Optimal Scheduling and Power Allocation in Cooperate-to-Join Cognitive Radio Networks

Mehmet Karaca, *Student Member, IEEE*, Karim Khalil, *Student Member, IEEE*, Eylem Ekici, *Senior Member, IEEE*, and Ozgur Ercetin, *Member, IEEE*

***Abstract*—In this paper, optimal resource allocation policies are characterized for wireless cognitive networks under the spectrum leasing model. We propose cooperative schemes in which secondary users share the time-slot with primary users in return for cooperation. Cooperation is feasible only if the primary system's performance is improved over the non-cooperative case. First, we investigate a scheduling problem where secondary users are interested in immediate rewards. Here, we consider both infinite and finite backlog cases. Then, we formulate another problem where the secondary users are guaranteed a portion of the primary utility, on a long-term basis, in return for cooperation. Finally, we present a power allocation problem where the goal is to maximize the expected net benefit defined as utility minus cost of energy. Our proposed scheduling policies are shown to outperform non-cooperative scheduling policies, in terms of expected utility and net benefit, for a given set of feasible constraints. Based on Lyapunov optimization techniques, we show that our schemes are arbitrarily close to the optimal performance at the price of reduced convergence rate.**

***Index Terms*—Cognitive radios, Lyapunov optimization, opportunistic scheduling, resource allocation, spectrum leasing.**

I. Introduction

COGNITIVE radio networks (CRNs) have recently been investigated extensively [2], [3]. The main advantage that CRN presents is the efficient utilization of the scarce radio spectrum resources. By opportunistically exploiting the under utilized spectrum, unlicensed (i.e., secondary) users can transmit over the licensed bands, provided that they do not *hurt* the performance of the licensed (i.e., primary) users.

Approaches to cognitive radio can be divided into two categories: *commons* model and *property-rights* model [4], [5]. In the commons model, the primary network is oblivious to the secondary network activity, and the aim of secondary users (SUs) is to detect and exploit the spectrum holes without interacting with the primary system. These spectrum holes represent the absence of primary activity either in time, frequency, or space. In the property-rights model (spectrum leasing), primary users (PUs) own the spectrum and are willing to lease it to SUs in return for some form of service, for instance, cooperation via relaying. Consider the following motivating scenario: In a cellular network, a licensed wireless user is far away from the base station and is experiencing low transmission rates. At the same time, a cognitive user is halfway between the licensed user and the base station and thus has better channel conditions. The cognitive user desires to access the channel to send some of its own data to the base station. After coordination, PU agrees to share a portion of its own time-slot with SU in exchange for SU relaying PU's data to the base station. In our work, we exploit this cooperative scheme between primary and secondary systems to improve the overall performance.

Scheduling is an essential problem for any shared resource. The problem becomes more challenging in a dynamic setting such as wireless networks where the channel capacity is time-varying due to multiple superimposed random effects such as mobility and multipath fading. Optimal scheduling in wireless networks has been extensively studied in the literature under various assumptions and purposes. In [6]–[8], the authors proposed scheduling algorithms in which both delay and channel conditions were taken into account. It has been shown that policies that exploit the time-varying nature of the wireless channel are at least as good as static policies [9]. In principle, these opportunistic policies schedule the user with the favorable channel conditions to increase the overall performance of the system. However, without imposing individual performance guarantees for each user in the system, opportunistic policies may lead to starvation of some users, for example, those far away from the base station in a cellular network. To mitigate this problem, fairness constraints are added to the problem formulation as in [9]–[11]. Minimum energy consumption is yet another important performance objective especially for mobile users with limited power supply [12], [13]. For example, in [12], the authors developed power optimal opportunistic scheduling policy assuming users have minimum rate constraints. The authors in [13] considered a utility-based power control framework for a cellular system. However, these works assumed a sing-hop system, and no cooperation among users was investigated.

Opportunistic scheduling was recently studied for cognitive radio networks under the commons model [14], [15]. In these works, Lyapunov optimization tools were used to design flow control, scheduling, and resource allocation algorithms and explicit performance bounds were derived. Using the technique

Manuscript received March 02, 2011; revised February 29, 2012 and June 27, 2012; accepted November 16, 2012; approved by IEEE/ACM TRANSACTIONS ON NETWORKING Editor S. Sarkar. This work was supported in part by the National Science Foundation under Grant No. CCF-0914912 and the European Commission under Marie Curie IRSES Grant PIRSES-GA-2010-269132 AGILENet. The material in this paper was presented in part at the IEEE International Conference on Computer Communications (INFOCOM), Shanghai, China, April 10–15, 2011.

M. Karaca and O. Ercetin are with the Faculty of Engineering and Natural Sciences, Sabanci University, 34956 Istanbul, Turkey (e-mail: mehmetkrc@sabanciuniv.edu; oercetin@sabanciuniv.edu).

K. Khalil and E. Ekici are with Department of Electrical and Computer Engineering, The Ohio State University, Columbus, OH 43210 USA (e-mail: khalilk@ece.osu.edu; ekici@ece.osu.edu).

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/TNET.2012.2230187

of virtual queues, the joint problem of stabilizing the queues of SUs in addition to satisfying long-term constraint on the collision probability or interference on the primary channels is transformed into a queue stability problem. In addition, cognitive radio system applying spectrum leasing approach was investigated in [16]–[18]. In [16], the authors presented a spectrum leasing scheme that allows PU to lease its own bandwidth for a fraction of time in exchange for enhanced performance guarantee via cooperation with SUs. As a result, more spectrum access opportunity is left for SU to transmit their own data. The authors in [17] and [18] developed a game-theoretic framework for a spectrum leasing in which PU actively participates in a non-cooperative game with SUs. In these works, PU plays an active role and allows SUs' access while meeting its own minimum quality-of-service (QoS) requirement. On the other hand, SUs aim to achieve energy-efficient transmissions as long as they do not cause excessive interference to PU.

In this paper, we propose optimal opportunistic scheduling policies for primary *and* secondary users in a cognitive radio network under the spectrum leasing model. To the best of the authors knowledge, this is the first work to consider scheduling of cooperative primary and secondary networks with multiple users sharing a common destination. For example, [16]–[18] considered only one primary transmitter and separate receivers for primary and secondary systems. Thus, the only coordination required is among the transmission between the single PU and a subset of SUs. In addition, the authors in [17] and [18] did not explicitly model the price paid by SUs to PUs to share the licensed spectrum. In our paper, we first consider the optimization of the total expected utility of both primary and secondary systems while satisfying an average performance constraint for each primary user in the network. Here, we develop a cooperative scheduling policy by which the performance is improved and shown to be *at least* as good as the original primary-only system. For a given time-slot, users cooperate using decode-and-forward multihop scheme [19] where SUs relay the messages of PUs to a common destination in a portion of the time-slot as a levy of using the already licensed spectrum for a fraction of that time-slot. The parameters specifying the cooperation strategy are the fraction of the time-slot during which SU relays PU's data and the fraction used to transmit SU's own data.

Next, another formulation is considered in which SUs are guaranteed some portion of the primary utility in an average sense in return for cooperation. This formulation presents a model of *banking* between primary and secondary systems where rewards are gained over the long term. Finally, we formulate a power control problem where the objective is to maximize the net benefit defined as the difference between the value of the utility and the cost of the energy consumption, under minimum requirement constraints on PUs. We employ Lyapunov optimization tools developed in [20] and [21] to analyze our proposed schemes and to derive explicit bounds on the performance achieved. We show that our proposed schemes can be pushed arbitrarily close to the optimal with a tradeoff between optimality and the convergence rate of the algorithms.

The rest of this paper is organized as follows. Section II presents the network model, the basic structure of the proposed cooperative schemes, and an introduction to the Lyapunov optimization technique. In Section III, we formulate the immediate rewards scheduling problem and derive both stationary and time-varying optimal policies. Then, in Section IV, we formulate and solve another version of the problem where constraints on the minimum performance of SUs and long-term rewards are considered. In Section V, we formulate and analyze the net benefit problem where we consider the cost for energy consumption. Numerical results are presented in Section VI. Finally, Section VII concludes the paper and presents possible future directions.

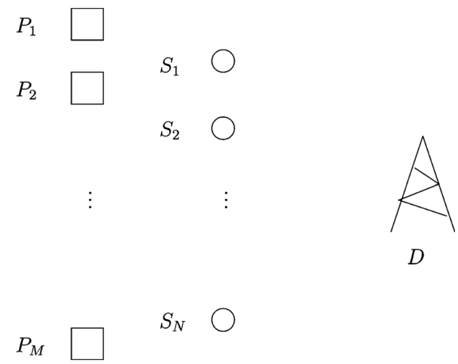

Fig. 1. Network model.

II. Network Model

A. Cognitive Network

Consider a cognitive radio network of M PUs and N SUs, all wishing to communicate with a common destination D as shown in Fig. 1. This destination can be viewed as a base station in a single cell of a cellular network or as an access point in a Wi-Fi network.

We consider a time-slotted system where the time-slot is the resource to be shared among different users. We adopt a non-interference model where only one user, either primary or secondary, is transmitting at any given time. Random channel gains between each user and other users in the network are assumed to be independent and identically distributed (i.i.d.) across time according to a general distribution and independent across users with values taken from a finite set. Moreover, we assume that channel gains are time-varying, but fixed over the time-slot duration. We assume the availability of perfect channel-state information of all channels at the scheduler, i.e., knowledge of channel coefficients immediately prior to transmission.

In the following analysis, we use the notation $R_m^{\rm p}(t)$, $R_n^{\rm s}(t)$ to denote the transmission rates from PU m to destination and from SU n to destination, respectively, at time-slot t. The corresponding random rate vectors are denoted as $\mathbf{R}^{\rm p}(t)$, $\mathbf{R}^{\rm s}(t)$. The transmission rate from PU m to SU n is denoted as $R_{mn}^{\rm r}(t)$, where the corresponding rate matrix is $\mathbf{R}^{\rm r}(t)$. The transmission rate is a function of the random channel conditions, and thus a measure of the channel quality. We assume that transmission rate processes are ergodic and bounded. As will be clear in Section II-B, since our scheme works by selecting a pair of users (primary and secondary) to transmit at a given time-slot, the utility achieved by a user is a function of the cooperating

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

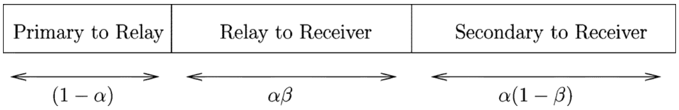

Fig. 2. General time-slot structure.

pair. Consequently, the utility function of a PU m when it cooperates with SU n at time-slot t is denoted as $U_{mn}(t)$. Similarly, the utility function of an SU n that cooperates with PU m is denoted as $V_{mn}(t)$. These utility functions are measures of the level of satisfaction of users, and thus they are generally assumed to be nondecreasing concave functions of the transmission rate.

B. Cooperative Scheme

To schedule transmissions of different users, a scheduling policy is required. In our cooperative framework, we allow the scheduling policy to either schedule a PU to transmit during a given time-slot, or to schedule a pair of primary and secondary users to share the time-slot, according to the channel conditions. The scheduling policy Q is a rule that selects the four-tuple (m, n, α, β) to transmit at time-slot t, where α and β specify the cooperation strategy the pair of primary and secondary users m, n use. In a time-slot t, the scheduling policy is a function of the rate vectors $\mathbf{R}^{\mathrm{p}}(t)$, $\mathbf{R}^{\mathrm{s}}(t)$, rate matrix $\mathbf{R}^{\mathrm{r}}(t)$, and possibly other variables related to past performance. Note that the scheduling policy we adopt is opportunistic in the sense that it exploits the time-varying nature of the wireless channel.

In our model, we focus on a cooperation based spectrum leasing scenario. Under this model, scheduling is done such that, if feasible, a pair of primary and secondary users cooperatively share a single time-slot to improve the performance of the original primary system and allow unlicensed users to access the licensed spectrum, where feasibility is to be defined. Cooperation is achieved as follows: For a fraction $(1-\alpha)$, $0 \le \alpha \le 1$, of the time-slot, PU m sends its data (intended to destination) to SU n (relay). In the remaining portion of the time-slot, the scheduled SU uses the channel to relay PU's data over a β fraction, $0 \le \beta \le 1$, and then transmits its own data during the rest of the time-slot, i.e., over $\alpha(1-\beta)$ fraction. A schematic of the time-slot structure is shown in Fig. 2. This cooperative scheme is a form of implementation of the spectrum-leasing cognitive radio framework where SUs help primary system improve its performance to access the licensed spectrum. By this scheme, our system is in fact trying to reap the benefits of a form of spatial diversity. We note that the structure of our scheme is similar to the cooperative scheme of [16], however we do not employ distributed space time coding and allow only one SU to cooperate in a given time-slot.

We set $n = 0$ by definition for the case when a PU m is scheduled to transmit directly to the destination without cooperating with SUs. This is the case when cooperation is either infeasible or leads to suboptimal utility values. We set $R^{\mathrm{r}}_{m0}(t) = R^{\mathrm{p}}_{m}(t)$, $m \in \{1, 2, \ldots, M\}$ and $\alpha = 0$ in such cases.

The utility function is taken to be a nonnegative nondecreasing concave function of the rate. This choice is of practical interest since a small increase in the rate in the low-rate regime is generally more appreciated than a small increase in the high-rate regime. Given a scheduling decision $Q = (m', n', \alpha, \beta)$, we define the utility of the selected primary and secondary users, $U_{m'n'}(Q,t)$ and $V_{m'n'}(Q,t)$, respectively, as

$$U_{m'n'}(Q,t) = \begin{cases} h_1(R^{\mathrm{p}}_{m'}(t)), & \text{if } n' = 0 \\ h_1((1-\alpha)R^{\mathrm{r}}_{m'n'}(t)), & \text{otherwise} \end{cases} \quad (1)$$

$$V_{m'n'}(Q,t) = \begin{cases} 0, & \text{if } n' = 0 \\ h_2(\alpha(1-\beta)R^{\mathrm{s}}_{n'}), & \text{otherwise} \end{cases} \quad (2)$$

in a given time-slot t. For all other primary and secondary users (m, n) such that $m \ne m'$ and $n \ne n'$, we set $U_{mn}(Q,t) = V_{mn}(Q,t) = 0$. In the following, we sometimes use the shorthand $U_{mn}(t)$ and $V_{mn}(t)$ in place of $U_{mn}(Q,t)$ and $V_{mn}(Q,t)$ for simplicity. Examples of utility functions that can be used include $h_i(x) = \log(1+x)$ and $h_i(x) = x$, $i \in \{1, 2\}$.

Note that we assume scheduler's knowledge of the transmission rates for the primary and secondary users at each time-slot. For the scheduler to choose a pair to transmit over a given slot rather than scheduling a PU for direct transmission, feasibility conditions should hold. For a time-slot t, the feasibility conditions can be summarized as follows:

$$0 < (1-\alpha)R^{\mathrm{r}}_{mn}(t) \le \alpha R^{\mathrm{s}}_{n}(t). \quad (3)$$

The strict inequality in (3) guarantees validity of the cooperation, whereas the second inequality asserts that SU n has a sufficiently good channel to relay primary transmission at a given time-slot t. Given α, it can be seen that the optimal value of β is given by

$$\beta^* = \frac{(1-\alpha)R^{\mathrm{r}}_{mn}}{\alpha R^{\mathrm{s}}_{n}}. \quad (4)$$

If $\beta < \frac{(1-\alpha)R^{\mathrm{r}}_{mn}}{\alpha R^{\mathrm{s}}_{n}}$, SU n does not have sufficient time to relay the data of m. If $\beta > \frac{(1-\alpha)R^{\mathrm{r}}_{mn}}{\alpha R^{\mathrm{s}}_{n}}$, then unnecessary time is wasted by SU n. Thus, in the following, we use the notation $Q = (m, n, \alpha)$ for the decision of a scheduling policy Q. Note that (3) implies $\beta \le 1$.

Since we are interested in the maximization of the total expected utility of both primary and secondary systems, (4) is required to ensure superiority over non-cooperative schemes as will be clear in Section III.

Let $\mathcal{F}$ be the set of feasible policies at a given time-slot. The set $\mathcal{F}$ is constructed from all the tuples (m, n, α) such that (3) holds for some $0 \le \alpha \le 1$. We set the tuple $(m, 0, 0) \in \mathcal{F}$ by definition.

Let the total utility of the system (both primary and secondary), when scheduling policy Q is employed at a given time-slot t, be $W(Q,t)$. Then

$$W(Q,t) = \sum_{m=1}^{M}\sum_{n=0}^{N} U_{mn}(Q,t) + V_{mn}(Q,t). \quad (5)$$

Note that when the scheduling policy Q selects the tuple (m', n', α), the system receives a reward of $W(Q,t) = U_{m'n'}(Q,t) + V_{m'n'}(Q,t)$. The total expected utility is defined as $\bar{W}(Q,t) \triangleq \mathbb{E}[W(Q,t)]$ where the expectation is taken over the random transmission rates (random channel conditions), and possibly over the randomized policy.

C. Lyapunov Drift With Optimization

In our work, we use Lyapunov drift and optimization tools to show the optimality of our schemes. The advantage of this tool is the ability to provide a simple way to find optimal scheduling algorithms for complex models and to prove their optimality. Basically, this simplicity comes from defining each constraint as a virtual queue and then transforming the problem into a network stability problem [20].

We first introduce two definitions: Let $Z_i(t)$, $i \in \{1,2,\ldots,L\}$ be a queue backlog process and $\mathbf{Z}(t) = (Z_1(t)Z_2(t)\cdots Z_L(t))$ in a network with L users. Suppose that the goal is to stabilize the backlog process $\mathbf{Z}(t)$ while maximizing the time average of a scalar-valued utility function $g(\cdot)$ of another process $\mathbf{R}(t)$. Suppose that the optimal value of $g(\cdot)$ is g^*. Define the following quadratic Lyapunov function and conditional Lyapunov drift

$$L(\mathbf{Z}(t)) \triangleq \frac{1}{2}\sum_{l=1}^{L} Z_l^2(t) \tag{6}$$

$$\Delta(\mathbf{Z}(t)) \triangleq \mathbb{E}\left[L(\mathbf{Z}(t+1)) - L(\mathbf{Z}(t))|\mathbf{Z}(t)\right]. \tag{7}$$

We restate a result of [21] that is critical to establish the optimality of our proposed schemes.

Theorem 1 (Lyapunov Optimization) [21]: For the scalar valued function $g(\cdot)$, if there exists positive constants K, ϵ, B, such that for all time-slots t and all unfinished work vectors $\mathbf{Z}(t)$ the Lyapunov drift satisfies the condition

$$\Delta(\mathbf{Z}(t)) - K\mathbb{E}[g(\mathbf{R}(t))|\mathbf{Z}(t)] \leq B - \epsilon\sum_{l=0}^{L} Z_l(t) - Kg^*$$

then the time-average utility and queue backlog satisfy

$$\liminf_{t\to\infty} \frac{1}{t}\sum_{\tau=0}^{t-1} \mathbb{E}[g(\mathbf{R}(\tau))] \geq g^* - \frac{B}{K} \tag{8}$$

$$\limsup_{t\to\infty} \frac{1}{t}\sum_{\tau=0}^{t-1}\sum_{l=1}^{L} \mathbb{E}[Z_l(\tau)] \leq \frac{B + K(\bar{g} - g^*)}{\epsilon} \tag{9}$$

where $\bar{g} = \limsup_{t\to\infty} \frac{1}{t}\sum_{\tau=0}^{t-1} \mathbb{E}[g(\mathbf{R}(\tau))]$.

We note that Theorem 1 is a modified version of [21, Theorem 5.4]. Specifically, in our analysis, the function $g(\cdot)$ represents the total utility of the system in a time-slot given by (5), which is function of the utility matrices (1) and (2) and the scheduling policy Q. Hence, our objective is to maximize the time average of the instantaneous utilities. On the other hand, in [21], the utility of each user is a function of the time average of the instantaneous rates. However, by using Jensen's inequality and noting that the utility function is concave, one can easily show that the same lower bounds in (8) and (9) apply for our objective. We note that by using T-slot Lyaponuv drift techniques [21], similar results can be derived for more general (i.e., correlated) channel processes.

In our problem, since the utility of individual primary and secondary users is bounded, it can be shown that the total utility $W(Q,t)$ is bounded. It follows that the total expected utility can be pushed arbitrarily close to the optimum by choosing K sufficiently large. However, this leads to increasing bound on the average queue size given in (9).

III. Primary Constraints and Immediate Rewards

In this section, the goal is to schedule the transmissions of primary and secondary users to achieve maximum average sum utility of primary and secondary systems while maintaining minimum performance levels for each PU. Here, SU n is allowed to access the spectrum only if cooperation improves the instantaneous utility of a PU m Hence, we define $\mathcal{F}_1$ as the set of tuples (m,n,α) satisfying the following condition:

$$R_m^{\mathrm{p}}(t) \leq (1-\alpha)R_{mn}^{\mathrm{r}}(t) \leq \alpha R_n^{\mathrm{s}}(t) \tag{10}$$

for some $0 < \alpha < 1$. This constraint sets an upper bound on the range of α for the possible cooperation between each pair (m,n). Note that $\mathcal{F}_1 \subset \mathcal{F}$. We discuss two types of scheduling policies. First, we consider stationary scheduling polices that depend only on the values of the rates $\mathbf{R}^{\mathrm{p}}(t)$, $\mathbf{R}^{\mathrm{s}}(t)$, $\mathbf{R}^{\mathrm{r}}(t)$. Then, we investigate the more general dynamic policies.

A. Problem Formulation

The optimal opportunistic scheduling problem with minimum performance constraints was previously solved in [9]. By including N SUs to the system, our model can be viewed as a generalization to the model in [9]. In addition, setting $N = 0$ in our scheme yields the scheme in [9] as will be shown in Section III-B.

The problem[1] can be stated formally as follows:

$$\max_{Q\in\mathcal{F}_1} \quad \bar{W}(Q,t)$$
$$\text{s.t.} \quad \mathbb{E}\left[\sum_{n=0}^{N} U_{mn}(Q,t)\right] \geq C_m \tag{11}$$

$m \in \{1,2,\ldots,M\}$, where C_m is the minimum performance constraint for each PU m. To compare to the non-cooperative system, an example of the choice of the constraints C_m is given at the end of Section III-B.

The aforementioned problem formulation along with (10) implies that SUs are rewarded access to the channel immediately during a time-slot if their cooperation improves the performance of the primary system.

B. Optimal Stationary Policy

In this section, we propose a stationary scheduling policy in a form similar to the optimal policies reported in [9] and show that it solves (11) for the given cognitive radio network.

Scheduling Algorithm $Q_{1a:}$*:* For every time-slot t and given the values of $U_{mn}(t)$ and $V_{mn}(t)$ for all pairs (m,n), the solution to (11) is given by

$$Q_{1\mathrm{a}} = \underset{(m,n,\alpha)\in\mathcal{F}_1}{\arg\max} \; \{\lambda_m^* U_{mn}(Q,t) + V_{mn}(Q,t)\} \tag{12}$$

[1]Another valid objective is the utility of expected rate, which we do not focus on in this paper. Our algorithms are not necessarily optimal with respect to this objective.

where λ_m^*, $m \in \{1, 2, \ldots, M\}$ are real-valued parameters satisfying:

1) $\min_m \lambda_m^* = 1$;
2) $\mathbb{E}\left[\sum_{n=0}^N U_{mn}(Q_{1\mathrm{a}}, t)\right] \geq C_m$ for all m;
3) if $\mathbb{E}\left[\sum_{n=0}^N U_{mn}(Q_{1\mathrm{a}}, t)\right] > C_m$, then $\lambda_m^* = 1$.

The following theorem shows the optimality of Algorithm $Q_{1\mathrm{a}}$ where the proof is provided in Appendix A.

Theorem 2: Scheduling Algorithm $Q_{1\mathrm{a}}$ solves (11).

The structure of the derived scheduling policy suggests that when PU m experiences unfavorable channel conditions, the associated parameter λ_m^* will be larger than unity. Then, it attains average utility that is only equal to its corresponding constraint. Otherwise, this PU is granted a utility strictly larger than its minimum requirement.

The policy in (12) is stationary since it only depends on the values of the utility functions. Note that for any time-slot t, given the values of $R_m^{\mathrm{p}}(t)$, $R_n^{\mathrm{s}}(t)$ and $R_{mn}^{\mathrm{r}}(t)$ for all m and n, the scheduler is able to construct the set of feasible policies $\mathcal{F}_1$ by associating ranges $r_\alpha^- \leq \alpha \leq r_\alpha^+$ and $r_\beta^- \leq \beta \leq r_\beta^+$ for each pair (m, n). These ranges are chosen to satisfy the feasibility conditions (10), where $0 \leq r_\alpha^-, r_\alpha^+, r_\beta^-, r_\beta^+ \leq 1$. Then, the scheduler decides which pair is relatively best according to (12). The choice of the pair (m, n) is a combinatorial optimization problem that may require discrete exhaustive search. The optimal value of α can be obtained since (12) can be shown to be concave in α by a simple application of the second derivative test [22]. In addition, when $h_1(.)$ and $h_2(.)$ are given, one can use the first derivative test to determine the optimal value of α analytically.

The parameters λ_m^*, $m \in \{1, 2, \ldots, M\}$ depend on the choice of $h_1(.)$, $h_2(.)$ and the distribution of the utility functions, which in turn depends on the distribution of the underlying channel variations. Hence, λ_m^* needs to be estimated online in practice. This can be carried out using stochastic approximation techniques similar to the one explained in [9]. An estimation technique is presented in Section VI.

Example: The above algorithm can be compared to non-cooperative algorithms as follows. Consider for example the utilitarian fairness constraints problem solved in [9] with the constraints $a_m = \gamma_m \bar{W}(\hat{Q})$ for each PU $m \in \{1, 2, \ldots, M\}$ where $\sum_{m=1}^M \gamma_m \leq 1$ and $\bar{W}(\hat{Q})$ is the average performance achieved under the optimal (primary-only) scheduling policy $\hat{Q}$. According to this definition of a_m, the problem is always feasible. Let Γ be the improvement factor with respect to the system with no cooperation. Note that SUs can have access to the channel if they help improving the performance of primary system. Thus, PUs achieve utility that is at least the same as that in the non-cooperative case (i.e., $\Gamma \geq 1$). On the other hand, Γ is upper-bounded such that $\Gamma \leq \Gamma_{\max}$ due to the capacity region constraint (see Section III-E). Hence, $1 \leq \Gamma \leq \Gamma_{\max}$.

Now consider a network of M primary and N secondary users such that the scheduler executes the optimal policy $\hat{Q}$ to schedule only the PUs but does not act on it and simultaneously executes and implements our cooperative scheduling policy $Q_{1\mathrm{a}}$. Since the scheduling policy $\hat{Q}$ converges as the number of time-slots $t \to \infty$ [9], we can set $C_m = \Gamma a_m$ in (11). As long as $1 \leq \Gamma \leq \Gamma_{\max}$, it follows that our cooperative scheme improves the performance of individual PUs over the non-cooperative scheme, and hence improves the overall performance.

C. Optimal Dynamic Policy

In this section, we solve (11) using the stochastic network optimization tool of [21]. This tool yields a scheduling policy that is similar in structure to (12). However, the policy derived in this subsection does not entail the computation of the online parameters λ_m^*.

Define the time-average expected utility as follows:

$$\bar{W}(Q) = \lim_{t \to \infty} \frac{1}{t} \sum_{\tau=0}^{t-1} \mathbb{E}[W(Q, \tau)] \tag{13}$$

where $W(Q, t)$ is defined in (5).

Let $U_m^t(Q, t) \triangleq \sum_{n=0}^N U_{mn}(Q, t)$, $V_n^t(Q, t) \triangleq \sum_{m=1}^M V_{mn}(Q, t)$. For each of the constraints in (11), we construct a virtual queue such that the queue dynamics is given by

$$X_m(t+1) = [X_m(t) - U_m^t(Q, t)]^+ + C_m \tag{14}$$

$m \in \{1, 2, \ldots, M\}$, where $[x]^+ \triangleq \max\{x, 0\}$. Note that stabilizing the queues in (14) is equivalent to satisfying the constraints in (11) since a queue is stable if the arrival rate is less than the service rate. We assume that $U_m^t(Q, t)$ and $V_n^t(Q, t)$ are bounded such that $U_m^t(Q, t) \leq U^{\max}$, $V_n^t(Q, t) \leq V^{\max}$ for all $m \in \{1, 2, \ldots, M\}$, $n \in \{1, 2, \ldots, N\}$, $t \geq 0$ and for all $Q \in \mathcal{F}_1$. These upper bounds are justified since we assume bounded transmission rates. Let $\mathbf{X}(t) = (X_1(t) X_2(t) \cdots X_M(t))$ be the vector of virtual queues. Define the following quadratic Lyapunov function and conditional Lyapunov drift:

$$L_1(\mathbf{X}(t)) \triangleq \frac{1}{2} \sum_{m=1}^M X_m^2(t) \tag{15}$$

$$\Delta_1(\mathbf{X}(t)) \triangleq \mathbb{E}\left[L_1(\mathbf{X}(t+1)) - L_1(\mathbf{X}(t)) | \mathbf{X}(t)\right]. \tag{16}$$

Define the following conditional expectation:

$$\bar{U}_m^t(Q, t) \triangleq \mathbb{E}[U_m^t(Q, t) | \mathbf{X}(t)]. \tag{17}$$

The following Lemma is useful in establishing the optimality of our algorithm.

Lemma 1: For every time-slot t and any policy Q, the Lyapunov drift in (16) can be upper-bounded as follows:

$$\begin{aligned} \Delta_1(\mathbf{X}(t)) - K\mathbb{E}[W(Q, t) | \mathbf{X}(t)] \leq B_1 + \sum_{m=1}^M X_m(t) C_m \\ - \sum_{m=1}^M (K + X_m(t)) \bar{U}_m^t(Q, t) - \sum_{n=1}^N K \bar{V}_n^t(Q, t) \end{aligned} \tag{18}$$

where $B_1 = \frac{1}{2}\left(\sum_{m=1}^M C_m^2 + M(U^{\max})^2\right)$ and K is a system parameter that characterizes a tradeoff between performance optimization and delay in the virtual queues.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Proof: The proof is given in Appendix B. ■

Now, we present our opportunistic scheduling algorithm that involves cooperation between primary and secondary users to achieve better performance.

Scheduling Algorithm $Q_{1\mathrm{b}}$*:* At each time-slot t, observe the virtual queue backlog $X_m(t)$ for each PU m and the transmission rates, and choose (m,n,α) solving the following optimization problem:

$$Q_{1\mathrm{b}} = \arg\max_{(m,n,\alpha)\in\mathcal{F}_1} \left\{ \left(1+\frac{X_m(t)}{K}\right) U_{mn}(t) + V_{mn}(t) \right\}. \tag{19}$$

Then, update the virtual queues according to the queue dynamics in (14).

Note that we assume knowledge of the utility functions and channel states at the scheduler at each time-slot. Hence, the queue states are known constants in the above optimization problem. Comparing to Algorithm $Q_{1\mathrm{a}}$, let $\tilde{\lambda}_m(t) \triangleq 1 + \frac{X_m(t)}{K} \geq 1$. It is clear that both algorithms have exactly the same form. However, contrary to the algorithm in Section III-B, $Q_{1\mathrm{b}}$ does not require the knowledge of the statistics of the channel states or need the computation of online parameters.

Compared to the non-cooperative scheme in [9] that requires M multiplications, it can be seen that $Q_{1\mathrm{b}}$ requires $2M(N+1)$ operations ($M(N+1)$ multiplications and $M(N+1)$ additions). In addition, an algorithm to compute the best α for each pair (m,n) is needed in our scheme. However, the complexity of our scheme can be reduced by allowing the base station to select only a subset of available SUs with strong channels to be considered in scheduling.

We analyze our algorithm using Lyapunov drift with optimization [21]. We define a class of policies that will be useful to prove the optimality of the scheduling algorithm $Q_{1\mathrm{b}}$. Consider the class of scheduling algorithms $\mathcal{S}$ that schedule users according to a stationary and possibly randomized function of only the transmission rates and independent of the queue states. It was shown in [20] and [21] that optimality is achieved within the class of stationary policies $\mathcal{S}$ for a large class of network flow problems including fairness problems. Since the channel states are chosen from a finite set and the set $\{(m,n,\alpha)|\ m \in \{1,2,\ldots,M\}, n \in \{1,2,\ldots,N\}, \alpha \in [0,1]\}$ is closed and bounded, we have the following lemma (which can be proved using similar argument as in [20, Corollary 1]). Let the feasibility region of (11) be Λ and let $\epsilon \triangleq (\epsilon\ \epsilon \cdots \epsilon)$.

Lemma 2: If the vector $\mathbf{C} = (C_1\ C_2 \cdots C_M)$ is feasible (i.e., $\mathbf{C} \in \Lambda$), then there exists a stationary randomized policy Q_{s_1} that solves (11) and satisfies the following:

$$\mathbb{E}[W(Q_{\mathrm{s}_1},t)] = \bar{W}_1^* \tag{20}$$

$$\mathbb{E}\left[U_m^t(Q_{\mathrm{s}_1},t)\right] \geq C_m, m \in \{1,2,\ldots,M\} \tag{21}$$

where $\bar{W}_1^*$ is the optimal performance for (11) over all scheduling policies. Moreover, if $\mathbf{C}$ is strictly interior to Λ, then there exists $\epsilon > 0$ such that $(\mathbf{C}+\epsilon) \in \Lambda$ and a stationary scheduling policy $Q_{\mathrm{s}_1(\epsilon)}$ satisfying

$$\mathbb{E}\left[U_m^t(Q_{\mathrm{s}_1(\epsilon)},t)\right] \geq C_m + \epsilon, m \in \{1,2,\ldots,M\} \tag{22}$$

with an optimal total average utility $\bar{W}_1^*(\epsilon)$ such that $\bar{W}_1^*(\epsilon) \leq \bar{W}_1^*$ where $\bar{W}_1^*(\epsilon) \to \bar{W}_1^*$ as $\epsilon \to 0$.

We are now ready to present bounds on the performance of our proposed algorithm $Q_{1\mathrm{b}}$. The following theorem shows that all the virtual queues are strongly stable [21]. Hence, all time-average constraints in (11) are satisfied.

Theorem 3: If $\mathbf{C}$ is strictly in the interior of Λ, then the proposed algorithm in $Q_{1\mathrm{b}}$ stabilizes the virtual queues and achieves the following bounds:

$$\liminf_{t\to\infty} \frac{1}{t}\sum_{\tau=0}^{t-1} \mathbb{E}[W(Q_{1\mathrm{b}},\tau)] \geq \bar{W}_1^* - \frac{B_1}{K} \tag{23}$$

$$\limsup_{t\to\infty} \frac{1}{t}\sum_{\tau=0}^{t-1} \mathbb{E}\left[\sum_{m=1}^{M} X_m^2(\tau)\right] \leq \frac{B_1 + KW^{\max}}{\epsilon_{\max}} \tag{24}$$

where $W^{\max} = U^{\max} + V^{\max}$ and $B_1 = \frac{1}{2}\left(\sum_{m=1}^{M} C_m^2 + M(U^{\max})^2\right)$ and $\epsilon_{\max}$ is the largest ϵ such that $\mathbf{C}+\epsilon \in \Lambda$.

Proof: Consider the upper bound given by Lemma 1. From Lemma 2, there exists a stationary policy $Q_{\mathrm{s}_1(\epsilon)}$ that satisfies the constraints (22). By definition of $Q_{1\mathrm{b}}$, $\mathrm{RHS}_{Q_{1\mathrm{b}}} \leq \mathrm{RHS}_{Q_{\mathrm{s}_1(\epsilon)}}$ where RHS_Q is the right-hand side (RHS) of inequality (18) evaluated for the policy Q. Now consider evaluating $\mathrm{RHS}_{Q_{\mathrm{s}_1(\epsilon)}}$ using (22). Expanding the RHS of (18) and using the property that the utility is independent of queue states, it is straightforward to see that $\mathrm{RHS}_{Q_{\mathrm{s}_1(\epsilon)}} = B_1 - \epsilon\sum_{m=1}^{M} X_m(t) - K\bar{W}_1^*(\epsilon)$. It follows that

$$\Delta_1(\mathbf{X}(t)) - K\mathbb{E}[W(Q_{1\mathrm{b}},t)|\mathbf{X}(t),\mathbf{Y}(t)] \leq \mathrm{RHS}_{Q_{1\mathrm{b}}}$$
$$\leq \mathrm{RHS}_{Q_{\mathrm{s}_1(\epsilon)}} = B_1 - \epsilon\sum_{m=1}^{M} X_m(t) - K\bar{W}_1^*(\epsilon)$$

which is in exactly the same form of the condition in Theorem 1. Applying the result of Theorem 1, we have the following bounds:

$$\liminf_{t\to\infty} \frac{1}{t}\sum_{\tau=0}^{t-1} \mathbb{E}[W(Q_{1\mathrm{b}},\tau)] \geq \bar{W}_1^*(\epsilon) - \frac{B_1}{K} \tag{25}$$

$$\limsup_{t\to\infty} \frac{1}{t}\sum_{\tau=0}^{t-1} \mathbb{E}\left[\sum_{m=1}^{M} X_m^2(\tau)\right] \leq \frac{B_1 + KW^{\max}}{\epsilon} \tag{26}$$

where (26) follows since $0 \leq W(Q,t) \leq W^{\max}$ for all Q. The choice of ϵ affects the bounds only and does not affect the policy $Q_{1\mathrm{b}}$. Therefore, (25) and (26) can be optimized separately. Taking $\epsilon \to 0$ in (25) yields (23), and taking $\epsilon = \epsilon_{\max}$ in (26) yields (24), concluding the proof. ■

From (24) and (23), it is clear that the parameter K specifies a tradeoff between optimality and the average length of the virtual queues. Thus, for large virtual queues, the system experiences larger transient times to achieve the optimal performance and hence needs more time to adapt to possible changes in channel statistics [21].

D. Finite Backlog Case

In Sections III-B and IV-C, we determined optimal stationary and dynamic policies solving (11) with the assumption of all

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

SUs having infinite backlogs. In this section, we investigate the solution of the same problem when SUs have *finite* backlogs. The main difference in this case is that SUs will not be willing to cooperate to relay primary traffic if they do not have sufficient data of their own to transmit to the base station. Hence, the achievable utilities for the infinite backlogs case constitute an upper bound for the finite backlogs case. Also note that we do not further elaborate on a system when both primary and secondary users have finite backlogs since the analysis of this more complicated model provides little additional insight.

Let $A_n(t)$ and $L_n(t)$ be the amount of data arriving to SU n and the current backlog of the same user at time t, respectively. Under scheduling decision $Q = (m', n', \alpha)$, the time evolution of SU backlog $L_n(t)$, $n \in \{1, 2, \ldots, N\}$ is given as follows:

$$L_n(t+1) = \begin{cases} \max\{L_n(t) - \alpha(1-\beta)R_n^{\mathrm{s}}(t), 0\} + A_n(t), & \text{if } n = n' \\ L_n(t) + A_n(t), & \text{otherwise.} \end{cases} \tag{27}$$

Unlike the infinite backlog case, when scheduled, SU n' can transmit only $\min\{L_{n'}(t), \alpha(1-\beta)R_{n'}^{\mathrm{s}}(t)\}$ amount of data at time-slot t. Hence, under scheduling decision $Q = (m', n', \alpha)$, the utility of SU with finite backlogs is modified as follows:

$$\hat{V}_{m'n'}(Q,t) = \begin{cases} 0, & \text{if } n' = 0 \\ h_2(\min\{L_{n'}(t), \alpha(1-\beta)R_{n'}^{\mathrm{s}}(t)\}), & \text{otherwise.} \end{cases} \tag{28}$$

The following scheduling algorithm opportunistically selects a primary–secondary user pair.

Scheduling Algorithm $Q_{1\mathrm{c}}$*:* At each time-slot t, observe the virtual queue backlog $X_m(t)$ for every PU m, the actual queue backlog $L_n(t)$ for every SU n, the transmission rates, and choose (m, n, α) according to the following optimization problem:

$$Q_{1\mathrm{c}} = \underset{(m,n,\alpha)\in\mathcal{F}}{\arg\max} \left\{ \left(1 + \frac{X_m(t)}{K}\right) U_{mn}(t) + \hat{V}_{mn}(t) \right\}. \tag{29}$$

Then, update the virtual and actual queues according to the queue dynamics in (14) and (27), respectively.

Note that the only difference between $Q_{1\mathrm{c}}$ and $Q_{1\mathrm{b}}$ is that $V_{mn}(t)$ in $Q_{1\mathrm{b}}$ is replaced by $\hat{V}_{mn}(t)$ in $Q_{1\mathrm{c}}$. However, unlike $Q_{1\mathrm{b}}$, $Q_{1\mathrm{c}}$ is a nonconvex optimization problem for a given (m, n) pair due to the definition of $\hat{V}_{mn}(t)$ given in (28). The problem can be transformed into a convex program using auxiliary variables as shown in proof of Proposition 1, which is presented in Appendix C.

Proposition 1: For a fixed primary–secondary user pair (m, n), let α^* be the solution of (19), and β^* be defined as in (4). If $\alpha^*(1-\beta^*) > \frac{L_n(t)}{R_n(t)}$, then the optimal α^* for (29) is given as

$$\alpha_f = \frac{L_n(t) + R_{mn}^{\mathrm{r}}(t)}{R_{mn}^{\mathrm{r}}(t) + R_n^{\mathrm{s}}(t)}$$

and if $\alpha^*(1-\beta^*) \le \frac{L_n(t)}{R_n(t)}$, then the optimal solution for (29) is $\alpha_f = \alpha^*$.

The optimality of $Q_{1\mathrm{c}}$ can be shown in a similar fashion as of $Q_{1\mathrm{b}}$, and hence it is omitted for brevity.

E. Note on Feasibility

In the algorithms developed in Sections III-B and III-C, we assumed the feasibility of the set of constraints on the primary users' performance. In fact, the feasibility region characterization depends on the distribution of the channel gains of PUs and SUs. In general, there is no closed-form expression for feasibility region even if the channel distributions are known [9]. Since our scheduling schemes can only improve the performance of the primary-only network as a special case, the feasibility region given in [9] is strictly a subset of the feasibility region of our policy. In addition, it can be shown, using similar a similar argument as in [9], that our feasibility region is convex. Specifically, the region is a subset of an M-dimensional space such that the vertex on the mth axis is $(0\ 0 \cdots \mathbb{E}[\tilde{U}_m] \cdots 0)$, where $\mathbb{E}[\tilde{U}_m]$ is the average utility achieved by applying our cooperative algorithm on a network composed of only the mth PU in addition to N SUs.

Considering the example presented in Section III-B, $\Gamma_{\max}$ specified the maximum gain the cooperative system can achieve over the non-cooperative counterpart. It is clear that $\Gamma_{\max}$ can be characterized by the boundary of the feasibility region. More specifically, if $(a_1\ a_2 \cdots a_M)$ is the performance vector in the non-cooperative system defined in Section III-B, and if the feasibility region of our cooperative system is Λ, then $\Gamma_{\max}$ is given by

$$\Gamma_{\max} = \max_{\Gamma \ge 1} \ \Gamma \quad \text{s.t.} \ (\Gamma a_1\ \Gamma a_2 \cdots \Gamma a_M) \in \Lambda. \tag{30}$$

The solution to (30) can be determined numerically after estimating the feasibility region. A rigorous analysis to compute $\Gamma_{\max}$ is beyond the scope of this paper.

IV. Secondary Constraints and Long-Term Rewards

A. Formulation and Optimal Algorithm

In this section, we study a generalized version of the problem studied in Section III. Here, a long-term constraint is imposed on the minimum performance of each SU. More specifically, a portion of the primary utility achieved by cooperation is guaranteed for each cooperating SU in an average sense. In fact, the formulation of the problem below allows for the idea of *banking* between primary and secondary users. That is, in contrast to Section III where immediate rewards granted to cooperating SU are not lower-bounded, here SUs are guaranteed a specific share over a large number of time-slots. This is achieved by allowing α to take values such that $\alpha \le 1$, i.e., we lift the constraint imposed in the first inequality of (10) and SU can possibly be scheduled the entire time-slot for its own transmission at some time-slots. The problem is formulated as follows:

$$\begin{aligned} \max_{Q\in\mathcal{F}} \quad & \bar{W}(Q) \\ \text{s.t.:} \quad & 1)\ \mathbb{E}\left[U_m^t(Q,t)\right] \ge C_m, \qquad m \in \{1, 2, \ldots, M\} \\ & 2)\ \mathbb{E}\left[V_n^t(Q,t)\right] \ge \mathbb{E}\left[\sum_{m=1}^{M} \phi(U_{mn}(Q,t))\right], \\ & \qquad n \in \{1, 2, \ldots, N\} \end{aligned} \tag{31}$$

where $\phi(\cdot)$ is a nonnegative, nondecreasing scalar-valued function that determines the amount of utility cooperating SUs are guaranteed with respect to the primary utility achieved through this cooperation. We assume that the constraints in (31) are within the feasibility region. Define $\nu_n(Q,t) \triangleq \sum_{m=1}^{M} \phi(U_{mn}(Q,t))$. For each of the constraints above, we construct virtual queues such that the queue dynamics are given by

$$X_m(t+1) = [X_m(t) - U_m^t(Q,t)]^+ + C_m \quad (32)$$
$$Y_n(t+1) = [Y_n(t) - V_n^t(Q,t)]^+ + \nu_n(Q,t) \quad (33)$$

$m \in \{1,2,\ldots,M\}$, $n \in \{1,2,\ldots,N\}$. We assume that $\nu_n(Q,t) \le \nu^{\max}$ for all $n \in \{1,2,\ldots,N\}$, $t \ge 0$ and for all $Q \in \mathcal{F}$. Let $\mathbf{Y}(t) = (Y_1(t)Y_2(t)\cdots Y_N(t))$ be the vector of virtual queues of SUs. Define the following quadratic Lyapunov function and conditional Lyapunov drift:

$$L_2(\mathbf{X}(t),\mathbf{Y}(t)) \triangleq \frac{1}{2}\sum_{m=1}^{M} X_m^2(t) + \frac{1}{2}\sum_{n=1}^{N} Y_n^2(t) \quad (34)$$

$$\Delta_2(\mathbf{X}(t),\mathbf{Y}(t)) \triangleq \mathbb{E}\Big[L_2(\mathbf{X}(t+1),\mathbf{Y}(t+1)) - L_2(\mathbf{X}(t),\mathbf{Y}(t))|\mathbf{X}(t),\mathbf{Y}(t)\Big]. \quad (35)$$

We also define the following conditional expectation:

$$\bar{\nu}_n(Q,t) \triangleq \mathbb{E}[\nu_n(Q,t)|\mathbf{X}(t),\mathbf{Y}(t)]. \quad (36)$$

The Lyapunov drift in (35) is bounded by the following lemma where the proof is very similar to the proof of Lemma 1 and is omitted for brevity.

Lemma 3: For every time-slot t, the Lyapunov drift defined in (35) can be upper-bounded as follows:

$$\begin{aligned}\Delta_2(\mathbf{X}(t),\mathbf{Y}(t)) &- K\mathbb{E}[W(Q,t)|\mathbf{X}(t),\mathbf{Y}(t)] \\ &\le B_2 + \sum_{m=1}^{M} X_m(t)C_m - \sum_{m=1}^{M}(K+X_m(t))\bar{U}_m^t(Q,t) \\ &- \sum_{n=1}^{N}(K+Y_n(t))\bar{V}_n^t(Q,t) + \sum_{n=1}^{N} Y_n(t)\bar{\nu}_n(Q,t),\end{aligned}$$

where

$$B_2 = \frac{1}{2}\left(\sum_{m=1}^{M} C_m^2 + M(U^{\max})^2 + N((\nu^{\max})^2 + (V^{\max})^2)\right)$$

and K is a system parameter that characterizes a tradeoff between performance optimization and unfinished work in the virtual queues.

Now we present our opportunistic scheduling algorithm.

Scheduling Algorithm Q_2*:* At each time-slot t, the scheduler observes the virtual queue states $X_m(t)$, $Y_n(t)$ and the transmission rates $R_m^{\mathrm{p}}(t)$, $R_{mn}^{\mathrm{r}}(t)$, and $R_n^{\mathrm{s}}(t)$ for all $m \in \{1,2,\ldots,M\}$ and $n \in \{1,2,\ldots,N\}$, and then solves the following optimization problem:

$$Q_2 = \arg\max_{(m,n,\alpha)\in\mathcal{F}} \left\{\left(1+\frac{X_m(t)}{K}\right)U_{mn}(t) + \left(1+\frac{Y_n(t)}{K}\right)V_{mn}(t) - \left(\frac{Y_n(t)}{K}\right)\phi(U_{mn}(t))\right\}.$$

The virtual queues are then updated according to the queue dynamics in (32) and (33).

The structure of the scheduling policy suggests that when a secondary virtual queue $Y_n(t)$ is congested, then the system has a *debt* to pay to SU n. This is accomplished by favoring instantaneous allocations that reduce this debt by increasing payments (i.e., higher weight for $V_{mn}(t)$) and reduced additional debt (i.e., lower weight for $\phi(U_{mn}(t))$). Therefore, it is possible that the system allocates an entire time-slot to an SU without requiring the relay of a PU's data. Similarly, it is also possible that an SU relays primary data without obtaining immediate share of that time-slot to transmit its own data.

B. Algorithm Analysis

The analysis follows the same approach as in Section III-C. Let $\bar{W}_2^*$ be the optimal time-average system utility achieved over all scheduling policies for (31), and consider the class of stationary randomized scheduling algorithms that are independent of the queue states.

Lemma 4: If vectors $\mathbf{C}$ and $\mathbb{E}[\nu] = \mathbb{E}[(\nu_1\nu_2\cdots\nu_N)]$ are feasible, then there exists a stationary randomized policy Q_{s_2} that solves (31) and satisfies the following:

$$\mathbb{E}[W(Q_{\mathrm{s}_2},t)] = \bar{W}_2^* \quad (37)$$
$$\mathbb{E}\left[U_m^t(Q_{\mathrm{s}_2},t)\right] \ge C_m, \qquad m \in \{1,2,\ldots,M\} \quad (38)$$
$$\mathbb{E}\left[V_n^t(Q_{\mathrm{s}_2},t)\right] \ge \mathbb{E}\left[\nu_n(Q_{\mathrm{s}_2},t)\right] \quad (39)$$

where $\bar{W}_2^*$ is the optimal performance for (31) over all scheduling policies. Moreover, if $\mathbf{C}$ and $\mathbb{E}[\nu]$ are strictly interior to the feasibility region, then there exists $\epsilon' > 0$ and a stationary scheduling policy $Q_{\mathrm{s}_2(\epsilon')}$ satisfying:

$$\mathbb{E}\left[U_m^t(Q_{\mathrm{s}_2(\epsilon')},t)\right] \ge C_m + \epsilon', \qquad m \in \{1,2,\ldots,M\} \quad (40)$$
$$\mathbb{E}\left[V_n^t(Q_{\mathrm{s}_2(\epsilon')},t)\right] \ge \mathbb{E}\left[\nu_n(Q_{\mathrm{s}_2},t)\right] + \epsilon' \quad (41)$$

with an optimal total average utility $\bar{W}_2^*(\epsilon')$ such that $\bar{W}_2^*(\epsilon') \le \bar{W}_2^*$ where $\bar{W}_2^*(\epsilon') \to \bar{W}_2^*$ as $\epsilon' \to 0$.

The following theorem is parallel to Theorem 3, and the proof uses Lemmas 3 and 4.

Theorem 4: If the constraints in (31) are feasible, then the proposed algorithm Q_2 stabilizes the virtual queues and achieves the following bounds:

$$\liminf_{t\to\infty} \frac{1}{t}\sum_{\tau=0}^{t-1} \mathbb{E}[W(Q_2,\tau)] \ge \bar{W}_2^* - \frac{B_2}{K}$$

$$\limsup_{t\to\infty} \frac{1}{t}\sum_{\tau=0}^{t-1} \mathbb{E}\left[\sum_{m=1}^{M} X_m^2(\tau) + \sum_{n=1}^{N} Y_n^2(\tau)\right] \le \frac{B_2 + KW^{\max}}{\epsilon'_{\max}}$$

where $\bar{W}_2^*$ is the optimal value for the time-average expected utility, $W^{\max} = U^{\max} + V^{\max}$, and $\epsilon'_{\max}$ is the largest ϵ' such that constraints (40) and (41) are feasible.

C. Downlink Case

The analysis and algorithms developed in this paper are mainly focusing on the uplink scenario. However, it can be easily seen that the same analysis and results can be applied in the downlink case. In this section, we state the differences for the sake of completeness. In the downlink case and when cooperation is feasible, the base station (or access point) D uses the first $(1-\alpha)$ fraction of the given time-slot to transmit to an SU n. Then, SU relays data to the scheduled PU m during the next $\alpha\beta$ fraction. The rest of the time-slot is dedicated to the downlink of SU's data. We denote the transmission rates from base station to primary and secondary users m and n as $\tilde{R}^{\mathrm{p}}_m(t)$ and $\tilde{R}^{\mathrm{s}}_n(t)$, respectively, at time-slot t. We also denote the transmission rate from SU n to PU m as $\tilde{R}^{\mathrm{r}}_{mn}$. Similar to the uplink case, the utility functions for a scheduled pair (m,n) are defined as

$$\tilde{U}_{mn}(Q,t) = \begin{cases} h_3(\tilde{R}^{\mathrm{p}}_m(t)), & \text{if } n=0 \\ h_3((1-\alpha)\tilde{R}^{\mathrm{s}}_n(t)), & \text{otherwise} \end{cases} \quad (42)$$

$$\tilde{V}_{mn}(Q,t) = \begin{cases} 0, & \text{if } n=0 \\ h_4(\alpha(1-\beta)\tilde{R}^{\mathrm{s}}_n), & \text{otherwise} \end{cases} \quad (43)$$

where $h_3(\cdot)$ and $h_4(\cdot)$ are some nonnegative nondecreasing concave functions. Feasibility regions can be defined as follows. For immediate rewards problem (11), it can be seen that optimal scheduling policy selects the tuple (m,n,α) from $\mathcal{F}_{d_1}$ satisfying

$$\tilde{R}^{\mathrm{p}}_m(t) \leq (1-\alpha)\tilde{R}^{\mathrm{s}}_n(t) \leq \alpha\tilde{R}^{\mathrm{r}}_{mn}(t) \quad (44)$$

for some $0<\alpha<1$. For the long-term rewards problem, the feasibility set $\mathcal{F}_d$ is constructed from all tuples (m,n,α) such that

$$0 < (1-\alpha)\tilde{R}^{\mathrm{s}}_n(t) \leq \alpha\tilde{R}^{\mathrm{r}}_{mn}(t) \quad (45)$$

for some $0 \leq \alpha \leq 1$, where

$$\beta^* = \frac{(1-\alpha)\tilde{R}^{\mathrm{s}}_n}{\alpha\tilde{R}^{\mathrm{r}}_{mn}} \quad (46)$$

for both cases. Using definitions (42)–(46) in Algorithms $Q_{1\mathrm{a}}$, $Q_{1\mathrm{b}}$, $Q_{1\mathrm{c}}$, Q_2, the same results can be obtained.

V. Maximization of Net Benefit

Transmission power control is an essential part of wireless communications, and it is especially important for wireless devices with limited energy resources. In this section, our objective is to investigate scheduling and power control policies when PUs and SUs are allowed to cooperate. For this purpose, we define "net benefit" of a user as the difference between the total average utility and the total weighted average energy consumption. Our objective is to determine optimal dynamic joint power control and scheduling policy that maximizes aggregate net benefits of primary and secondary systems considering immediate rewards.

Let $E^{\mathrm{p}}_{mn}(Q,t)$ and $E^{\mathrm{s}}_{mn}(Q,t)$ be the consumed energy by PU m and SU n under policy Q at time-slot t, respectively. Also we define the net benefit of primary and secondary users, $F_{mn}(Q,t)$ and $G_{mn}(Q,t)$, as follows [13], [23]:

$$F_{mn}(Q,t) = U_{mn}(Q,t) - \rho\cdot E^{\mathrm{p}}_{mn}(Q,t) \quad (47)$$

$$G_{mn}(Q,t) = V_{mn}(Q,t) - \rho\cdot E^{\mathrm{s}}_{mn}(Q,t) \quad (48)$$

where ρ is the cost per unit transmission energy. Let the total net benefit of both systems when policy Q is employed at a given time-slot t, be $H(Q,t)$. Then

$$H(Q,t) = \sum_{m=1}^{M}\sum_{n=0}^{N} F_{mn}(Q,t) + G_{mn}(Q,t). \quad (49)$$

Define $F^t_m(Q,t) \triangleq \sum_{n=0}^{N} F_{mn}(Q,t)$, $G^t_n(Q,t) \triangleq \sum_{m=1}^{M} G_{mn}(Q,t)$, $E^{\mathrm{p}}_m(Q,t) \triangleq \sum_{n=0}^{N} E^{\mathrm{p}}_{mn}(Q,t)$. These values are upper-bounded such that $F^t_m(Q,t) \leq F^{\max}$, $G^t_n(Q,t) \leq G^{\max}$ and $E^{\mathrm{p}}_m(Q,t) \leq E^{\max}$.

The total expected net benefit for time-slot t is defined as $\bar{H}(Q,t) \triangleq \mathbb{E}[H(Q,t)]$. Now, we consider the following optimization problem:

$$\begin{aligned} \max_{Q\in\mathcal{G}} \quad & \bar{H}(Q,t) \\ \text{s.t.} \quad & 1)\ \mathbb{E}\left[U^t_m(Q,t)\right] \geq C_m, \qquad m\in\{1,2,\ldots,M\} \\ & 2)\ \mathbb{E}\left[G^t_n(Q,t)\right] \geq 0, \qquad n\in\{1,2,\ldots,N\} \end{aligned} \quad (50)$$

where C_m is the minimum performance constraint for each PU m, and $\mathcal{G}$ is the set of feasible policies to be defined later. If SU n relays the data of PU m, then the net benefit of SU may be negative. In this case, SU n may not be willing to join cooperation. Hence, we impose the nonnegativity constraint on the expected net benefit of secondary system in (50).

Here, the scheduling policy Q is a rule that selects $(m,n,\alpha,\beta,P^{\mathrm{p}}_{mn},P^{\mathrm{s}}_{mn})$ at time-slot t where P^{p}_{mn} and P^{s}_{mn} are the transmission powers from PU m to SU n and from SU n to destination to relay data of PU m, respectively. Under this model, if cooperation is infeasible, PU m transmits directly to the destination at power of P^{p}_{m0}. In addition, SU n transmits its own data to the destination using the power level P^{s}_{0n}. We use the same notational convention with the data rates R.

Even though our results are general for all channel-state distributions, in numerical evaluations, we assume all channels to be *Gaussian*. We represent the uplink channel for PU m, cross channel between PU m and SU n $(m\in\{1,2,\ldots,M\}, n\in\{1,2,\ldots,N\})$ with a power gain (magnitude square of the channel gains) $l_{m0}(t)$ and $l_{mn}(t)$, respectively, at time-slot t. We normalize the power gains such that the (additive Gaussian) noise has unit variance. In the following, we employ information theoretic expressions for the achievable data rates. Thus

$$R^{\mathrm{p}}_{mn}(t) = \log(1+P^{\mathrm{p}}_{mn}(t)l_{mn}(t))$$

$$R^{\mathrm{s}}_{mn}(t) = \log(1+P^{\mathrm{s}}_{mn}(t)l_n(t)).$$

In addition, peak power constraints are imposed in the network such that

$$0 \leq P^{\mathrm{p}}_{mn}, \leq P^m_{\max} \quad (51)$$

$$0 \leq P^{\mathrm{s}}_{mn}, P^{\mathrm{s}}_n \leq P^n_{\max} \quad (52)$$

$\forall\, m\in\{0,1,2,\ldots,M\}$ and $n\in\{0,1,2,\ldots,N\}$. Given these definitions, the optimization problem in (50) along with constraints (3), (51), and (52) is a nonconvex optimization problem. Hence, it is hard to find a closed-form solution for this problem.

Here, we consider a simplified version of the problem, where the values of α and β are taken as constants satisfying the following:

$$(1-\alpha) = \alpha\beta = \alpha(1-\beta) = \frac{1}{3}. \tag{53}$$

We fix the duration of each of the three phases in our cooperative scheme to one third of the time-slot as implied by (53). Clearly, this is a suboptimal solution since we do not consider α and β as decision variables. Nevertheless, this model has practical applicability. In many real networks, including cellular networks [24], users are assigned fixed time-slots to complete their transmissions. In such systems, transmission power control is the main tool to adjust transmission rates.

Since transmission durations are equal, the following holds for all $m \in \{1,2,\dots,M\}$ and $n \in \{1,2,\dots,N\}$:

$$R^{\mathrm{p}}_{mn}(t) = R^{\mathrm{s}}_{mn}(t). \tag{54}$$

Thus, given instantaneous channel conditions and for a given value of P^{p}_{mn}, we can determine P^{s}_{mn}. Hence, we use the notation $Q = (m, n, P^{\mathrm{p}}_{mn}, P^{\mathrm{s}}_{0n})$ to denote the joint scheduling and power control decision of policy Q.

Let $\mathcal{G}$ be the set of tuples $(m, n, P^{\mathrm{p}}_{mn}, P^{\mathrm{s}}_{0n})$ satisfying (51)–(53) and the following condition:

$$R^{\mathrm{p}}_{m0}(t) \le \frac{1}{3} R^{\mathrm{p}}_{mn}(t) \tag{55}$$

where $\forall\ m \in \{1,2,\dots,M\}$, $n \in \{1,2,\dots,N\}$. Clearly, the above inequality states that SU n is allowed to join cooperation if it improves the instantaneous utility of PU m. Moreover, set $\mathcal{G}$ is constructed from all $(m, n, P^{\mathrm{p}}_{mn}, P^{\mathrm{s}}_{0n})$. $E^{\mathrm{p}}_{mn}(Q,t)$ and $E^{\mathrm{s}}_{mn}(Q,t)$ under scheduling policy Q are given as follows:

$$E^{\mathrm{p}}_{mn}(Q,t) = \begin{cases} P^{\mathrm{p}}_{m0}, & \text{if } n = 0 \\ \frac{1}{3}P^{\mathrm{p}}_{mn}, & \text{otherwise} \end{cases} \tag{56}$$

$$E^{\mathrm{s}}_{mn}(Q,t) = \begin{cases} 0, & \text{if } n = 0 \\ \frac{1}{3}P^{\mathrm{s}}_{mn} + \frac{1}{3}P^{\mathrm{s}}_{0n}, & \text{otherwise.} \end{cases} \tag{57}$$

We solve (50) using the stochastic network optimization tool of [21]. For the first set of constraints in (50), we construct the system of virtual queues defined in (14). Also, for each of the second set of constraints in (50), we construct virtual queues $S_n(t)$ with queue dynamics given as follows:

$$S_n(t+1) = [S_n(t) - G^t_n(Q,t)]^+ \tag{58}$$

$n \in \{1,2,\dots,N\}$. Now we present our joint power allocation and scheduling algorithm.

Scheduling Algorithm Q_3*:* At each time-slot t, observe the virtual queue states $X_m(t)$ and $S_n(t)$ and select $(m, n, P^{\mathrm{p}}_{mn}, P^{\mathrm{s}}_{0n})$ solving the following optimization problem:

$$Q_3 = \underset{(m,n,P^{\mathrm{p}}_{mn},P^{\mathrm{s}}_{0n})\in\mathcal{G}}{\arg\max} \left\{ \left(1 + \frac{X_m(t)}{K}\right) U_{mn}(t) - \rho \cdot E^{\mathrm{p}}_{mn}(t) + \left(1 + \frac{S_n(t)}{K}\right) G_{mn}(t) \right\}.$$

Theorem 5: If the constraints in (50) are feasible, then the proposed algorithm Q_3 stabilizes the virtual queues and achieves the following bounds:

$$\liminf_{t\to\infty} \frac{1}{t}\sum_{\tau=0}^{t-1} \mathbb{E}[H(Q,\tau)] \ge \bar{H}^* - \frac{B_3}{K}, \quad \limsup_{t\to\infty} \frac{1}{t}\sum_{\tau=0}^{t-1} \mathbb{E}\left[\sum_{m=1}^{M} X^2_m(\tau) + \sum_{n=1}^{N} S^2_n(\tau)\right] \le \frac{B_3 + KH^{\max}}{\varepsilon_{\max}}$$

where $\bar{H}^*$ is the solution of (50) among class of policies $\mathcal{G}$, $H_{\max}$ is the maximum net benefit that can be achieved at any time-slot, $B_3 = \frac{1}{2}\left(\sum_{m=1}^{M} C^2_m + M(U^{\max})^2 + N(G^{\max})^2\right)$, and $\varepsilon_{\max}$ can be defined similarly as $\epsilon_{\max}$ in Section III-C.

Proof: The proof is similar to the proof of Theorem 3 and involves finding an upper bound on the conditional Lyapunov drift for a quadratic Lyapunov function of virtual queues $X_m(t)$ and $S_n(t)$ and uses Lemma 4. ∎

It can be seen that when there is no cost for power consumption, i.e., $\rho = 0$, the scheduling algorithm Q_3 has the same form as the scheduling algorithm $Q_{1\mathrm{b}}$. In this sense, $Q_{1\mathrm{b}}$ can be considered as a specific case of Q_3.

VI. Numerical Results

In this section, we present four experiments to evaluate the performance of our proposed scheduling and power control schemes. First, we simulate a wireless network with $M = 4$ PUs and a varying number of SUs where we present a comparison between our cooperative scheduling scheme and the optimal non-cooperative scheme. Channel states vary randomly between "Good" and "Bad" for primary and secondary users. Transmission rates corresponding to the channel states {"Good", "Bad"} are set to $\{100,\ 15\}$ units/slot, and channel states evolve independently across users and across time. For all pairs (m,n), the transmission rates are given by $R^{\mathrm{p}}_m(t) = \{100,\ 10\}$ with probability $\{0.5,\ 0.5\}$, $R^{\mathrm{r}}_{mn}(t) = \{100,\ 10\}$ with probability $\{0.6,\ 0.4\}$, and $R^{\mathrm{s}}_n(t) = \{100,\ 10\}$ with probability $\{0.6,\ 0.4\}$. Given these channel statistics, we run the simulation for 200 000 time-slots, which is sufficient for the convergence of algorithm $Q_{1\mathrm{a}}$ for the above channel statistics. For the utility functions, we employ the functions $h_1(x) = h_2(x) = \log(1+x)$.

For the constraints on the performance of PUs in the non-cooperative system, we adopt a *fair sharing* policy, that is, the achievable primary system utility is to be divided evenly among the PUs. We set the constraints C_m in (11) as in the example given in Section III-B for the sake of comparison with non-cooperative systems. Applying scheduling policy $Q_{1\mathrm{a}}$, we let $\Gamma = 1.01$ and use a stochastic approximation approach to estimate the parameters λ^*_m, $m \in \{1,2,\dots,M\}$ as follows. First, from the constraints on the PU performance, we see that for $m \in \{1,2,\dots,M\}$, λ^*_m is the root to the following equation:

$$f_m(\lambda_m) = (\lambda_m - 1)\left(\mathbb{E}\left[\sum_{n=0}^{N} U_{mn}(Q_{1\mathrm{a}}, t)\right] - C_m\right).$$

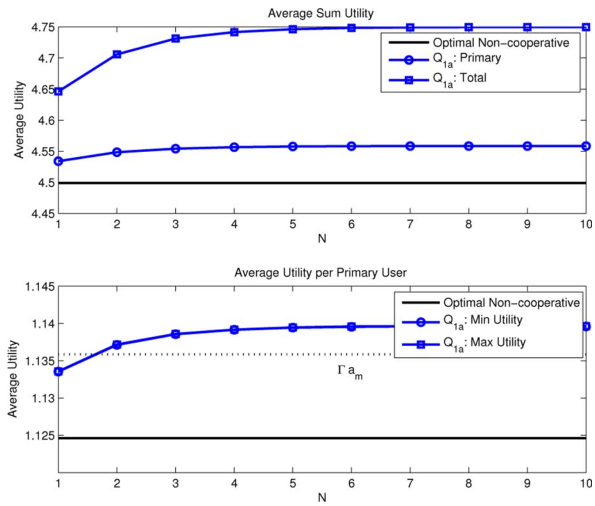

Fig. 3. Average total system utility versus N.

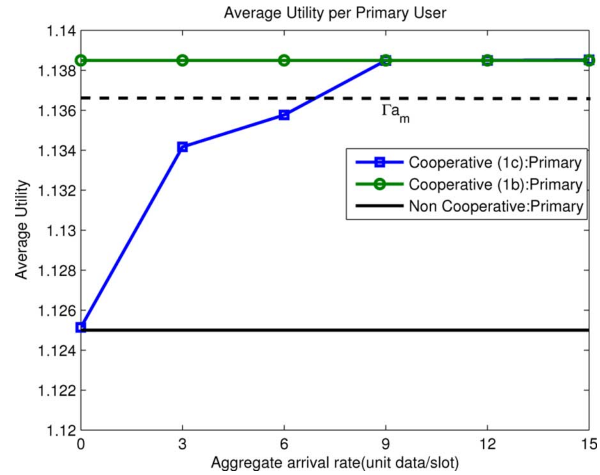

Fig. 4. Average PU utility versus SU arrival rate.

However, since we only have knowledge about the instantaneous channel gains, we need to estimate the distribution of the utility functions. Hence, using the observation we have, we can write an estimate g_m^k of f_m^k as

$$g_m^k(\lambda_m) = (\lambda_m^k - 1)\left(\sum_{n=0}^{N} U_{mn}(Q_{1\mathrm{a}}, t) - C_m\right)$$

where k is the iteration index. Since this estimator is unbiased $\left(\mathbb{E}[g_m^k - f_m^k(\lambda_m^k)] = 0, \forall\, m\right)$, then we can use a stochastic approximation algorithm of the form

$$\lambda_m^{k+1} = \lambda_m^k - \delta^k g_m^k$$

where δ^k can be taken to be $1/k$ [9]. For a given time-slot t, it can be shown that (12) is a concave function in α. The optimization is then done over all pairs so that α satisfies the condition (10), then the tuple $(m, n, \alpha) \in \mathcal{F}_1$ that maximizes (12) is selected by the scheduler at this time-slot. For a pair (m, n), since the objective function is concave and the constraint is linear in α, then Karush–Kuhn–Tucker (KKT) conditions are both necessary and sufficient to solve (12), along with (10) [22].

In Fig. 3, the average system utility is plotted with respect to the number of cognitive users. The cooperative scheme achieves higher average system utility compared to the non-cooperative scheme. For $N = 1$, the constraints are infeasible to achieve, however the $Q_{1\mathrm{a}}$ still performs better than the non-cooperative policy. For $N \geq 1$, the constraints are feasible. Moreover, exploiting the opportunity that relaying offers, we could achieve nonzero secondary system average utility. Fig. 3 also shows the per-user (primary) performance. It can be seen that the smallest per-user performance is still better than the non-cooperative case with at least the value Γ. Scheduling policy $Q_{1\mathrm{b}}$ yields the same results as $Q_{1\mathrm{a}}$ under the same channel parameter selection, using $K = 5000$.

In the second experiment, we evaluate the performance of our scheduling policy when SUs have finite backlogs. Fig. 4 depicts the average utility per PU versus average aggregate arrival rate when SUs have finite backlogs. In this experiment, we apply scheduling policy $Q_{1\mathrm{c}}$ for the same channel statistics used in the first simulation and set $M = 4$, $N = 3$. It is assumed that at $t = 0$, all queues of SUs are empty, i.e., $L_n(0) = 0\ \forall\, n$, and new data arrive at secondary users with the same arrival rate according to independent Bernoulli processes. If all SUs have no data to transmit at a time-slot t, i.e., $L_n(t) + A_n(t) = 0\ \forall\, n$, they do not join cooperation. In this case, PUs transmit to the base station directly. Hence, the primary utility achieved is the same as in the non-cooperative case. As the arrival rate increases, SUs start to be backlogged and join the network. Therefore, the utility of primary users increases as well. When secondary users have sufficiently large backlog, primary users can achieve the utility obtained in the infinite backlog case at most. Next, in the third experiment, we simulate the long-term rewards scenario and apply scheduling policy Q_2 when $M = 4$ and the number of SUs is fixed at $N = 5$ for the same channel statistics used in the first simulation. In Fig. 5, the running average of the expected utility of SU 3 up to time t is plotted and compared to the average primary utility achieved through cooperation with SU 3 for $\phi(U_{mn}(Q, t)) = bU_{mn}(Q, t)$, and $b = 1.2$. In other words, for every packet of primary traffic SU n relays, SU n is rewarded by being scheduled to send 1.2 of its own packets on the long term. In this experiment, we set $C_m = 3.5\ \forall\, m$. Stability is achieved for all primary and secondary virtual queues.

Finally, we evaluate the performance of the scheduling algorithm Q_3 in terms of net benefit, average utility, and energy consumption. We assume a scenario where $M = 4$, $\rho = 1$ and for varying number of N. The utility achieved by primary users in the non-cooperative system where the objective is to maximize the net benefit and $\rho = 1$ are taken as the minimum requirements in this experiment. We set $C_m = 1.095\ \forall\, m$ $(\Gamma = 1)$. We assume that all users have infinite backlog, and $P_{\max}^m = P_{\max}^n = 10$. In addition, the channel gains are adjusted so that the transmission rates in the first simulation with the maximum transmission power can be obtained. In Fig. 6, the average system net benefit with respect to the number of SUs is plotted. Clearly, our cooperative scheme performs better than non-cooperative scheme in terms of average system net benefit. Note that all constraints in (50) are satisfied in this experiment. In Fig. 7, the average utility and energy consumption of primary and secondary

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

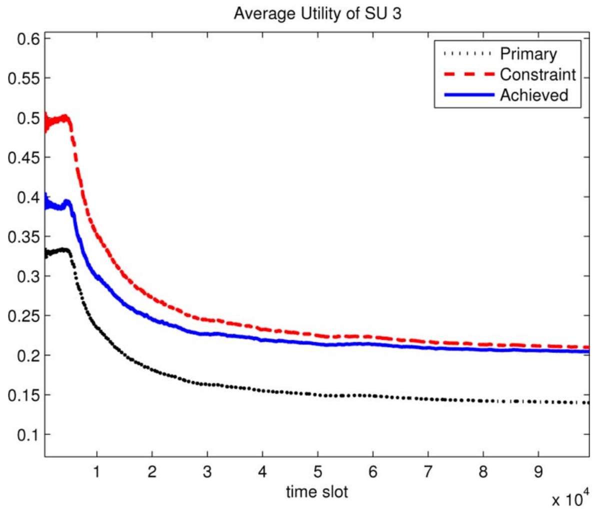

Fig. 5. Running average utility of SU 3.

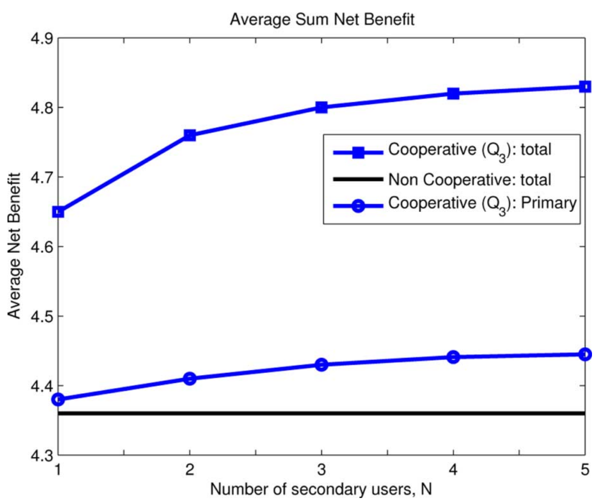

Fig. 6. Average net benefit versus N.

users are plotted for increasing values of ρ. We set $M = 4$, $N = 5$, and $C_m = 1 \;\forall\; m$. If users do not pay a penalty cost for transmission power (i.e., $\rho = 0$), they transmit at the maximum transmission power, thus the total average energy consumption is maximized. It is seen that the average utility achieved per a primary user, which is equal to 1, is same as the minimum requirement of the corresponding PU. However, energy consumption of primary system reduces by allowing SUs to join cooperation. That is, PUs conserve energy by cooperating with SUs and applying power control algorithm Q_3. Thus, primary net benefit increases as shown in Fig. 6. As ρ increases, SUs become unwilling to join cooperation and prefer not to transmit in order to satisfy their net benefit requirements. Consequently, their average utility and energy consumption decrease. When SUs leave cooperation, PUs transmit directly to destination, and only the minimum primary requirements are achieved since the energy expenditure is costly. The average energy consumption is fixed at approximately 0.025 in this case.

VII. Conclusion

We have developed optimal scheduling policies by exploiting the time-varying channel conditions and realizing the benefits of

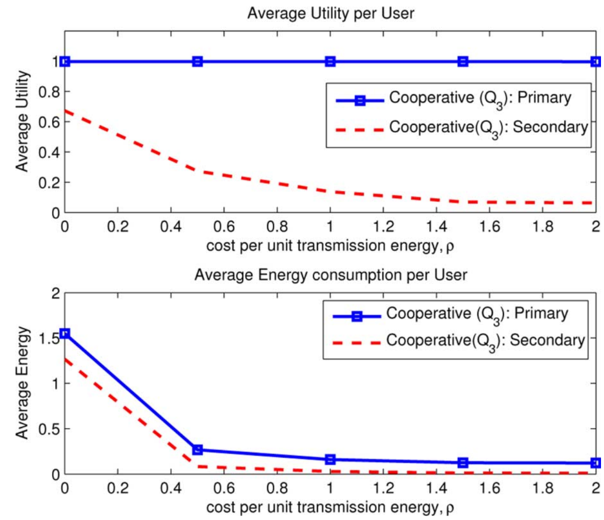

Fig. 7. Average utility and energy consumption versus ρ.

cooperative transmission in a cognitive radio network. The proposed model is based on the idea that secondary users can have opportunity to transmit their own data if they can improve the performance of a primary user via cooperation. First, we have studied the immediate reward strategy considering the cases in which secondary users have infinite and finite backlogs. Then, long-term rewards is studied where we introduce the idea of banking between primary and secondary users. In this banking model, secondary users are guaranteed a portion of the primary utility on a long-term basis, instead of immediate utility. Finally, we have investigated the energy-utility tradeoff by considering the power control and scheduling problems jointly where the objective is to maximize the net benefit. Numerical results show that our cooperative schemes improve the basic primary network performance in addition to giving SUs the ability to communicate in return for their cooperation. In addition, energy consumption of primary system can be reduced by allowing secondary users to join the network. Possible directions for future work include designing lower-complexity algorithms. Another line of research would be to study the delay performance of such a cooperative scheme.

Appendix A
Proof of Theorem 2

The proof is similar to the proof of the optimal policies in [9]. However, the scheduling policy $Q_{1\mathrm{a}}$ in our work decides a tuple of three variables each time-slot instead of only choosing a PU.

In the following, we drop the parameter t. Let Q be a scheduling policy satisfying $\mathbb{E}\left[\sum_{n=0}^{N} U_{mn}(Q)\right] \geq C_m$ for all $m \in \{1, 2, \ldots, M\}$. Then

$$\bar{W}(Q) \leq \bar{W}(Q) + \sum_{m=1}^{M} (\lambda_m^* - 1)\left(\sum_{n=0}^{N} \mathbb{E}\left[U_{mn}(Q)\right] - C_m\right)$$
$$= \sum_{m=1}^{M}\sum_{n=0}^{N} \mathbb{E}[U_{mn}(Q) + V_{mn}(Q)]$$
$$+ \sum_{m=1}^{M}\sum_{n=0}^{N} (\lambda_m^* - 1)\mathbb{E}[U_{mn}(Q)] - \sum_{m=1}^{M} (\lambda_m^* - 1) C_m.$$

After rearranging the right-hand side of the above inequality, we have

$$\bar{W}(Q) \leq \sum_{m=1}^{M}\sum_{n=0}^{N}\mathbb{E}[\lambda_m^* U_{mn}(Q) + V_{mn}(Q)] - \sum_{m=1}^{M}(\lambda_m^* - 1)C_m$$

where the first inequality follows since $\lambda_m^* \geq 1$. From the definition of $Q_{1\mathrm{a}}$, we have

$$\sum_{m=1}^{M}\sum_{n=0}^{N}\mathbb{E}[\lambda_m^* U_{mn}(Q) + V_{mn}(Q)] \leq \sum_{m=1}^{M}\sum_{n=0}^{N}\mathbb{E}[\lambda_m^* U_{mn}(Q_{1\mathrm{a}}) + V_{mn}(Q_{1\mathrm{a}})].$$

Therefore, we can write

$$\begin{aligned}\bar{W}(Q) &\leq \sum_{m=1}^{M}\sum_{n=0}^{N}\mathbb{E}[\lambda_m^* U_{mn}(Q_{1\mathrm{a}}) + V_{mn}(Q_{1\mathrm{a}})] - \sum_{m=1}^{M}(\lambda_m^* - 1)C_m\\ &= \bar{W}(Q_{1\mathrm{a}}) + \sum_{m=1}^{M}(\lambda_m^* - 1)\left(\sum_{n=0}^{N}\mathbb{E}[U_{mn}(Q_{1\mathrm{a}})] - C_m\right)\\ &= \bar{W}(Q_{1\mathrm{a}})\end{aligned}$$

where the last step follows from the properties of λ_m^*, completing the proof.

Appendix B
Proof of Lemma 1

We use the simplified notation $\bar{U}_m^t$ in place of $\bar{U}_m^t(Q,t)$. From the dynamics of the virtual queues (14), we can write

$$X_m^2(t+1) \leq X_m^2(t) + C_m^2 + (U_m^t(t))^2 - 2X_m(t)[U_m^t(t) - C_m]$$

for $m \in \{1, 2, \ldots, M\}$, where the above inequality follows from the fact that $\left([a]^+\right)^2 \leq (a)^2 \forall\, a$. Therefore, the Lyapunov drift in (16) can be upper-bounded as

$$\Delta_1(\mathbf{X}(t)) \leq \sum_{m=1}^{M}\frac{1}{2}C_m^2 + \frac{1}{2}\mathbb{E}[(U_m^t(t))^2|\mathbf{X}(t)] - X_m(t)\bar{U}_m^t + X_m(t)C_m.$$

Using the bounds on the utility functions $U^{\max}$, we have

$$\Delta_1(\mathbf{X}(t)) \leq B_1 + \sum_{m=1}^{M}X_m(t)(C_m - \bar{U}_m^t)$$

where $B_1 = \frac{1}{2}\left(\sum_{m=1}^{M}C_m^2 + M(U^{\max})^2\right)$. Subtracting the term $K\mathbb{E}[W(Q,t)|\mathbf{X}(t)]$ from both sides, expanding terms, rearranging terms, and using $V_0^t(t) = 0\ \forall\, t$, (18) follows.

Appendix C
Proof of Proposition 1

By using a similar idea as in [25], we introduce an auxiliary variable π_{mn} to obtain a convex problem. For a given (m,n) pair, the solution of the following problem gives the optimal α:

$$\max_{\alpha,\pi_{mn}} \left\{\tilde{\lambda}_m(t)h_1\left((1-\alpha)R_{mn}^{\mathrm{r}}(t)\right) + h_2(\pi_{mn})\right\} \quad (59)$$

$$\text{s.t.}\quad R_m^{\mathrm{p}}(t) \leq (1-\alpha)R_{mn}^{\mathrm{r}}(t) \quad (60)$$

$$(1-\alpha)R_{mn}^{\mathrm{r}}(t) \leq \alpha R_n^{\mathrm{s}}(t) \quad (61)$$

$$L_n(t) \geq \pi_{mn} \quad (62)$$

$$\alpha(R_{mn}^{\mathrm{r}}(t) + R_n^{\mathrm{s}}(t)) - R_{mn}^{\mathrm{r}}(t) \geq \pi_{mn}. \quad (63)$$

Note that the objective function is modified and is expressed with respect to the auxiliary variable π_{mn} and α. It is clear that (59) is jointly concave in α and π_{mn} (i.e., Hessian matrix of the objective function is negative semidefinite since all eigenvalues of it are negative). Hence, at a given time-slot, first-order optimality conditions given by KKT equations are sufficient for global optimality of (59)–(63).

In the following, we drop the parameter t. Let μ_3 and μ_4 be the dual variables associated to the constraints in (62) and (63), respectively. Let μ_3^*, μ_4^*, α^* and π_{mn}^* denote the values taken by μ_3, μ_4, α and π_{mn} at the optimal solution respectively. From KKT conditions, the following equalities hold:

$$\frac{dh_2}{d\pi_{mn}}|_{\pi_{mn}=\pi_{mn}^*} - \mu_3^* - \mu_4^* = 0 \quad (64)$$

$$\mu_3^*(L_n - \pi_{mn}^*) = 0 \quad (65)$$

$$\mu_4^*\left(\alpha^* - \frac{\pi_{mn}^* + R_{mn}^{\mathrm{r}}}{R_{mn}^{\mathrm{r}} + R_n^{\mathrm{s}}}\right) = 0. \quad (66)$$

It is clear that the objective function in (59) decreases with α. If $\pi_{mn}^* = L_n$, at the optimal solution α takes its minimum value, which is given by (63). Thus, α^* is given as follows:

$$\alpha^* = \frac{L_n + R_{mn}^{\mathrm{r}}}{R_{mn}^{\mathrm{r}} + R_n^{\mathrm{s}}}.$$

If $\pi_{mn}^* \neq L_n$, (64) and (65) yield the following equalities:

$$\mu_3^* = 0$$

$$\mu_4^* = \frac{dh_2}{d\pi_{mn}}|_{\pi_{mn}=\pi_{mn}^*} > 0.$$

From (66), α^* is given by

$$\alpha^* = \frac{\pi_{mn}^* + R_{mn}^{\mathrm{r}}}{R_{mn}^{\mathrm{r}} + R_n^{\mathrm{s}}}.$$

Thus

$$\pi_{mn}^* = \alpha^*(R_{mn}^{\mathrm{r}} + R_n^{\mathrm{s}}) - R_{mn}^{\mathrm{r}}. \quad (67)$$

Clearly, (67) states that α^* is obtained following the same steps in infinite backlog case. This completes the proof.

References

[1] K. Khalil, M. Karaca, O. Ercetin, and E. Ekici, "Optimal scheduling in cooperate-to-join cognitive radio networks," in *Proc. IEEE INFOCOM*, 2011, pp. 3002–3010.

[2] J. Mitola, III and G. Q. Maguire Jr., “Cognitive radio: Making software radios more personal,” *IEEE Pers. Commun. Mag.*, vol. 6, no. 4, pp. 13–18, Aug. 1999.

[3] I. F. Akyildiz, W. Y. Lee, M. Vuran, and S. Mohanty, “NeXt generation/dynamic spectrum access/cognitive radio wireless networks: A survey,” *Comput. Netw.*, vol. 50, no. 13, pp. 2127–2159, Sep. 2006.

[4] J. M. Peha, “Approaches to spectrum sharing,” *IEEE Commun. Mag.*, vol. 43, no. 2, pp. 10–12, Feb. 2005.

[5] J. O. Neel, “Analysis and design of cognitive radio networks and distributed radio resource management algorithms,” Ph.D., Virginia Polytechnic Institute, Blacksburg, VA, 2006.

[6] S. Shakkottai and A. Stolyar, “Scheduling algorithms for a mixture of real-time and non-real-time data in HDR,” Bell Laboratories, Tech. Rep., 2000.

[7] M. Andrews, K. Kumaran, K. Ramanan, A. Stolyar, P. Whiting, and R. Vijayakumar, “Providing quality of service over a shared wireless link,” *IEEE Trans. Magn.*, vol. 39, no. 2, pp. 150–153, Feb. 2001.

[8] I. H. Hou and P. R. Kumar, “Utility-optimal scheduling in time-varying wireless networks with delay constraints,” in *Proc. Mobihoc*, 2010, pp. 31–40.

[9] X. Liu, E. K. P. Chong, and N. B. Shroff, “A framework for opportunistic scheduling in wireless networks,” *Comput. Netw.*, vol. 41, pp. 451–474, 2003.

[10] S. Borst and P. Whiting, “Dynamic rate control algorithms for HDR throughput optimization,” in *Proc. IEEE INFOCOM*, 2001, vol. 2, pp. 976–985.

[11] Y. Liu and E. Knightly, “Opportunistic fair scheduling over multiple wireless channels,” in *Proc. IEEE INFOCOM*, 2003, vol. 2, pp. 1106–1115.

[12] A. Bhorkar, A. Karandikar, and V. Borka, “Power optimal opportunistic scheduling,” in *Proc. IEEE GLOBECOM*, 2006, pp. WLC39-3:1–WLC39-3:5.

[13] M. Xiao, N. B. Shroff, and E. Chong, “Utility-based power control in cellular wireless systems,” in *Proc. IEEE INFOCOM*, 2001, vol. 1, pp. 412–421.

[14] R. Urgaonkar and M. J. Neely, “Opportunistic scheduling with reliability guarantees in cognitive radio networks,” *IEEE Trans. Mobile Comput.*, vol. 8, no. 6, pp. 766–777, Jun. 2009.

[15] M. Lotfinezhad, B. Liang, and E. S. Sousa, “Optimal control of constrained cognitive radio networks with dynamic population size,” in *Proc. IEEE INFOCOM*, Mar. 2010, pp. 1–9.

[16] O. Simeone, I. Stanojev, S. Savazzi, Y. Bar-Ness, U. Spagnolini, and R. Pickholtz, “Spectrum leasing to cooperating secondary ad hoc networks,” *IEEE J. Sel. Areas Commun.*, vol. 26, no. 1, pp. 203–213, Jan. 2008.

[17] S. K. Jayaweera and T. Li, “Dynamic spectrum leasing in cognitive radio networks via primary-secondary user power control games,” *IEEE Trans. Wireless Commun.*, vol. 8, no. 6, pp. 3300–3310, Jun. 2009.

[18] S. K. Jayaweera, G. Vazquez-Vilar, and C. Mosquera, “Dynamic spectrum leasing: A new paradigm for spectrum sharing in cognitive radio networks,” *IEEE Trans. Veh. Technol.*, vol. 59, no. 5, pp. 2328–2339, Jun. 2010.

[19] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, “Cooperative diversity in wireless networks: Efficient protocols and outage behavior,” *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[20] M. J. Neely, “Energy optimal control for time varying wireless networks,” *IEEE Trans. Inf. Theory*, vol. 52, no. 7, pp. 2915–2934, Jul. 2006.

[21] L. Georgiadis, M. J. Neely, and L. Tassiulas, *Resource Allocation and Cross-Layer Control in Wireless Networks*. Hanover, MA: NOW, 2006, vol. 1, Foundations and Trends in Networking.

[22] S. Boyd and L. Vandenbeghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[23] P. Liu, P. Zhang, S. Jordan, and M. L. Honig, “Single-cell forward link power allocation using pricing in wireless networks,” *IEEE Trans. Wireless Commun.*, vol. 3, no. 2, pp. 533–543, Mar. 2004.

[24] J. Schiller, *Mobile Communication*. Reading, MA: Addison-Wesley, 2000.

[25] M. J. Neely, “Super-fast delay trade-offs for utility optimal fair scheduling in wireless networks,” *IEEE J. Sel. Areas Commun.*, vol. 24, no. 8, pp. 1489–1501, Aug. 2006.

Mehmet Karaca (S’08) received the B.S. degree in telecommunication engineering from Istanbul Technical University, Istanbul, Turkey, in 2006, and the M.S. degree in electronics engineering from Sabanci University, Istanbul, Turkey, in 2008, is currently pursuing the Ph.D. degree in electronics engineering at Sabanci University.

His research interests include the design and analysis of scheduling and resource allocation algorithms for wireless networks, stochastic optimization, cognitive radio networks, and machine learning.

Karim Khalil (S’10) received the B.S. degree in electronics and communications engineering from Cairo University, Cairo, Egypt, in 2007, and the M.S. degree in wireless communications from Nile University, Cairo, Egypt, in 2009, and is currently pursuing the Ph.D. degree in electrical and computer engineering at The Ohio State University, Columbus.

From 2007 to 2009, he was a Research Assistant with the Wireless Intelligent Networks Center (WINC), Nile University. His research interests are wireless communications, cognitive radio networks, information-theoretic security, and game theory.

Mr. Khalil is a student member of the IEEE Communications Society.

Eylem Ekici (S‘99–M’02–SM’11) received the B.S. and M.S. degrees in computer engineering from Bogazici University, Istanbul, Turkey, in 1997 and 1998, respectively, and the Ph.D. degree in electrical and computer engineering from the Georgia Institute of Technology, Atlanta, in 2002.

Currently, he is an Associate Professor with the Department of Electrical and Computer Engineering, The Ohio State University, Columbus. His current research interests include nanoscale communication systems, vehicular communication systems, cognitive radio networks, resource management, and analysis of network architectures and protocols.

Dr. Ekici is an Associate Editor of *Computer Networks*, *Mobile Computing and Communications Review*, and the IEEE/ACM TRANSACTIONS ON NETWORKING.

Ozgur Ercetin (M’12) received the B.S. degree in electrical and electronics engineering from the Middle East Technical University, Ankara, Turkey, in 1995, and the M.S. and Ph.D. degrees in electrical engineering from the University of Maryland, College Park, in 1998 and 2002, respectively.

Since 2002, he has been with the Faculty of Engineering and Natural Sciences, Sabanci University, Istanbul, Turkey. He was also a Visiting Researcher with HRL Labs, Malibu, CA; Docomo USA Labs, Palo Alto, CA; and The Ohio State University, Columbus. His research interests are in the field of computer and communication networks with emphasis on fundamental mathematical models, architectures and protocols of wireless systems, and stochastic optimization.