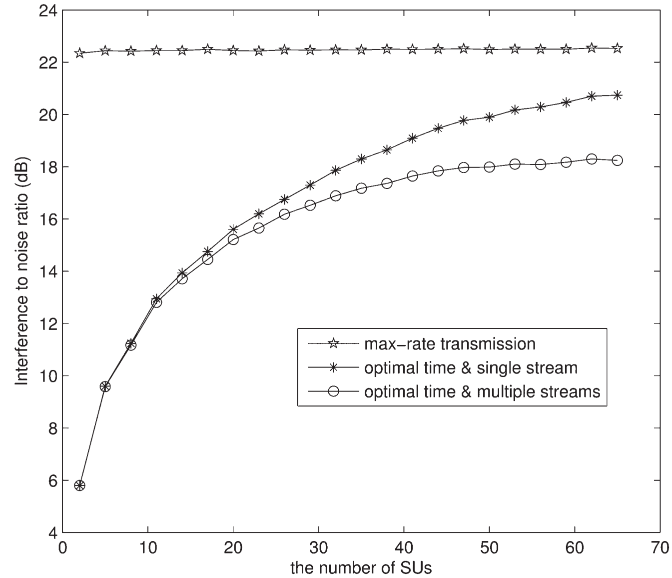


Fig. 3. Average interference power at primary receiver vs. the number of SUs in the statistical CSI scenario.

the baseline, the energy reduction of "optimal time & multiple streams" is up to 78% when the secondary system is under-utilized. During simulations, we have observed that when the secondary system is under-utilized or when it experiences a strong interference from the primary system, the optimal way for each SU to perform transmit precoding is to do the single data stream transmission. This can be seen in Fig. 2: when the secondary system has less than 8 SUs, both the "optimal time & single stream" and "optimal time & multiple streams" produce the same result. As the traffic load increases, the optimal number of data streams will increase to exploit the multiplexing gain and further reduce the energy consumption. Compared with the "optimal time & single stream" scheme, the energy consumption of the "optimal time & multiple streams" scheme is reduced by 23% when the secondary system has 65 users.

B. Interference to the Primary Receivers

Fig. 3 shows the average interference power from the SUs to the primary receivers in the statistical CSI scenario, which is measured by the interference-to-noise ratio (in dB). In the "max-rate" scheme, the average interference-to-noise ratio is smaller than the required 25 dB. This is because the secondary transmitter needs to satisfy the interference constraint at both primary receivers in each simulated network. Therefore, the interference-to-noise ratio at each primary receiver is smaller than 25 dB in many simulation runs. The interference to the primary system is also alleviated in the "optimal time & single stream" and "optimal time & multiple streams" schemes. The interference power to the primary system adapts to the traffic load of the secondary system. Moreover, we can observe that in the "optimal time & multiple streams" scheme, the interference power increases slower compared with the "optimal time & single stream" scheme.

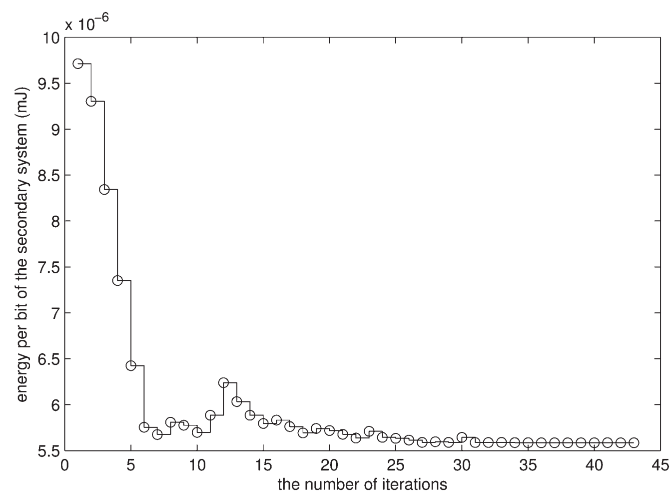


Fig. 4. The energy consumption of the secondary network as a function of the number of iterations in the statistical CSI scenario. The secondary network has 35 users with the topology shown in Fig. 1.

C. Convergence

As shown in Section IV-A, the SPG method is guaranteed to converge to the optimal time allocation in the statistical CSI scenario. Fig. 4 shows the convergence performance for the sample random secondary network with 35 users, as shown in Fig. 1. We find that the objective function is reduced significantly during the first 10 iterations. Another observation is that the objective function in the SPG method may not monotonically decrease in every iteration, which has also been discussed in [30]. Fig. 5 shows the convergence performance, plotted in terms of the average number of iterations. The stopping criterion is $\|\mathbf{t}^{(n+1)} - \mathbf{t}^{(n)}\|_\infty < 10^{-5}$. We observe that the number of iterations first increases and then decreases with the number of SUs, and always stays below 85 iterations. This

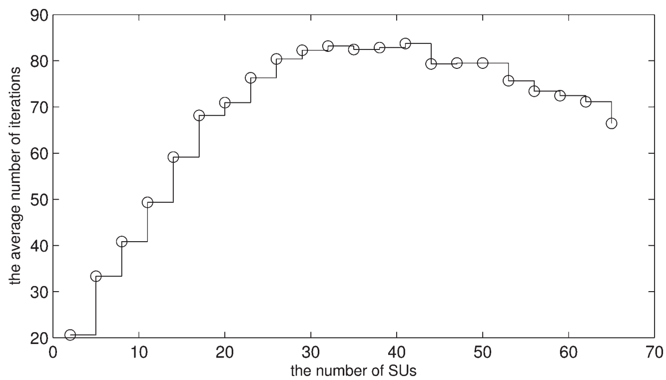


Fig. 5. Average number of iterations needed to converge to the optimal solution in the statistical CSI scenario.

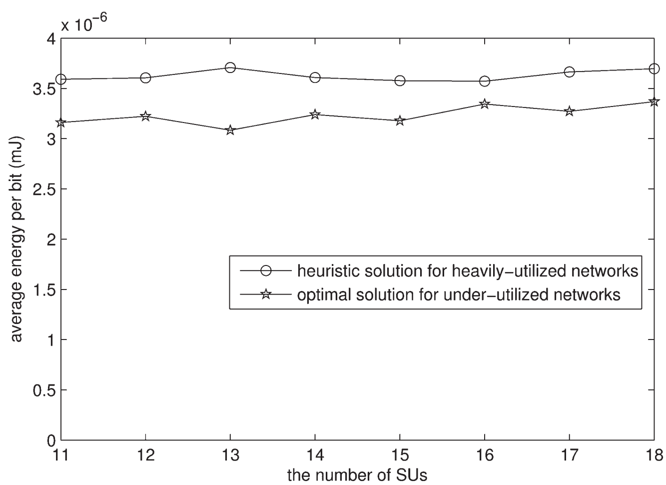


Fig. 6. Performance gap of the heuristic algorithm in the perfect CSI scenario with heavy network load.

indicates that the SPG method has a good convergence rate even in a secondary system with many users. Note that in the perfect CSI scenario, the SPG method is also guaranteed to converge to the optimal solution with similar convergence performance when the secondary system is under-utilized.

D. Optimality

Only in the case of perfect CSI scenario with heavy network load, we need to resort to the heuristic algorithm proposed in Section V-B. Fig. 6 shows the performance of the heuristic algorithm. We investigate the secondary networks which are at the boundary between under-utilized and heavily-utilized (i.e., K ranges from 11 to 18). Given each K, among the 1000 tested networks, some are under-utilized and some are heavily-utilized, which is due to the randomness of the SUs' locations. The under-utilized secondary networks have better average channel gains on the secondary links compared to those in the heavily-utilized networks. The average minimum energy consumption for these under-utilized secondary networks could serve as an appropriate baseline in order to show the performance obtained with the heuristic algorithm. As can be seen from Fig. 6, the heuristic algorithm performs very close to the optimal values in the under-utilized networks. The gap is within 20%. However, it is difficult to prove formal bounds on sub-optimality.

VII. Conclusion

In this paper, we have considered jointly energy-optimal time allocation and precoding in MIMO CR networks. The problem formulations turn out to be non-convex optimization problems. We successfully tackle the non-convexity by applying an optimization decomposition technique. As a result, under statistical CSI, the global optimal solution can be found efficiently; under perfect CSI, the global optimal solution can be obtained efficiently when the secondary system is under-utilized. We know that since the number of data streams in the transmit covariance matrix is discrete by nature, the energy consumption is a piece-wise defined function in the time variable. We have shown that in the statistical CSI scenario, the optimal energy consumption is continuous, first-order differentiable, strictly convex and monotonically decreasing in the time variable, and how this allows to use the SPG algorithm to quickly find the optimal time allocation. In a slotted system, where time is a discrete variable, the optimal number of time slots can be found with a greedy algorithm. Given the optimal time allocation, the optimal transmit covariance matrix is obtained by water-filling. Only in the case of perfect CSI and when the secondary system is heavily-utilized, we need resort to a heuristic algorithm that iterates between two optimization modules.

Note that spectrum efficiency can be further improved if the SUs send simultaneous uplink transmissions with spatial multiplexing. However, in this case, the SUs will interfere with each other. Finding the optimal transmit covariance matrices for the SUs is more challenging. Considering the trade-off between the diversity gain and the spatial multiplexing gain for energy-efficient MIMO CR networks is left for future studies.

Appendix A
Proof of Proposition 1

The expessions of $\lim_{t_{S_k}\to\tau_{S_k}^{-}(m_{S_k})} E_{S_k}(t_{S_k})$ and $\lim_{t_{S_k}\to\tau_{S_k}^{+}(m_{S_k})} E_{S_k}(t_{S_k})$ are shown in (19) and (20), shown at the bottom of the next page. Therefore, $E_{S_k}(t_{S_k})$ is a continuous function in t_{S_k}.

The first order derivative of $E_{S_k}(t_{S_k})$ with respect to t_{S_k} is

$$\frac{dE_{S_k}}{dt_{S_k}} = \left(D_{S_k}^{*}\left(t_{S_k}\right) - \frac{R_{S_k}}{wt_{S_k}}\right)\left(\frac{\exp\left(\frac{R_{S_k}}{wt_{S_k}}\right)}{\prod_{i=1}^{D_{S_k}^{*}(t_{S_k})}\lambda_{S_k,i}}\right)^{\frac{1}{D_{S_k}^{*}\left(t_{S_k}\right)}} - \sum_{i=1}^{D_{S_k}^{*}\left(t_{S_k}\right)}\frac{1}{\lambda_{S_k,i}}.$$

The expressions of $\lim_{t_{S_k}\to\tau_{S_k}^-(m_{S_k})}\frac{dE_{S_k}}{dt_{S_k}}$ and $\lim_{t_{S_k}\to\tau_{S_k}^+(m_{S_k})}\frac{dE_{S_k}}{dt_{S_k}}$ are shown in (21) and (22), shown at the bottom of the next page, respectively. Thus, we have

$$\lim_{t_{S_k}\to\tau_{S_k}^-\left(m_{S_k}\right)}\frac{dE_{S_k}}{dt_{S_k}}=\lim_{t_{S_k}\to\tau_{S_k}^+\left(m_{S_k}\right)}\frac{dE_{S_k}}{dt_{S_k}}.$$

Therefore, $\frac{dE_{S_k}}{dt_{S_k}}$ is a continuous function in t_{S_k}.

The second order derivative of $E_{S_k}(t_{S_k})$ with respect to t_{S_k} is

$$\frac{d^2E_{S_k}}{dt_{S_k}^2}=\frac{R_{S_k}^2}{w^2t_{S_k}^3D_{S_k}^*\left(t_{S_k}\right)}\left(\frac{\exp\left(\frac{R_{S_k}}{wt_{S_k}}\right)}{\prod_{i=1}^{D_{S_k}^*\left(t_{S_k}\right)}\lambda_{S_k,i}}\right)^{\frac{1}{D_{S_k}^*\left(t_{S_k}\right)}},$$

which is always positive for any positive t_{S_k}. However, it is not continuous in t_{S_k}, with the non-continuous points at $t_{S_k}=\tau_{S_k}(m_{S_k})$, $(m_{S_k}=\{1,\cdots,(W_{S_k}-1)\})$.

$$\begin{aligned}
\lim_{t_{S_k}\to\tau_{S_k}^-\left(m_{S_k}\right)}E_{S_k}\left(t_{S_k}\right)&=\tau_{S_k}\left(m_{S_k}\right)\left(\left(m_{S_k}+1\right)\left(\frac{\exp\left(\frac{R_{S_k}}{w\tau_{S_k}\left(m_{S_k}\right)}\right)}{\prod_{i=1}^{m_{S_k}+1}\lambda_{S_k,i}}\right)^{\frac{1}{m_{S_k}+1}}-\sum_{i=1}^{m_{S_k}+1}\frac{1}{\lambda_{S_k,i}}\right),\\
&=\tau_{S_k}\left(m_{S_k}\right)\left(\left(m_{S_k}+1\right)\left(\frac{\exp\left(\left(\sum_{i=1}^{m_{S_k}}\log\lambda_{S_k,i}\right)-m_{S_k}\log\lambda_{S_k,\left(m_{S_k}+1\right)}\right)}{\prod_{i=1}^{m_{S_k}+1}\lambda_{S_k,i}}\right)^{\frac{1}{m_{S_k}+1}}-\sum_{i=1}^{m_{S_k}+1}\frac{1}{\lambda_{S_k,i}}\right),\\
&=\tau_{S_k}\left(m_{S_k}\right)\left(\left(m_{S_k}+1\right)\left(\frac{\prod_{i=1}^{m_{S_k}}\lambda_{S_k,i}}{\lambda_{S_k,\left(m_{S_k}+1\right)}^{m_{S_k}}\prod_{i=1}^{m_{S_k}+1}\lambda_{S_k,i}}\right)^{\frac{1}{m_{S_k}+1}}-\sum_{i=1}^{m_{S_k}+1}\frac{1}{\lambda_{S_k,i}}\right)\\
&=\tau_{S_k}\left(m_{S_k}\right)\left(\frac{m_{S_k}}{\lambda_{S_k,\left(m_{S_k}+1\right)}}-\sum_{i=1}^{m_{S_k}}\frac{1}{\lambda_{S_k,i}}\right)
\end{aligned}\tag{19}$$

$$\begin{aligned}
\lim_{t_{S_k}\to\tau_{S_k}^+\left(m_{S_k}\right)}E_{S_k}\left(t_{S_k}\right)&=\tau_{S_k}\left(m_{S_k}\right)\left(m_{S_k}\left(\frac{\exp\left(\frac{R_{S_k}}{w\tau_{S_k}\left(m_{S_k}\right)}\right)}{\prod_{i=1}^{m_{S_k}}\lambda_{S_k,i}}\right)^{\frac{1}{m_{S_k}}}-\sum_{i=1}^{m_{S_k}}\frac{1}{\lambda_{S_k,i}}\right),\\
&=\tau_{S_k}\left(m_{S_k}\right)\left(m_{S_k}\left(\frac{\exp\left(\left(\sum_{i=1}^{m_{S_k}}\log\lambda_{S_k,i}\right)-m_{S_k}\log\lambda_{S_k,\left(m_{S_k}+1\right)}\right)}{\prod_{i=1}^{m_{S_k}}\lambda_{S_k,i}}\right)^{\frac{1}{m_{S_k}}}-\sum_{i=1}^{m_{S_k}}\frac{1}{\lambda_{S_k,i}}\right),\\
&=\tau_{S_k}\left(m_{S_k}\right)\left(\frac{m_{S_k}}{\lambda_{S_k,\left(m_{S_k}+1\right)}}-\sum_{i=1}^{m_{S_k}}\frac{1}{\lambda_{S_k,i}}\right)
\end{aligned}\tag{20}$$

Next we show that $\frac{dE_{S_k}}{dt_{S_k}}$ is always negative for any positive t_{S_k}. Since $\frac{d^2E_{S_k}}{dt^2_{S_k}}$ is always positive, this means that $\frac{dE_{S_k}}{dt_{S_k}}$ is an increasing function in t_{S_k}. Thus, we know that $\frac{dE_{S_k}}{dt_{S_k}}$ satisfies inequality (23), shown at the bottom of the page. The second line of (23) follows from (12), which states that when t_{S_k} approaches infinity, the optimal number of data streams equals 1. Therefore, we can find that the optimal energy consumption $E_{S_k}(t_{S_k})$ is a strictly convex, continuous, first-order differentiable, and monotonically decreasing function in t_{S_k}.

$$\begin{aligned}
\lim_{t_{S_k}\to\tau^-_{S_k}(m_{S_k})} \frac{dE_{S_k}}{dt_{S_k}} &= \left(\left(m_{S_k}+1\right) - \frac{R_{S_k}}{w\tau_{S_k}\left(m_{S_k}\right)}\right)\left(\frac{\exp\left(\frac{R_{S_k}}{w\tau_{S_k}\left(m_{S_k}\right)}\right)}{\prod_{i=1}^{m_{S_k}+1}\lambda_{S_k,i}}\right)^{\frac{1}{m_{S_k}+1}} - \sum_{i=1}^{m_{S_k}+1}\frac{1}{\lambda_{S_k,i}},\\
&= \left(\left(m_{S_k}+1\right) - \frac{R_{S_k}}{w\tau_{S_k}\left(m_{S_k}\right)}\right)\left(\frac{\exp\left(\left(\sum_{i=1}^{m_{S_k}}\log\lambda_{S_k,i}\right) - m_{S_k}\log\lambda_{S_k,\left(m_{S_k}+1\right)}\right)}{\prod_{i=1}^{m_{S_k}+1}\lambda_{S_k,i}}\right)^{\frac{1}{m_{S_k}+1}} - \sum_{i=1}^{m_{S_k}+1}\frac{1}{\lambda_{S_k,i}},\\
&= \left(\left(m_{S_k}+1\right) - \frac{R_{S_k}}{w\tau_{S_k}\left(m_{S_k}\right)}\right)\frac{1}{\lambda_{S_k,\left(m_{S_k}+1\right)}} - \sum_{i=1}^{m_{S_k}+1}\frac{1}{\lambda_{S_k,i}},\\
&= \left(m_{S_k} - \frac{R_{S_k}}{w\tau_{S_k}\left(m_{S_k}\right)}\right)\frac{1}{\lambda_{S_k,\left(m_{S_k}+1\right)}} - \sum_{i=1}^{m_{S_k}}\frac{1}{\lambda_{S_k,i}}
\end{aligned} \tag{21}$$

$$\begin{aligned}
\lim_{t_{S_k}\to\tau^+_{S_k}(m_{S_k})} \frac{dE_{S_k}}{dt_{S_k}} &= \left(m_{S_k} - \frac{R_{S_k}}{w\tau_{S_k}\left(m_{S_k}\right)}\right)\left(\frac{\exp\left(\frac{R_{S_k}}{w\tau_{S_k}\left(m_{S_k}\right)}\right)}{\prod_{i=1}^{m_{S_k}}\lambda_{S_k,i}}\right)^{\frac{1}{m_{S_k}}} - \sum_{i=1}^{m_{S_k}}\frac{1}{\lambda_{S_k,i}},\\
&= \left(m_{S_k} - \frac{R_{S_k}}{w\tau_{S_k}\left(m_{S_k}\right)}\right)\left(\frac{\exp\left(\left(\sum_{i=1}^{m_{S_k}}\log\lambda_{S_k,i}\right) - m_{S_k}\log\lambda_{S_k,\left(m_{S_k}+1\right)}\right)}{\prod_{i=1}^{m_{S_k}}\lambda_{S_k,i}}\right)^{\frac{1}{m_{S_k}}} - \sum_{i=1}^{m_{S_k}}\frac{1}{\lambda_{S_k,i}},\\
&= \left(m_{S_k} - \frac{R_{S_k}}{w\tau_{S_k}\left(m_{S_k}\right)}\right)\frac{1}{\lambda_{S_k,\left(m_{S_k}+1\right)}} - \sum_{i=1}^{m_{S_k}}\frac{1}{\lambda_{S_k,i}}
\end{aligned} \tag{22}$$

$$\begin{aligned}
\frac{dE_{S_k}}{dt_{S_k}} &< \lim_{t_{S_k}\to\infty}\left(\left(D^*_{S_k}\left(t_{S_k}\right) - \frac{R_{S_k}}{wt_{S_k}}\right)\left(\frac{\exp\left(\frac{R_{S_k}}{wt_{S_k}}\right)}{\prod_{i=1}^{D^*_{S_k}\left(t_{S_k}\right)}\lambda_{S_k,i}}\right)^{\frac{1}{D^*_{S_k}\left(t_{S_k}\right)}} - \sum_{i=1}^{D^*_{S_k}\left(t_{S_k}\right)}\frac{1}{\lambda_{S_k,i}}\right)\\
&= \left.\left(D^*_{S_k}\left(t_{S_k}\right)\left(\frac{1}{\prod_{i=1}^{D^*_{S_k}\left(t_{S_k}\right)}\lambda_{S_k,i}}\right)^{\frac{1}{D^*_{S_k}\left(t_{S_k}\right)}} - \sum_{i=1}^{D^*_{S_k}\left(t_{S_k}\right)}\frac{1}{\lambda_{S_k,i}}\right)\right|_{D^*_{S_k}\left(t_{S_k}\right)=1} = \frac{1}{\lambda_{S_k,1}} - \frac{1}{\lambda_{S_k,1}} = 0
\end{aligned} \tag{23}$$

APPENDIX B
PROOF OF THEOREM 2

If the optimal solution to Problem (16) satisfies Constraint (1c), the optimal solutions of Problems (16) and (1) are equivalent. For each S_k, substituting the optimal transmit covariance matrix solution in Lemma 1 into Constraint (1c), we have

$$\operatorname{tr}\left(\mathbf{H}_{P_j,S_k}\mathbf{V}_{S_k}\tilde{\mathbf{Q}}^*_{S_k}\mathbf{V}^H_{S_k}\mathbf{H}^H_{P_j,S_k}\right) = \operatorname{tr}\left(\tilde{\mathbf{Q}}^*_{S_k}\mathbf{B}_{P_j,S_k}\right) \le \phi_{P_j} \qquad \forall j. \tag{24}$$

According to (9) and (10), we have

$$\begin{aligned}
&\operatorname{tr}\left(\tilde{\mathbf{Q}}^*_{S_k}\mathbf{B}_{P_j,S_k}\right)\\
&= \sum_{i=1}^{D^*_{S_k}}\left(\left(\frac{\exp\left(\frac{R_{S_k}}{wt_{S_k}}\right)}{\prod_{i=1}^{D^*_{S_k}}\lambda_{S_k,i}}\right)^{\frac{1}{D^*_{S_k}}} - \frac{1}{\lambda_{S_k,i}}\right) b^{ii}_{P_j,S_k}\\
&= \left(\frac{\exp\left(\frac{R_{S_k}}{wt_{S_k}}\right)}{\prod_{i=1}^{D^*_{S_k}}\lambda_{S_k,i}}\right)^{\frac{1}{D^*_{S_k}}}\left(\sum_{i=1}^{D^*_{S_k}} b^{ii}_{P_j,S_k}\right) - \sum_{i=1}^{D^*_{S_k}}\frac{b^{ii}_{P_j,S_k}}{\lambda_{S_k,i}}.
\end{aligned}$$

Therefore, inequality (24) is equivalent to

$$\begin{aligned}
t_{S_k} &\ge \frac{R_{S_k}}{w\left(\sum_{i=1}^{D^*_{S_k}}\log\lambda_{S_k,i} + D^*_{S_k}\log\frac{\phi_{P_j}+\sum_{i=1}^{D^*_{S_k}} b^{ii}_{P_j,S_k}/\lambda_{S_k,i}}{\sum_{i=1}^{D^*_{S_k}} b^{ii}_{P_j,S_k}}\right)}\\
\Rightarrow t_{S_k} &\ge \frac{R_{S_k}}{w\left(\sum_{i=1}^{D^*_{S_k}}\log\lambda_{S_k,i} + D^*_{S_k}\log\eta_{S_k}\right)}, \qquad \forall j.
\end{aligned}$$

REFERENCES

[1] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-efficiency of MIMO and cooperative MIMO techniques in sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 6, pp. 1089–1098, Aug. 2004.

[2] M. Kobayashi and G. Caire, "Joint beamforming and scheduling for a multi-antenna downlink with imperfect transmitter channel knowledge," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 7, pp. 1468–1477, Sep. 2007.

[3] A. B. Gershman, N. D. Sidiropoulos, S. Shahbazpanahi, M. Bengtsson, and B. Ottersten, "Convex optimization-based beamforming: From receive to transmit and network designs," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 62–75, May 2010.

[4] W. Yu and T. Lan, "Transmitter optimization for the multi-antenna downlink with per-antenna power constraints," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2646–2660, Jun. 2007.

[5] H. Kim, C.-B. Chae, G. de Veciana, and R. W. Health, Jr., "A cross-layer approach to energy efficiency for adaptive MIMO systems exploiting spare capacity," *IEEE Trans. Wireless Commun.*, vol. 8, no. 8, pp. 4264–4275, Aug. 2009.

[6] L. Fu, H. Kim, J. Huang, S. C. Liew, and M. Chiang, "Energy conservation and interference mitigation: From decoupling property to win-win strategy," *IEEE Trans. Wireless Commun.*, vol. 10, no. 11, pp. 3943–3955, Nov. 2011.

[7] K. T. Phan, S. A. Vorobyov, N. D. Sidiropoulos, and C. Tellambura, "Spectrum sharing in wireless networks via QoS-aware secondary multicast beamforming," *IEEE Trans. Signal Process.*, vol. 57, no. 6, pp. 2323–2335, Jun. 2009.

[8] Y. Huang, Q. Li, W.-K. Ma, and S. Zhang, "Robust multicast beamforming for spectrum sharing-based cognitive radios," *IEEE Trans. Signal Process.*, vol. 60, no. 1, pp. 527–533, Jan. 2012.

[9] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201–220, Feb. 2005.

[10] D. Gesbert, M. Shafi, D. Shiu, P. J. Smith, and A. Naguib, "From theory to practice: An overview of MIMO space-time coded wireless systems," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 3, pp. 281–302, Apr. 2003.

[11] Y.-C. Liang, K.-C. Chen, G. Y. Li, and P. Mähönen, "Cognitive radio networking and communications: An overview," *IEEE Trans. Veh. Technol.*, vol. 60, no. 7, pp. 3386–3407, Sep. 2011.

[12] G. Zheng, S. Song, K.-K. Wong, and B. Ottersten, "Cooperative cognitive networks: Optimal, distributed and low-complexity algorithms," *IEEE Trans. Signal Process.*, vol. 61, no. 11, pp. 2778–2790, Jun. 2013.

[13] N. D. Sidiropoulos, T. N. Davidson, and Z.-Q. Luo, "Transmit beamforming for physical-layer multicasting," *IEEE Trans. Signal Process.*, vol. 54, no. 6, pp. 2239–2251, Jun. 2006.

[14] R. Zhang and Y.-C. Liang, "*Exploiting Multi-Antennas for Opportunistic Spectrum Sharing in Cognitive Radio Networks*," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 88–102, Feb. 2008.

[15] G. Zheng, K.-K. Wong, and B. Ottersten, "Robust cognitive beamforming with bounded channel uncertainties," *IEEE Trans. Signal Process.*, vol. 57, no. 12, pp. 4871–4881, Dec. 2009.

[16] R. Zhang, Y.-C. Liang, and S. Cui, "Dynamic resource allocation in cognitive radio networks: A convex optimization perspective," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 102–114, May 2010.

[17] L. Zhang, Y.-C. Liang, and Y. Xin, "Joint beamforming and power allocation for multiple access channels in cognitive radio networks," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 1, pp. 38–51, Jan. 2008.

[18] L. Zhang, Y. Xin, and Y.-C. Liang, "Weighted sum rate optimization for cognitive radio MIMO broadcast channels," *IEEE Trans. Wireless Commun.*, vol. 8, no. 6, pp. 2950–2959, Jun. 2009.

[19] Y. J. Zhang and A. M.-C. So, "Optimal spectrum sharing in MIMO cognitive radio networks via semidefinite programming," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 2, pp. 362–373, Feb. 2011.

[20] M. Razaviyayn, M. Sanjabi, and Z.-Q. Luo, "Linear transceiver design for interference alignment: Complexity and computation," *IEEE Trans. Inf. Theory*, vol. 58, no. 5, pp. 2896–2910, May 2012.

[21] G. Zhang, S. Ma, K.-K. Wong, and T.-S. Ng, "Robust beamforming in cognitive radio," *IEEE Trans. Wireless Commun.*, vol. 9, no. 2, pp. 570–576, Feb. 2010.

[22] L. Fu, Y. J. Zhang, and J. Huang, "Energy efficient transmissions in MIMO cognitive radio networks," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 11, pp. 2420–2431, Nov. 2013.

[23] E. Telatar, "Capacity of multi-antenna Gaussian channels," AT&T Bell Lab., Murray Hill, NJ, USA, Tech. Memo, Jun. 1995.

[24] T. S. Rappaport, *Wireless Communications: Principles and Practice*, 2nd ed. Upper Saddle River, NJ, USA, Prentice-Hall, 2002.

[25] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K., Cambridge Univ. Press, 2004.

[26] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed. Belmont, MA, USA, Athena Scientific, 1999.

[27] E. Birgin, J. Martínez, and M. Raydan, "Nonmonotone spectral projected gradient methods on convex sets," *SIAM J. Optim.*, vol. 10, no. 4, pp. 1196–1211, 2000.

[28] L. Grippo, F. Lampariello, and S. Lucidi, "A nonmonotone line search technique for Newtons method," *SIAM J. Numer. Anal.*, vol. 23, no. 4, pp. 707–716, Aug. 1986.

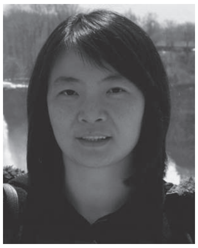
[29] J. Barzilai and J. Borwein, "Two-point step size gradient methods," *IMA J. Numer. Anal.*, vol. 8, no. 1, pp. 141–148, 1988.

[30] M. Schmidt, "Graphical model structure learning with l_1-regularization," Ph. D. dissertation, Univ. British Columbia, Vancouver, BC, Canada, 2010.

[31] D. P. Palomar and J. R. Fonollosa, "Practical algorithms for a family of waterfilling solutions," *IEEE Trans. Signal Process.*, vol. 53, no. 2, pp. 686–695, Feb. 2005.

[32] T. Ibaraki and N. Katoh, *Resource Allocation Problems: Algorithmic Approaches*. Cambridge, MA, USA, MIT Press, 1988.

[33] L. Vandenberghe, S. Boyd, and S.-P. Wu, "Determinant maximization with linear matrix inequality constraints," *SIAM J. Matrix Anal. Appl.*, vol. 19, no. 2, pp. 499–533, 1998.



Liqun Fu (S'08–M'11) received the Ph.D. degree in information engineering from The Chinese University of Hong Kong in 2010. She is currently a Postdoctoral Researcher with the ACCESS Linnaeus Centre, KTH Royal Institute of Technology. During 2009, she was a Visiting Scholar in the Department of Electrical Engineering, Princeton University. From 2010-2013, she was a Post-doctoral Research Fellow at the Institute of Network Coding, The Chinese University of Hong Kong. Her research interests are in the area of wireless communication and networking, with current focus on wireless greening, resource allocation, distributed protocol design, and signal processing. She is the awardee of the CUHK Global Scholarship for Research Excellence in 2009. She has served as a TPC Co-Chair of INFOCOM workshop on GCCCN 2014, and TPC member for ICC, Globecom, PIMRC, and WCNC.

Mikael Johansson received the M.Sc. and Ph.D. degrees in electrical engineering from Lund University, Sweden, in 1994 and 1999, respectively. He held postdoctoral positions at Stanford University and U.C. Berkeley before joining KTH in 2002, where he now serves as Full Professor. He has published two books and over hundred papers, several of which are highly cited and have received recognition in terms of best paper awards. He has served on the editorial boards of *Automatica* and the IEEE TRANSACTIONS ON CONTROL OF NETWORKED SYSTEMS, as well as on the program committee for several top-conferences organized by IEEE and ACM. He has played a leading role in several national and international research projects in control and communications.

Mats Bengtsson (M'00–SM'06) received the M.S. degree in computer science from Linköping University, Linköping, Sweden, in 1991 and the Tech. Lic. and Ph.D. degrees in electrical engineering from the KTH Royal Institute of Technology, Stockholm, Sweden, in 1997 and 2000, respectively.

From 1991 to 1995, he was with Ericsson Telecom AB Karlstad. He is currently an Associate Professor with the Signal Processing Department, School of Electrical Engineering, KTH. His research interests include statistical signal processing and its applications to communications, multi-antenna processing, cooperative communication, radio resource management, and propagation channel modelling. He served as Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING 2007–2009 and was a member of the IEEE SPCOM Technical Committee 2007–2012.