This paper was presented as part of the main technical program at IEEE INFOCOM 2011

Optimal Scheduling in Cooperate-to-Join Cognitive Radio Networks

Karim Khalil*, Mehmet Karaca†, Ozgur Ercetin†and Eylem Ekici*
*Department of Electrical and Computer Engineering
The Ohio State University, Columbus, OH 43210 USA
Email: {khalilk,ekici}@ece.osu.edu.
†Faculty of Engineering and Natural Sciences
Sabanci University, 34956 Orhanli-Tuzla, Istanbul/Turkey
Email: mehmetkrc@su.sabanciuniv.edu, oercetin@sabanciuniv.edu

***Abstract*—Optimal transmission scheduling in wireless cognitive networks is considered under the spectrum leasing model. We propose a cooperative scheme in which secondary nodes share the time slot with primary nodes in return for cooperation. Cooperation is feasible only if the system's performance is improved over the non-cooperative case. First, we investigate a scenario where secondary users are interested in immediate rewards. 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. In both scenarios, our proposed schemes are shown to outperform non-cooperative scheduling schemes, in terms of both individual and total expected utility, 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.**

I. Introduction

Cognitive Radio Networks (CRN) have recently been investigated extensively [1], [2]. The main advantage that CRN present is the efficient utilization of the scarce radio spectrum resources. By opportunistically exploiting the underutilized spectrum, unlicensed (i.e., secondary) users can transmit over the licensed bands, given that they do not *hurt* the licensed (i.e., primary) users' performance.

Approaches to cognitive radio can be divided into two categories: *commons* model and *property-rights* model [3]. In the commons model, the primary network is oblivious to the secondary network activity and the aim of secondary nodes 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 nodes own the spectrum and are willing to lease it to secondary nodes 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 achievable rates. At the same time, a cognitive node half way between the licensed user and the base station has more favorable channel conditions. The cognitive user wishes to access the channel and communicate with the base station. After coordination, the primary user agrees to share a portion of its own time slot with the secondary user in exchange for secondary user relaying the primary user's data to the base station. In our work, we exploit this cooperative scenario between primary and secondary nodes to improve the overall system performance.

This work is supported in part by NSF under grant number CCF-0914912.

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. It has been shown that policies that exploit the time varying nature of the wireless channel to schedule users are at least as good as static policies [4]. 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, this type of scheduling results in unfair sharing of resources and may lead to starvation of some users, for example, those far away from the base station in a cellular network. To mitigate this, fairness constraints are added to the problem formulation.

Opportunistic scheduling was recently studied for cognitive radio networks under the commons model [5], [6]. 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 of virtual queues, the joint problem of stabilizing the queues of secondary nodes 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 this paper, we propose optimal opportunistic scheduling policies for primary *and* secondary nodes in a cognitive radio network in two cases. First, we consider the optimization of the total expected utility while satisfying an average performance constraint for each primary node in the network. Here, we develop a cooperative scheduling policy by which the performance is improved and shown to be *at least* the same as the original primary-only system. In this cooperative scenario, during a time slot, nodes cooperate using decodeand-forward multihop scheme [7] where secondary nodes relay the messages of primary nodes 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 the secondary node relays the primary user's data and the fraction used to transmit the secondary user's own data. The problem is formulated in an optimization framework where the objective is to maximize the average system performance subject to long term constraints on the primary nodes' performance.

In the second part of the paper, another formulation is considered in which the secondary nodes 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. We employ Lyapunov optimization tools developed in [8], [9] 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 the paper is organized as follows. Section II presents the network model, the basic structure of the proposed cooperative scheme and an introduction to the Lyapunov analysis. In Section III, we introduce the formulation of the scheduling problem with constraints on the minimum achievable performance of the primary system. The structure of the optimal policy is then derived and shown to hold for both stationary and time varying policies. Then, in Section IV, we formulate and solve another version of the problem where constraints on the minimum performance of the secondary nodes are added to the original set of constraints. Numerical results are presented in Section V. Finally, Section VI concludes the paper and presents possible future directions.

II. NETWORK MODEL

A. The Cognitive Network

Consider a cognitive radio network of M primary users and N secondary users, all wishing to communicate with a common destination as shown in Figure 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 nodes. We adopt a non-interference model where only one node, either primary or secondary, is transmitting at any given time[1]. Random channel gains between each node and other nodes 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.

[1]The same analysis and results in the paper can be applied for the downlink case where a primary user is the destination and the base station is the primary transmitter. Conditions (3), (4) and (10) changes accordingly. We consider this case in the journal version.

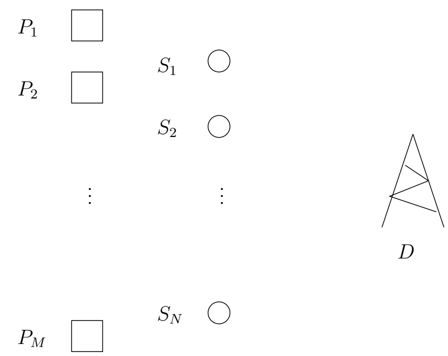

Fig. 1. Network Model.

In the following analysis, we use the notation $R_m^p(t)$, $R_n^s(t)$ to denote the *achievable* transmission rates from primary node m to destination and from secondary node n to destination, respectively, at time slot t. The corresponding random rate vectors are denoted as $\mathbf{R}^p(t)$, $\mathbf{R}^s(t)$. The achievable rate from primary node m to secondary node n is denoted as $R_{mn}^r(t)$, where the corresponding rate matrix is $\mathbf{R}^r(t)$. 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 the next subsection, since our scheme works by selecting a pair of nodes (primary and secondary) to transmit at a given time slot, the utility achieved by a node is a function of the cooperating pair. Consequently, the utility function of a primary user m when it cooperates with secondary user n at time slot t is denoted as $U_{mn}(t)$. Similarly, the utility function of a secondary user n that cooperates with primary user m is denoted as $V_{mn}(t)$. These utility functions are measures of the level of satisfaction of users and thus they are generally non-decreasing concave functions of the transmission rate.

Throughout the paper, we assume that both primary and secondary users adopt fixed transmission power and adaptive transmission rate strategies so that every node transmits at its achievable rate in each time slot. We further assume all users to be backlogged so that they always have packets to transmit to the destination.

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 primary user 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 $\mathbf{R}^p(t)$, $\mathbf{R}^s(t)$, rate matrix $\mathbf{R}^r(t)$ 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 nodes 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 $(1-\alpha)$, $0 \leq \alpha \leq 1$, of the time slot, the primary node m sends its data (intended to destination) to secondary (relay) node n. In the remaining portion of the time slot, the scheduled secondary node uses the channel to relay the primary user's data over a β fraction, $0 \leq \beta \leq 1$, and then, transmits its own data during the rest of the time slot, i.e., over $\alpha(1-\beta)$ fraction. This cooperative scheme is a form of implementation of the spectrum leasing cognitive radio framework where secondary users 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 [10], however, we do not employ distributed space time coding and allow only one secondary node to cooperate in a given time slot.

We set $n = 0$ by definition for the case when a primary node m is scheduled to transmit directly to the destination without cooperating with secondary nodes. This case arises when cooperation is either infeasible or leads to suboptimal utility values. We set $R^r_{m0}(t) = R^p_m(t)$, $m \in \{1, 2, \cdots, M\}$ and $\alpha = 0$ in such cases.

The utility function is taken to be a non-decreasing concave function of the transmission 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

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

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

in a given time slot t. For all other primary and secondary nodes (m, n) such that $m \neq m'$ and $n \neq n'$, 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. An example of utility functions that can be used is $h(x) = \log(1 + x)$.

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

$$0 < (1-\alpha)R^r_{mn}(t) \leq \alpha R^s_n(t). \quad (3)$$

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

$$\beta^* = \frac{(1-\alpha)R^r_{mn}}{\alpha R^s_n} \quad (4)$$

where β^* is the optimal value of β. If $\beta < \frac{(1-\alpha)R^r_{mn}}{\alpha R^s_n}$, secondary node n does not have sufficient time to relay the data of m. If $\beta > \frac{(1-\alpha)R^r_{mn}}{\alpha R^s_n}$, then unnecessary time is wasted by node 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 $\beta \leq 1$.

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 $0 \leq \alpha \leq 1$. 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,

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

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 achievable 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 deal with performance optimization and queue stability problems simultaneously in a unified framework. Moreover, the concept of virtual queues simplifies the analysis of the system when long term performance constraints are imposed [8]. In fact, the problem is transformed into a network stability problem.

We first introduce two definitions: Let $Z_i(t)$, $i \in \{1, 2, \cdots, L\}$ be a queue backlog process and $\mathbf{Z}(t) = (Z_1(t) Z_2(t) \cdots Z_L(t))$ in a network with L nodes. 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

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

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

We restate a result of [9] that is critical to establish the optimality of our proposed scheme.

Theorem 1: (Lyapunov Optimization) [9] 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 the Theorem in [9], since we are interested in the average of the utility function. 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), (2) and the scheduling policy Q.

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

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

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 $\mathbf{R}^p(t), \mathbf{R}^s(t), \mathbf{R}^r(t)$. Then we investigate the more general time-varying policies.

A. Problem Formulation

The optimal opportunistic scheduling problem with minimum performance constraints was previously solved in [4]. By including N secondary nodes to the system, our model can be viewed as a generalization to the model in [4]. In addition, setting $N = 0$ in our scheme yields the scheme in [4] as will be shown in the next subsection.

The problem can be stated formally as follows:

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

$m \in \{1,2,\ldots,M\}$, where C_m is the minimum performance constraint for each primary user 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 the secondary users 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 subsection, we propose a stationary scheduling policy in a form similar to the optimal policies reported in [4], and show that it solves (11) for the given cognitive radio network.

Scheduling Algorithm Q_{1a}:

For every time slot t and the given values of $U_{mn}(t)$ and $V_{mn}(t)$ for all m,n, the solution to the scheduling problem (11) is given by

$$Q_{1a} = \underset{(m,n,\alpha)\in\mathcal{F}_1}{\operatorname{argmax}} \left\{\lambda_m^* U_{mn}(Q,t) + V_{mn}(Q,t)\right\}, \tag{12}$$

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

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

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

Proof: The proof is provided in Appendix A. ■

The structure of the derived scheduling policy suggests that when a primary user m experiences unfavorable channel conditions, the associated parameter λ_m^* will be larger than one. Then, it attains average utility that is only equal to its corresponding constraint. Otherwise, this primary user 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 $R_m^p(t)$, $R_n^s(t)$ and $R_{mn}^r(t)$ for all m and n, the scheduler is able to construct the set of feasible policies $\mathcal{F}_1$ by associating the 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 in (10), where $0 \leq r_\alpha^-, r_\alpha^+, r_\beta^-, r_\beta^+ \leq 1$. Then it decides which pair (or single primary user) are relatively best according to (12). The choice of the pair (m,n)

is a combinatorial optimization problem which may require discrete exhaustive search. The optimal value of α can be obtained easily since (12) can be shown to be concave in α.

The parameters $\lambda_m^*, m \in \{1, 2, \ldots, M\}$ 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 [4]. An estimation technique is presented in Section V.

Example: The above algorithm can be compared to non-cooperative algorithms as follows. Consider for example the utilitarian fairness constraints problem solved in [4] with the constraints $a_m = \gamma_m \bar{W}(\hat{Q})$ for each primary user $m \in \{1, 2, \ldots, M\}$ 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 $\Gamma,\ 1 \leq \Gamma \leq \Gamma_{max}$ be the improvement factor with respect to the system with no cooperation where Γ_{max} is to be specified by the boundary of the feasibility region (see Section III-D). Now consider a network of M primary and N secondary nodes such that the scheduler executes the optimal policy $\hat{Q}$ to schedule only the primary nodes but does not act on it and simultaneously executes and implements our cooperative scheduling policy Q_{1a}. Since the scheduling policy $\hat{Q}$ converges as the number of time slots $t \to \infty$ [4], 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 primary nodes over the non-cooperative scheme, and hence improves the overall performance.

C. Optimal Time Varying Policy

In this subsection, we solve problem (11) using the stochastic network optimization tool of [9]. This tool yields a scheduling policy that is similar in structure to (12). However, the policy derived in this subsection does not need the computation of the online parameters λ_m^*.

Define the time average expected utility as follows.

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

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. Let $U_m^t(Q, t) \leq U^{max}$, $V_n^t(Q, t) \leq V^{max}$ 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:

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

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

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

$$\Delta_1(\mathbf{X}(t)) - K\mathbb{E}[W(Q, t) | \mathbf{X}(t)] \leq B_1 + 2 \sum_{m=1}^{M} X_m(t) C_m - \sum_{m=1}^{M} (K + 2X_m(t)) \bar{U}_m^t(Q, t) - \sum_{n=1}^{N} K \bar{V}_n^t(Q, t), \tag{18}$$

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

Proof: The proof is given in Appendix B. ■

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

Scheduling Algorithm Q_{1b}:

At each time slot t, observe the virtual queue backlog $X_m(t)$ for each primary user m and the achievable transmission rates, and choose (m, n, α) solving the following optimization problem.

$$Q_{1b} = \underset{(m,n,\alpha) \in \mathcal{F}_1}{\operatorname{argmax}} \left\{ \left(1 + \frac{2X_m(t)}{K}\right) U_{mn}(t) + V_{mn}(t) \right\}$$

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_{1a}, let $\tilde{\lambda}_m(t) \triangleq 1 + \frac{2X_m(t)}{K} \geq 1$. It is clear that both algorithms have exactly the same form. However, contrary to the algorithm in Section III-B, Q_{1b} does not require the knowledge of the statistics of the channel states or need the computation of online parameters.

We analyze our algorithm using the Lyapunov drift with optimization [9]. We define a class of policies that will be useful to prove the optimality of the scheduling algorithm Q_{1b}. Consider the class of scheduling algorithms $\mathcal{S}$ that schedules users according to a stationary and possibly randomized function of only the achievable rates and independent of the queue states. It was shown in [8], [9] that the 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 $\{(m, n, \alpha) \mid m \in \{1, 2, \cdots, M\}, n \in \{1, 2, \cdots, N\}, \alpha \in [0, 1]\}$ is closed and bounded, we have the following lemma (which can be proved using similar arguments as in [8]). Let the feasibility region of (11) be Λ and let $\boldsymbol{\epsilon} \triangleq (\epsilon\ \epsilon \cdots\ \epsilon)$.

Lemma 4: 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:

$$\mathbb{E}[W(Q_{s_1}, t)] = \bar{W}_1^*, \tag{19}$$

$$\mathbb{E}\left[U_m^t(Q_{s_1}, t)\right] \geq C_m, m \in \{1, 2, \ldots, M\}, \tag{20}$$

where $\bar{W}_1^*$ is the optimal performance for the problem (11) over all scheduling policies. Moreover, if $\mathbf{C}$ is strictly interior to Λ, then there exists $\epsilon > 0$ with $(\mathbf{C}+\boldsymbol{\epsilon}) \in \Lambda$ and a stationary scheduling policy $Q_{s_1(\epsilon)}$ satisfying:

$$\mathbb{E}\left[U_m^t(Q_{s_1(\epsilon)}, t)\right] \geq C_m + \epsilon, m \in \{1, 2, \ldots, M\} \tag{21}$$

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

We are now ready to present bounds on the performance of our proposed algorithm Q_{1b}. The following Theorem shows that all the virtual queues are strongly stable [9]. Hence, all time average constraints in (11) are satisfied.

Theorem 5: If $\mathbf{C}$ is strictly in the interior of Λ, then the proposed algorithm in Q_{1b} stabilizes the virtual queues and achieves the following bounds:

$$\liminf_{t\to\infty} \frac{1}{t} \sum_{\tau=0}^{t-1} \mathbb{E}[W(Q_{1b}, \tau)] \geq \bar{W}_1^* - \frac{B_1}{K}, \tag{22}$$

$$\limsup_{t\to\infty} \frac{1}{t} \sum_{\tau=0}^{t-1} \mathbb{E}\left[\sum_{m=1}^{M} X_m^2(t)\right] \leq \frac{B_1 + KW^{max}}{\epsilon_{max}}, \tag{23}$$

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

Proof: Consider the upper bound given by Lemma 3. From Lemma 4, there exists a stationary policy $Q_{s_1(\epsilon)}$ that satisfies the constraints (21). By the definition of Q_{1b}, $RHS_{Q_{1b}} \leq RHS_{Q_{s_1(\epsilon)}}$ where RHS_Q is the right hand side (RHS) of inequality (18) evaluated for the policy Q. Now consider evaluating $RHS_{Q_{s_1(\epsilon)}}$ using (21). Expanding the RHS of (18) and using the property that the utility is independent of queue states, it is straightforward to see that $RHS_{Q_{s_1(\epsilon)}} = B_1 - 2\epsilon \sum_{m=1}^{M} X_m(t) - K\bar{W}_1^*(\epsilon)$. It follows that

$$\Delta_1(\mathbf{X}(t)) - K\mathbb{E}[W(Q_{1b}, t)|\mathbf{X}(t), \mathbf{Y}(t)] \leq RHS_{Q_{1b}}$$
$$\leq RHS_{Q_{s_1(\epsilon)}} = B_1 - 2\epsilon \sum_{m=1}^{M} X_m(t) - K\bar{W}_1^*(\epsilon),$$

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

$$\liminf_{t\to\infty} \frac{1}{t} \sum_{\tau=0}^{t-1} \mathbb{E}[W(Q_{1b}, \tau)] \geq \bar{W}_1^*(\epsilon) - \frac{B_1}{K}, \tag{24}$$

$$\limsup_{t\to\infty} \frac{1}{t} \sum_{\tau=0}^{t-1} \mathbb{E}\left[\sum_{m=1}^{M} X_m^2(t)\right] \leq \frac{B_1 + KW^{max}}{2\epsilon}, \tag{25}$$

where (25) 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_{1b}. Therefore, (24) and (25) can be optimized separately. Taking $\epsilon \to 0$ in (24) yields (22) and taking $\epsilon = \frac{\epsilon_{max}}{2}$ in (25) yields (23), concluding the proof. ■

D. A 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 statistics of the channel conditions. Since our scheduling schemes can only improve the performance of the primary-only network, the feasibility region given in [4] is strictly a subset of the feasibility region of our policy. In addition, it can be shown, using similar techniques as in [4], 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 primary user in addition to N secondary users.

Considering the example presented in Section III-B, Γ_{max} specified the maximum gain the cooperative system can achieve over the non-cooperative counterpart. It is clear that Γ_{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 Subsection III-B, and if the feasibility region of our cooperative system is Λ, then Γ_{max} is given by:

$$\Gamma_{max} = \max_{\Gamma \geq 1}\ \Gamma$$
$$\text{s.t. } (\Gamma a_1\ \Gamma a_2 \cdots \Gamma a_M) \in \Lambda. \tag{26}$$

The solution to (26) can be determined numerically if the channel statistics are known. A more rigorous characterization of the feasibility region is beyond the scope of this work and is part of our future work.

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 secondary user. More specifically, a portion of the primary utility achieved by cooperation is guaranteed for each cooperating secondary node, in an average sense. In fact, the formulation of the problem below allows for the idea of *banking* between primary and secondary nodes. That is, in contrast to the immediate rewards of Section III, here, secondary nodes are guaranteed a specific share over a large number of time slots. This is achieved by allowing α to take values such that $\alpha \leq 1$, i.e., we lift the constraint imposed in the first inequality of (10). The problem is formulated as follows:

$$\begin{aligned} &\max_{Q\in\mathcal{F}} \ \bar{W}(Q) \\ \text{s.t.:}1)\ &\mathbb{E}\left[U_m^t(Q,t)\right] \geq C_m,\ m\in\{1,2,\ldots,M\}, \\ 2)\ &\mathbb{E}\left[V_n^t(Q,t)\right] \geq \mathbb{E}\left[\sum_{m=1}^{M}\phi(U_{mn}(Q,t))\right], \\ &n\in\{1,2,\ldots,N\}, \end{aligned} \tag{27}$$

where $\phi(\cdot)$ is a non-negative, non-decreasing scalar-valued function. We assume that the constraints in (27) 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 a virtual queue such that the queue dynamics are given by

$$X_m(t+1) = [X_m(t) - U_m^t(Q,t)]^+ + C_m, \tag{28}$$
$$Y_n(t+1) = [Y_n(t) - V_n^t(Q,t)]^+ + \nu_n(Q,t), \tag{29}$$

$m \in \{1,2,\ldots,M\}$, $n \in \{1,2,\ldots,N\}$. We assume that $\nu_n(Q,t) \leq \nu^{max}$ for all $n \in \{1,2,\ldots,N\}$, $t \geq 0$ 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 secondary nodes. Define the following quadratic Lyapunov function and conditional Lyapunov drift

$$L_2(\mathbf{X}(t),\mathbf{Y}(t)) \triangleq \sum_{m=1}^{M} X_m^2(t) + \sum_{n=1}^{N} Y_n^2(t), \tag{30}$$

$$\begin{aligned} &\Delta_2(\mathbf{X}(t),\mathbf{Y}(t)) \triangleq \\ &\mathbb{E}\left[L_2(\mathbf{X}(t+1),\mathbf{Y}(t+1)) - L_2(\mathbf{X}(t),\mathbf{Y}(t))|\mathbf{X}(t),\mathbf{Y}(t)\right]. \end{aligned} \tag{31}$$

We also define the following conditional expectation.

$$\bar{\nu}_n(Q,t) \triangleq \mathbb{E}[\nu_n(Q,t)|\mathbf{X}(t),\mathbf{Y}(t)]. \tag{32}$$

The Lyapunov drift in (31) is bounded by the following Lemma where the proof is very similar to the proof of Lemma 3 and is omitted for brevity.

Lemma 6: For every time slot t, the Lyapunov drift defined in (31) can be upper bounded as follows.

$$\begin{aligned} &\Delta_2(\mathbf{X}(t),\mathbf{Y}(t)) - K\mathbb{E}[W(Q,t)|\mathbf{X}(t),\mathbf{Y}(t)] \\ &\quad \leq B_2 + 2\sum_{m=1}^{M} X_m(t)C_m - \sum_{m=1}^{M}(K+2X_m(t))\bar{U}_m^t(Q,t) \\ &\qquad - \sum_{n=1}^{N}(K+2Y_n(t))\bar{V}_n^t(Q,t) + \sum_{n=1}^{N} 2Y_n(t)\bar{\nu}_n(Q,t), \end{aligned}$$

where $B_2 = \sum_{m=1}^{M} C_m^2 + M(U^{max})^2 + N((\nu^{max})^2 + (V^{max})^2)$ 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 state of the virtual queues $X_m(t)$, $Y_n(t)$ and the achievable rates $R_m^p(t)$, $R_{mn}^r(t)$ and $R_n^s(t)$ for all $m \in \{1,2,\ldots,M\}$ and $n \in \{1,2,\ldots,N\}$, and then solves the following optimization problem:

$$\begin{aligned} Q_2 = \underset{(m,n,\alpha)\in\mathcal{F}}{\operatorname{argmax}} &\left\{\left(1+\frac{2X_m(t)}{K}\right)U_{mn}(t)\right. \\ &\left. + \left(1+\frac{2Y_n(t)}{K}\right)V_{mn}(t) - \left(\frac{2Y_n(t)}{K}\right)\phi(U_{mn}(t))\right\} \end{aligned}$$

The virtual queues are then updated according to the queue dynamics in (28), (29).

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 secondary user 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 $\phi(U_{mn}(t))$). Therefore, it is possible that the system allocates an entire time slot to a secondary user without requiring the relay of a primary user's data. Similarly, it is also possible that a secondary user relays primary data without obtaining immediate share of that time slot to transmit its own data.

B. Algorithm Analysis

The analysis follows the same strategy as in Section III-C. Let $\bar{W}_2^*$ be the optimal time average system utility achieved over all scheduling policies for the problem (27), and consider the class of stationary randomized scheduling algorithms that are independent of the queue states.

Lemma 7: If vectors $\mathbf{C}$ and $\mathbb{E}[\boldsymbol{\nu}] = \mathbb{E}[(\nu_1\ \nu_2\cdots\ \nu_N)]$ are feasible, then there exists a stationary randomized policy Q_{s_2} that solves (27) and satisfies the following:

$$\mathbb{E}[W(Q_{s_2},t)] = \bar{W}_2^*, \tag{33}$$
$$\mathbb{E}\left[U_m^t(Q_{s_2},t)\right] \geq C_m,\ m\in\{1,2,\ldots,M\}, \tag{34}$$
$$\mathbb{E}\left[V_n^t(Q_{s_2},t)\right] \geq \mathbb{E}\left[\nu_n(Q_{s_2},t)\right], \tag{35}$$

where $\bar{W}_2^*$ is the optimal performance for the problem (27) over all scheduling policies. Moreover, if $\mathbf{C}$ and $\mathbb{E}[\boldsymbol{\nu}]$ are strictly interior to the feasibility region, then there exists $\epsilon' > 0$ and a stationary scheduling policy $Q_{s_2(\epsilon')}$ satisfying:

$$\mathbb{E}\left[U_m^t(Q_{s_2(\epsilon')},t)\right] \geq C_m + \epsilon',\ m\in\{1,2,\ldots,M\}, \tag{36}$$
$$\mathbb{E}\left[V_n^t(Q_{s_2(\epsilon')},t)\right] \geq \mathbb{E}\left[\nu_n(Q_{s_2},t)\right] + \epsilon', \tag{37}$$

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

Theorem 8: If the constraints in (27) are feasible, then the proposed algorithm Q_2 stabilizes the virtual queues and achieves the following bounds.

$$\liminf_{t\to\infty}\frac{1}{t}\sum_{\tau=0}^{t-1}\mathbb{E}[W(Q_2,\tau)] \geq \bar{W}_2^* - \frac{B_2}{K},$$
$$\limsup_{t\to\infty}\frac{1}{t}\sum_{\tau=0}^{t-1}\mathbb{E}\left[\sum_{m=1}^{M}X_m^2(t) + \sum_{n=1}^{N}Y_n^2(t)\right] \leq \frac{B_2 + KW^{max}}{\epsilon'_{max}},$$

where $\bar{W}_2^*$ is the optimal value for the time average expected utility, $B_2 = \sum_{m=1}^{M} C_m^2 + M(U^{max})^2 + N((\nu^{max})^2 +$

$(V^{max})^2)$, $W^{max} = U^{max} + V^{max}$ and ϵ'_{max} is the largest ϵ' such that constraints (36) and (37) are feasible.

Proof: The proof uses Lemma 6 and Lemma 7 and is similar to the proof of Theorem 5. ■

V. Numerical Results

In this section, we simulate a wireless network with $M = 4$ primary nodes and a varying number of secondary nodes all communicating with a common destination. First, 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 and evolve independently across users and across time. For all pairs (m, n), the transmission rates are set to $R^p_m(t) = \{100,\ 15\}$ units/slot with probability $\{0.5,\ 0.5\}$, $R^r_{mn}(t) = \{100,\ 15\}$ units/slot with probability $\{0.6,\ 0.4\}$, and $R^s_n(t) = \{100,\ 15\}$ units/slot 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_{1a} 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 primary users performance 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 primary users. We set the constraints C_m in (11) as in the example given is Section III-B for the sake of comparison with non-cooperative systems. Applying scheduling policy Q_{1a}, we let $\Gamma = 1.01$ and use a stochastic approximation approach to estimate the parameters $\lambda^*_m, m \in \{1, 2, \dots, M\}$ as follows. First, from the constraints on the primary user performance, we see that for $m \in \{1, 2, \cdots, M\}$, λ^*_m is the root to the following equation.

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

But 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^k_m of f^k_m as:

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

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

$$\lambda^{k+1}_m = \lambda^k_m - \delta^k g^k_m$$

where δ^k can be taken to be $1/k$ [4].

For a given time slot t, it can be shown that (12) is a concave function in α. The optimization over α is then done over all pairs so that α satisfies condition (10) yielding α^*_{mn} for every pair (m, n). Then the tuple $(m, n, \alpha^*_{mn}) \in \mathcal{F}_1$ that maximizes (12) is selected by the scheduler at this time slot. For each pair (m, n), since the objective function (12) is concave in α and the constraint is linear, then Karush-Kuhn-Tucker conditions are both necessary and sufficient to solve the problem (12), along with (10) [11].

In Figure 2, the average system utility is plotted with respect to the number of cognitive nodes. As shown in the figure, 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 policy Q_{1a} still performs better than the non-cooperative policy. For $N > 1$, the constraints are feasible. Moreover, exploiting the opportunity relaying offers, we could achieve non zero secondary system average utility. Figure 2 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.

Next, we apply scheduling policy Q_2 when $M = 4$ and the number of secondary users is fixed at $N = 5$ for the same channel statistics used in the first simulation. In Figure 3, the running average of the expected utility of secondary node 3 up to time t is plotted and compared to the average primary utility achieved through cooperation with secondary node 3 for $\phi(U_{mn}(Q, t)) = bU_{mn}(Q, t)$, and $b = 1.5$. In this experiment, we set $C_m = 3.5\ \forall m$. Stability is achieved for all primary and secondary virtual queues and hence the constraints for the primary and secondary users are met.

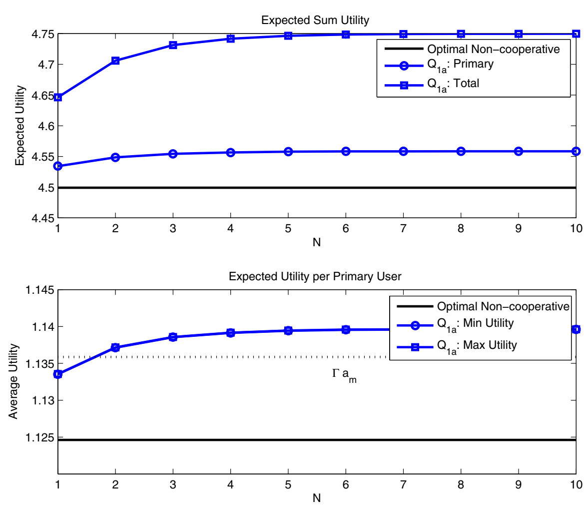

Fig. 2. Average total system utility.

VI. Conclusion

We have studied optimal transmission scheduling policies in a cognitive radio network under a cooperation based spectrum leasing model for the heavy traffic scenario. We show that our cooperative scheme improves the basic primary network performance and allows secondary nodes to access the licensed spectrum in return for cooperation. Immediate and long term rewards for secondary users are studied where we introduce the idea of banking between primary and secondary users. Possible extensions include considering the joint control and scheduling problem and the effect of finite buffers.

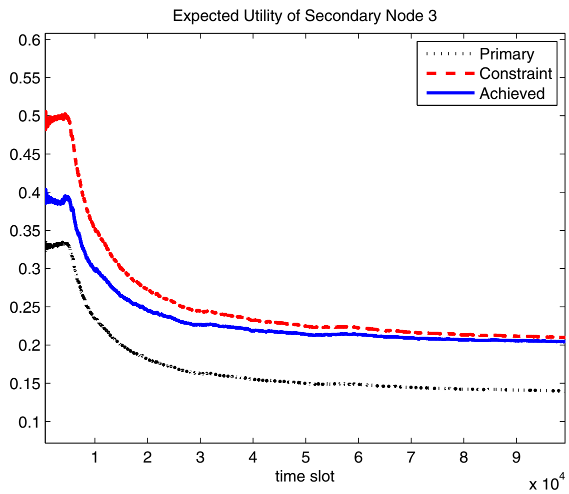

Fig. 3. Performance of secondary node 3.

Appendix A
Proof of Theorem 2

The proof is similar to the proof of the optimal policies in [4]. However, the scheduling policy Q in our work decides a tuple of three variables each time slot instead of only choosing a primary user.

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, it follows that

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

$$= \sum_{m=1}^{M}\sum_{n=0}^{N}\mathbb{E}[U_{mn}(Q) + V_{mn}(Q)]$$

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

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

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

$$\lambda_m^* U_{mn}(Q) + V_{mn}(Q) \leq \lambda_m^* U_{mn}(Q_{1a}) + V_{mn}(Q_{1a})$$

Therefore, we can write

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

$$- \sum_{m=1}^{M}(\lambda_m^* - 1)C_m = \bar{W}(Q_{1a})$$

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

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

Appendix B
Proof of Lemma 3

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

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

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

$$\Delta_1(\mathbf{X}(t)) \leq B_1 + 2\sum_{m=1}^{M}X_m(t)(C_m - \bar{U}_m^t) \quad (38)$$

where $B_1 = \sum_{m=1}^{M} C_m^2 + M(U^{max})^2$. Subtracting the term $K\mathbb{E}[W(Q,t)|\mathbf{X}(t)]$ from both sides of (38), expanding and rearranging terms, and using $V_0^t(t) = 0 \; \forall \; t$, (18) follows.

References

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

[2] 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.

[3] J. O. Neel, "Analysis and design of cognitive radio networks and distributed radio resource management algorithms," Ph.D. dissertation, Virginia Polytechnic Institute, Sep. 2006.

[4] X. Liu, E. K. P. Chong, and N. B. Shroff, "A framework for opportunistic scheduling in wireless networks," *Comput. Netw.*, vol. 41, no. 4, pp. 451–474, Mar. 2003.

[5] 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.

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

[7] 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.

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

[9] L. Georgiadis, M. J. Neely, and L. Tassiulas, *Resource allocation and cross-layer control in wireless networks*. Foundations and Trends in Networking, 2006.

[10] 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.

[11] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.