Energy-Efficient Nonstationary Spectrum Sharing

Yuanzhang Xiao and Mihaela van der Schaar, *Fellow, IEEE*

***Abstract*—We develop a novel design framework for energy-efficient spectrum sharing among autonomous users who aim to minimize their energy consumptions subject to minimum throughput requirements. Most existing works proposed *stationary* spectrum sharing policies, in which users transmit at *fixed* power levels. Since users transmit simultaneously under stationary policies, to fulfill minimum throughput requirements, they need to transmit at high power levels to overcome interference. To improve energy efficiency, we construct *nonstationary* spectrum sharing policies, in which the users transmit at *time-varying* power levels. Specifically, we focus on TDMA (time-division multiple access) policies in which one user transmits at each time (but not in a round-robin fashion). The proposed policy can be implemented by each user running a low-complexity algorithm in a decentralized manner. It achieves high energy efficiency even when the users have *erroneous* and *binary* feedback about their interference levels. Moreover, it can adapt to *dynamic entry and exit* of users. The proposed policy is also deviation-proof, namely autonomous users will find it in their self-interests to follow it. Compared to existing policies, the proposed policy can achieve an energy saving of up to 90% under a large number of users.**

***Index Terms*—Repeated games, tntervention, power control, flow control.**

I. Introduction

A KEY challenge in wireless networks is determining efficient solutions for autonomous users to share the spectrum. In cognitive radio networks where the users are differentiated as primary users (PUs) and secondary users (SUs), we also require SUs to access the spectrum without degrading PUs' quality of service (QoS). To be more general, we consider cognitive radio networks in this work, and design spectrum sharing policies that achieve efficient spectrum usage and protect PUs' QoS. Our work can be easily applied to wireless networks in which users are not differentiated as PUs and SUs.

Spectrum sharing policies, which specify the PUs' and SUs' transmission schedules and transmit power levels, are essential to achieve spectrum and energy efficiency. Research on designing spectrum sharing policies can be roughly divided in two main categories. The research in the first category formulates the spectrum sharing problem as a utility maximization problem subject to the users' maximum transmit power constraints [1]–[10][19]–[22][29]. Many works in this category [1]–[7][19]–[22][29] define the utility function as an increasing function of the signal-to-interference-and-noise-ratio (SINR), while neglecting to consider the energy consumption of the resulting spectrum sharing policies. Some other works in this category [8]–[10] define the utility function as the ratio of throughput to transmit power, in order to maximize the spectrum efficiency per energy consumption. Research in the second category [11]–[18] formulates the spectrum sharing problem as an energy consumption minimization problem subject to the users' minimum throughput requirements. In this formulation, the users' throughput requirements can be explicitly specified. Hence, the spectrum efficiency is guaranteed with the minimal energy consumption. The work in this paper pertains to this second category of research works.

One major limitation of existing works in the second category [11]–[18] is that they restrict attention to a simple class of spectrum sharing policies that require the users to transmit at *fixed* power levels as long as the environment (e.g. the number of users, the channel gains) does not change[1]. We call this class of spectrum sharing policies *stationary*. The stationary policies are not energy efficient, because due to multi-user interference, the users need to transmit at high power levels to fulfill the minimum throughput constraints. To improve energy efficiency, we study *nonstationary*[2] spectrum sharing policies. Specifically, we focus on TDMA (time-division multiple access) spectrum sharing policies, a class of nonstationary policies in which the users transmit in a TDMA fashion. TDMA policies can achieve high spectrum efficiency that is not achievable under stationary policies, and greatly improve the energy efficiency of the stationary policies, because of the following two reasons. First, there is no multi-user interference in TDMA policies. Second, TDMA policies allow users to adaptively switch between transmission and dormancy, depending on the average throughput they have achieved, for the purpose of energy saving. Note that in the optimal TDMA policies we propose, users usually do not transmit in the simple round-robin fashion, because of the heterogeneity in their minimum throughput requirements and channel conditions (see Section IV for a motivating example that shows the sub-optimality of round-robin TDMA policies).

Another limitation of existing works in the second category [11]–[18] is the assumption that each user's receiver can perfectly estimate the local interference temperature (i.e. the

Manuscript received March 19, 2013; revised November 30, 2013; accepted January 6, 2014. The editor coordinating the review of this paper and approving it for publication was Y. J. Zhang.

This work was supported by NSF Grant No. 0830556.

The authors are with the Electrical Engineering Department, University of California, Los Angeles, CA 90095 USA (e-mail: yxiao@seas.ucla.edu, mihaela@ee.ucla.edu).

Digital Object Identifier 10.1109/TCOMM.2014.012614.130953

[1]Although some spectrum sharing policies [11]–[18] go through a transient period of adjusting the power levels before converging to the optimal power levels, the users maintain the fixed power levels after the convergence.

[2]We use "nonstationary", instead of "dynamic", to describe the proposed policy, because "dynamic spectrum sharing" has been extensively used to describe general spectrum sharing policies in cognitive radio, where SUs access the channel opportunistically. In this sense, our policy is dynamic. However, our nonstationary policy is different from other dynamic spectrum sharing policies, in that the power levels are time-varying.

interference and noise power level), and can accurately feed it back to its transmitter. However, in practice, users cannot perfectly estimate the interference temperature, and can only send limited (quantized) feedback.

In this paper, we provide a novel design framework to construct nonstationary spectrum sharing policies that achieve PUs' and SUs' minimum throughput requirements with minimal energy consumptions, even when the users have erroneous and very limited (only binary) feedback about their local interference temperatures. We first prove a key property of the optimal TDMA spectrum sharing policy: each user should choose the same power level whenever it transmits. This property enables us to solve the policy design problem in two tractable steps: first determine the optimal power levels before run-time, and then determine the transmission schedule at run-time. We then propose a low-complexity distributed instantaneous throughput selection (ITS) algorithm for the users to determine their optimal power levels before run-time, and a low-complexity distributed longest-distance-first (LDF) scheduling algorithm to determine the transmission schedule at run-time. We prove that both algorithms converge linearly independent of the number of users (i.e. the distance from the optimal solution decreases exponentially, resulting in a logarithmic convergence time). The proposed policy can also adapt to the dynamic entry and exit of users without affecting the convergence of existing users. Moreover, it is deviation-proof, meaning that a user cannot improve its energy efficiency over the proposed policy while still fulfilling the throughput requirement. In this way, autonomous users will find it in their self-interest to adopt the policy.

The rest of the paper is organized as follows. We give detailed comparisons against existing works in Section II. Section III describes the system model for spectrum sharing. Section IV gives a motivating example to show the performance gain achieved by nonstationary policies and the necessity of deviation-proof policies. We formulate and solve the policy design problem in Section V and Section VI, respectively. Simulation results are presented in Section VII. Finally, Section VIII concludes the paper.

II. Related Works

In this section, we provide a comprehensive comparison between the proposed scheme and existing works. Although only some works [11]–[18] use the same problem formulation as ours, we compare against a wide range of related works [1]–[29] to highlight the technical novelty of our work, and to illustrate that the works [1]–[10][19]–[29] proposed under different problem formulations cannot be adapted to our setting.

A. Stationary Spectrum Sharing Policies

Most existing works propose stationary spectrum sharing policies. We compare against them in Table I. Note that throughout this section, the feedback is the information on interference and noise power levels sent from a user's receiver to its transmitter.

TABLE I
Comparisons Against Stationary Policies.

	Energy -efficient	Feedback (Overhead)	User number	Deviation -proof
[1]–[5]	No	Error-free, unquantized	Fixed	No
[6][7]	No	Error-free, unquantized (Large)	Fixed	Yes
[8]–[16]	Yes	Error-free, unquantized (Large)	Fixed	Yes
[17][18]	Yes	Error-free, unquantized (Large)	Varying	Yes
[19]–[21]	No	Error-free, unquantized (Large)	Fixed	Yes
Proposed	Yes	Erroneous, binary (One-bit)	Varying	Yes

TABLE II
Comparisons Against Nonstationary Policies.

	[22]	[23]	[24]–[26]	Proposed
Energy -efficient	No	No	No	Yes
Power control	Yes	No	No	Yes
Users	Heterogenous	Homogenous	Homogenous	Heterogenous
Feedback (Overhead)	Error-free unquantized (Large)	Erroneous binary (One-bit)	Error-free binary (One-bit)	Erroneous binary (One-bit)
User number	Fixed	Fixed	Fixed	Varying
Deviation-proof	Yes	No	No	Yes

B. Nonstationary Spectrum Sharing Policies

There have been some works that develop nonstationary policies using repeated games [22], Markov decision processes (MDPs) [23], and multi-art bandit [24]–[26]. We summarize the major differences between the existing nonstationary policies and our proposed policy in Table II.

C. Comparison With Our Previous Work

Most related to this work is our previous work [29]. However, the design frameworks proposed in [29] and in this work are significantly different because the design objectives are different. In [29], we aimed to design TDMA spectrum sharing policies that maximize the users' total throughput without considering energy efficiency. Under this design objective, each user will transmit at the maximum power level in its slot, as long as the interference temperature constraint is not violated. Hence, what we optimized was *only the transmission schedule of the users*. In this work, since we aim to minimize the energy consumption subject to the minimum throughput requirements, we need to optimize *both the transmission schedule and the users' transmit power levels*, which makes the design problem more challenging. Moreover, this work considers the scenario in which users enter and leave the network, which is not considered in [29].

D. Comparison With Theoretical Frameworks

Our results on nonstationary policies build on the concept of "self-generating sets" proposed in the game theory literature [27]. Self-generating sets are used to analyze repeated games with imperfect monitoring. For example, the Folk Theorem in repeated games with imperfect monitoring in [28] builds on the concept of self-generating sets. However, we cannot apply this concept straightforwardly or in a way similar as in [28] for the following reasons. The self-generating set is defined as a fixed point of a set-valued mapping. The work [27] defined the set-valued mapping, and proved an important property of the fixed point of this set-valued mapping (i.e. the

TABLE III
RELATED THEORETICAL FRAMEWORKS.

	Constructive	Feedback	User number
[27]	No	N/A	Fixed
[28]	No	Erroneous, high-granularity	Fixed
Proposed	Yes	Erroneous, binary	Varying

self-generating set): every payoff vector in the self-generating set can be achieved at an equilibrium. However, although [27] discovered this important property, it did not show how to construct a self-generating set. Without constructing the self-generating set, we do not know what payoff vectors can be achieved at the equilibria or how to achieve them.

The concept of self-generating sets is applied in [28] to prove the Folk theorem in repeated games with imperfect monitoring. However, our work is fundamentally different from [28] in two aspects. First, the results in [28] are not constructive: they focus on what payoff vectors can be achieved, but not how to achieve them. In contrast, given a target payoff vector, we explicitly construct the policy to achieve it. Second, the results in [28] require a high-granularity feedback signal, namely the cardinality of feedback signals should be proportional to the number of power levels a user can choose. In contrast, by exploiting the structure of the spectrum sharing problem, we prove that binary feedback is sufficient to achieve optimality in the considered scenarios.

In Table III, we summarize the key differences between our work and [27][28].

III. SYSTEM MODEL

A. Model For Spectrum Sharing in Cognitive Radio Networks

We consider a cognitive radio network that consists of M primary users and N secondary users transmitting in a single frequency channel. The set of PUs and that of SUs are denoted by $\mathcal{M} \triangleq \{1, 2, \dots, M\}$ and $\mathcal{N} \triangleq \{M+1, M+2, \dots, M+N\}$, respectively. A wireless network in which users are not differentiated as PUs and SUs is a special case of our model with $M = 0$. Each user[3] has a transmitter and a receiver. The channel gain from user i's transmitter to user j's receiver is g_{ij}. Each user i chooses its power level p_i from a compact set $\mathcal{P}_i \subseteq \mathbb{R}_+$. We assume that $0 \in \mathcal{P}_i$, namely user i can choose not to transmit. The set of joint power profiles is denoted by $\boldsymbol{\mathcal{P}} = \prod_{i=1}^{M+N} \mathcal{P}_i$, and the joint power profile of all the users is denoted by $\boldsymbol{p} = (p_1, \dots, p_{M+N}) \in \boldsymbol{\mathcal{P}}$. Let $\boldsymbol{p}_{-i}$ be the power profile of all the users other than user i. Each user i's throughput is a function of the joint power profile, namely $r_i : \boldsymbol{\mathcal{P}} \to \mathbb{R}_+$. Since the users cannot jointly decode their signals, each user i treats the interference from the other users as noise, and obtains the following throughput at the power profile $\boldsymbol{p}$ [1]–[21]:

$$r_i(\boldsymbol{p}) = \log_2 \left(1 + \frac{p_i g_{ii}}{\sum_{j \in \mathcal{M} \cup \mathcal{N}, j \neq i} p_j g_{ji} + \sigma_i^2} \right),$$

where σ_i^2 is the noise power at user i's receiver.

We define user i's local interference temperature $I_i(\boldsymbol{p}_{-i})$ as the interference and noise power level at its receiver, namely $I_i(\boldsymbol{p}_{-i}) \triangleq \sum_{j \neq i} p_j g_{ji} + \sigma_i^2$. Each user's receiver measures the interference temperature with errors and feedback the quantized measurement to its transmitter. We assume that each user i uses a *unbiased* estimator with an additive estimation error to obtain the estimate $\hat{I}_i \triangleq I_i + \varepsilon_i$, where ε_i is the estimation error with zero mean, whose probability distribution function f_{ε_i} is known to user i. We also assume that each user i uses the following simple *two-level* quantizer Q_i:

$$Q_i(\hat{I}_i(\boldsymbol{p}_{-i})) = \begin{cases} \bar{I}_i, & \text{if } \hat{I}_i(\boldsymbol{p}_{-i}) > \theta_i \\ \underline{I}_i, & \text{otherwise} \end{cases}, \tag{1}$$

where θ_i is user i's quantization threshold, and $\bar{I}_i$ and $\underline{I}_i$ are two reconstruction values. We assume that the quantizer preserves the mean value of $\hat{I}_i(\boldsymbol{p}_{-i})$ when there is no multi-user interference. In other words, when $\boldsymbol{p}_{-i} = \mathbf{0}$ (i.e. when $I_i(\boldsymbol{p}_{-i}) = \sigma_i^2$), the quantizer should satisfy

$$\mathbb{E}_{\varepsilon_i}\{Q_i(\hat{I}_i(\boldsymbol{p}_{-i})|_{\boldsymbol{p}_{-i}=\mathbf{0}})\} = \mathbb{E}_{\varepsilon_i}\{\hat{I}_i(\boldsymbol{p}_{-i})|_{\boldsymbol{p}_{-i}=\mathbf{0}}\} = \sigma_i^2.$$

This property can be easily satisfied by setting

$$\begin{array}{l} \bar{I}_i = \int_{x-\sigma_i^2 \in \text{supp}(f_{\varepsilon_i}),\ x \geq \theta_i} x \cdot f_{\varepsilon_i}(x - \sigma_i^2) dx \\ \underline{I}_i = \int_{x-\sigma_i^2 \in \text{supp}(f_{\varepsilon_i}),\ x < \theta_i} x \cdot f_{\varepsilon_i}(x - \sigma_i^2) dx \end{array}, \tag{2}$$

where $\text{supp}(f_{\varepsilon_i})$ is the support of the distribution function f_{ε_i}. In practice, it is easy to implement an unbiased estimator and a simple two-level quantizer as in (1) and (2). As we will show later, such an estimator and a quantizer are sufficient to achieve the optimal performance.

Remark 1: Here is an intuition why an unbiased estimator and the two-level quantizer in (1) and (2) are good enough for us. For user i to achieve a minimum throughput r_i, given the feedback $Q_i(\hat{I}_i)$, its transmit power level $\hat{p}_i$ should be $\hat{p}_i = (2^{r_i} - 1) \cdot Q_i(\hat{I}_i)/g_{ii}$. In a TDMA policy, there is no multi-user interference (i.e. $\boldsymbol{p}_{-i} = \mathbf{0}$) when user i transmits. Hence, using an unbiased estimator and the quantizer in (1) and (2), user i's expected transmit power level is $\mathbb{E}_{\varepsilon_i}\{\hat{p}_i\} = \mathbb{E}_{\varepsilon_i}\left\{(2^{r_i} - 1) \cdot Q(\hat{I}_i)/g_{ii}\right\} = (2^{r_i} - 1)\mathbb{E}_{\varepsilon_i}\{Q(\hat{I}_i)\}/g_{ii} = (2^{r_i} - 1)\sigma_i^2/g_{ii}$, which is exactly the transmit power level when user i perfectly knows the interference temperature σ_i^2. In contrast, under a non-TDMA policy, there is multi-user interference. In this case, one user's erroneous and quantized feedback affects its own transmit power level, which in turn affects the others' transmit power levels through the interference. Hence, an unbiased estimator and a simple two-level quantizer in (1) and (2) may result in performance loss under non-TDMA policies.

Since each user i adopts a two-level quantizer, its feedback from the receiver to the transmitter is binary. Then we can further reduce the feedback overhead as follows. Each user i's receiver informs its transmitter of the two reconstruction values $\bar{I}_i$ and $\underline{I}_i$ only once, at the beginning, after which the receiver sends a signal, probably in the form of a simple probe, only when the estimated interference temperature $\hat{I}_i$ exceeds the quantization threshold θ_i. The event of receiving or not receiving the probing signal, which is sent only when $\hat{I}_i > \theta_i$, is enough to indicate user i's transmitter which one of the two reconstruction values it should choose. Since the probing signal indicates high interference temperature, we call it the

[3]We refer to a primary user or a secondary user as a user in general, and will specify the type of users only when necessary.

distress signal as in [12],[18]. With some abuse of definition, we denote user i's distress signal as $y_i \in Y = \{0, 1\}$ with $y_i = 1$ representing the event that user i's distress signal is sent (i.e. $\hat{I}_i > \theta_i$). We write $\rho_i(y_i|\mathbf{p})$ as the conditional probability distribution of user i's distress signal y_i given power profile $\boldsymbol{p}$, which is calculated as

$$\rho_i(y_i = 1|\boldsymbol{p}) = \int_{x>\theta_i - I_i(\boldsymbol{p}_{-i})} f_{\varepsilon_i}(x)dx.$$

B. Spectrum Sharing Policies

The system is time slotted at $t = 0, 1, 2, \ldots$. At the beginning of time slot t, each user i chooses its transmit power p_i^t, and achieves the throughput $r_i(\boldsymbol{p}^t)$. At the end of time slot t, each user j who transmits ($p_j^t > 0$) sends its distress signal $y_j^t = 1$ if the estimate $\hat{I}_j$ exceeds the threshold θ_j. We define $y \in Y$ as the *system distress signal*, indicating whether there exists a user who has sent its distress signal, namely $y = 1$ if there exists j such that $p_j > 0$ and $y_j = 1$, and $y = 0$ otherwise. The conditional distribution is denoted $\rho(y|\boldsymbol{p})$, which is calculated as $\rho(y = 0|\boldsymbol{p}) = \Pi_{j:p_j>0}\rho_j(y_j = 0|\boldsymbol{p})$. Note that the system distress signal is not a physical signal sent in the system, but rather a logical signal summarizing the status of the system. From now on, we refer to the system distress signal simply as the distress signal.

Each user i determines the transmit power level p_i^t based on the history of distress signals. The history of distress signals is $h^t = \{y^0; \ldots; y^{t-1}\} \in Y^t$ for $t \geq 1$, and $h^0 = \varnothing$ for $t = 0$. Then each user i's strategy π_i is a mapping from the set of all the possible histories to its action set, namely $\pi_i : \cup_{t=0}^{\infty} Y^t \to \mathcal{P}_i$. The *spectrum sharing policy*, denoted by $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_{M+N})$, is the joint strategy profile of all the users. Hence, user i's transmit power level at time slot t is determined by $p_i^t = \pi_i(h^t)$, and the users' joint power profile is determined by $\boldsymbol{p}^t = \boldsymbol{\pi}(h^t)$.

We classify all the spectrum sharing policies into two categories, stationary and nonstationary policies. As in [33, pp. 22] and [34, Sec. 5.5.2], stationary policies always choose the same action under the same state, while nonstationary policies may choose different actions under the same state. In our model, the state can be considered as the system parameters (e.g. the number of users, the channel conditions, etc.). Hence, a spectrum sharing policy $\boldsymbol{\pi}$ is *stationary* if and only if for all $i \in \mathcal{N}$, for all $t \geq 0$, and for all $h^t \in Y^t$, we have $\pi_i(h^t) = p_i^{\mathrm{stat}}$, where $p_i^{\mathrm{stat}} \in \mathcal{P}_i$ is a constant. A spectrum sharing policy is *nonstationary* if it is not stationary. In this paper, we restrict our attention to a special class of nonstationary polices, namely TDMA policies (with fixed transmit power levels). A spectrum sharing policy $\boldsymbol{\pi}$ is a TDMA policy if at most one user transmits in each time slot. TDMA policies are optimal when the interference among the users is strong [31], which is often the case when the number of users is large. We will illustrate how TDMA policies outperform stationary policies through a simple example in Section IV and through extensive simulations in Section VII.

Remark 2: In the formal definition of a nonstationary policy, it seems that each user needs to keep track of the history of all the past distress signals at each time slot. However, as we will see from the longest-distance-first scheduling algorithm that implements the proposed policy, each user only needs a finite memory.

C. Definition of Spectrum and Energy Efficiency

We characterize the spectrum and energy efficiency of a spectrum sharing policy by the users' *discounted* average throughput and *discounted* average energy consumption, respectively. Each user discounts its future throughput and energy consumption because of its *delay-sensitive* application (e.g. video streaming) [19]–[22][29]. A user running a more delay-sensitive application discounts more (with a lower discount factor). Assuming as in [19]–[22][29] that all the users have the same discount factor $\delta \in [0, 1)$, user i's average throughput is

$$U_i(\boldsymbol{\pi}) = \mathbb{E}_{h^0,h^1,\ldots}\left\{(1-\delta)\sum_{t=0}^{\infty}\delta^t \cdot u_i(\boldsymbol{\pi}(h^t))\right\}.$$

Similarly, user i's average energy consumption is the expected discounted average transmit power per time slot, written as

$$P_i(\boldsymbol{\pi}) = \mathbb{E}_{h^0,h^1,\ldots}\left\{(1-\delta)\sum_{t=0}^{\infty}\delta^t \cdot \pi_i(h^t))\right\}.$$

Each user i aims to minimize its average energy consumption $P_i(\boldsymbol{\pi})$ while fulfilling a minimum throughput requirement $R_i^{\min}$. From one user's perspective, it has the incentive to deviate from a given spectrum sharing policy, if by doing so it can fulfill the minimum throughput requirement with a lower average energy consumption. Hence, we can define deviation-proof policies as follows.

Definition 1: A spectrum sharing policy $\boldsymbol{\pi}$ is deviation-proof if for all $i \in \mathcal{M} \cup \mathcal{N}$, we have

$$\pi_i = \arg\min_{\pi_i'} P_i(\pi_i', \boldsymbol{\pi}_{-i}), \text{ subject to } R_i(\pi_i', \boldsymbol{\pi}_{-i}) \geq R_i^{\min},$$

where $\boldsymbol{\pi}_{-i}$ is the strategy profile of all the users except i.

IV. Motivation For Optimal Nonstationary TDMA Policies

Before formally describing the design framework, we provide a motivating example to show the advantage of the proposed optimal nonstationary TDMA policy, compared to the stationary policy and round-robin policies, in terms of both the energy efficiency and the computational complexity. Consider a simple network with three symmetric SUs. They have the same direct channel gain of $g_{ii} = 1$, the same cross channel gain of $g_{ij} = 0.25$, the same noise power $\sigma_i^2 = 5$ mW, the same minimum throughput requirement of $R_i^{\min} = 1.5$ bits/s/Hz, and the same discount factor of $\delta = 0.6$.

A. Energy Efficiency

We illustrate the policies and their performances in Table IV. The power levels are the transmit power levels of the 3 users whenever they transmit. In the optimal constant policy, users 1,2,3 all transmit all the time, at the same power level of 186 mW. In TDMA policies, whether round-robin or the proposed optimal policy, users do not all transmit all the time. For round-robin policies, we compute the optimal policy given the cycle length by determining the optimal (in terms of average long-term energy consumption across users) order of transmission in a cycle and the corresponding power levels. In

TABLE IV
ILLUSTRATION AND PERFORMANCE OF POLICIES.

Policies	Power levels (mW)	Transmission schedule (first 12 time slots)	Average energy
Optimal stationary [11]–[18]	(186,186,186)	simultaneous transmission	186 mW
Optimal round-robin (cycle length $L = 3$)	(33,144,1432)	123 123 123 123	108 mW
Optimal round-robin (cycle length $L = 4$)	(43,212,249)	1233 1233 1233	48 mW
Optimal nonstationary (proposed)	(108,108,108)	1233 2321 3231	36 mW

the optimal round-robin policy with cycle of length 3, user 1 transmits first at a low power level (33 mW), user 2 transmits after user 1 at a higher power level (133 mW) to compensate for having to wait for transmission and user 3 transmits last at a still higher power level (1432 mW) to compensate for having to wait still longer. In the cycle of length 4, again user 1 transmits at the lowest power level, user 2 transmits at a middle power level, and user 3 transmits at the highest power level, but the last two power levels are closer together (than in the cycle of length 3) because user 3 transmits more often. In the optimal nonstationary policy, the users all transmit at the same constant power level (108 mW) whenever they transmit; this works because the order in which they transmit is constantly changing. In the last column of Table 1, the discounted average energy per user per time slot is calculated. Notice that the cycle of length 3 is slightly more efficient than the constant policy, the cycle of length 4 is much more efficient, but the optimal nonstationary policy is more efficient still. Indeed, the optimal policy achieves 80%, 67% and 25% energy savings compared to the optimal constant policy, the optimal round-robin policy with cycle of 3 and with cycle of 4, respectively. Importantly, the energy savings are even more significant when the number of users is large (see Sec. VII).

B. Computational Complexity

Remarkably, not only is the proposed optimal nonstationary policy much more efficient than round-robin policies, it is much easier to compute. To get a hint of why this is so, note that in a round-robin policy, the user's performance is determined not only by the number of slots in a cycle but also by the positions of the slots since users are discounting their future throughput (due to delay sensitivity). For a given number of users $M+N$ and a given cycle length L, the number of nontrivial round-robin schedules (the ones in which each user gets at least one slot) is greater than $(M+N)^{L-(M+N)}$. So searching among these schedules will be totally impractical even if L is moderately larger than $M+N$ - but in order to achieve efficiency close to the optimal non-stationary policy, the cycle length L must be much larger than $M+N$. For instance, for the 3-user case above, achieving energy efficiency within 10% of the optimal nonstationary policy requires that the cycle length L be at least 8, and so requires searching among the thousands (5796) of different nontrivial schedules of cycle length 8 and finding power levels for each user. Even this small problem is computationally intensive. For a moderate number of users - say 10 - and a cycle length of 20 - this means searching more than ten billion (i.e. 10^{10}) schedules and finding power levels for each user - a completely intractable problem. However, we will propose a simple algorithm to compute the optimal nonstationary policy - both the schedule and the power levels - whose complexity grows only *linearly* with the number of users.

V. THE DESIGN PROBLEM FORMULATION

Our goal is to construct a deviation-proof TDMA policy that fulfills all the users' minimum throughput requirements and optimizes a certain energy efficiency criterion. The energy efficiency criterion can be represented by a function defined on all the users' average energy consumptions, $E(P_1(\boldsymbol{\pi}), \ldots, P_{M+N}(\boldsymbol{\pi}))$. Note, importantly, that the energy efficiency criterion can also reflect the priority of the PUs over the SUs. For example, the energy efficiency criterion can be the weighted sum of all the users' energy consumptions, i.e. $E(P_1(\boldsymbol{\pi}), \ldots, P_{M+N}(\boldsymbol{\pi})) = \sum_{i\in\mathcal{M}\cup\mathcal{N}} w_i \cdot P_i(\boldsymbol{\pi})$ with $w_i \geq 0$ and $\sum_{i\in\mathcal{M}\cup\mathcal{N}} w_i = 1$. Each user i's weight w_i indicates the importance of this user. We can set higher weights for PUs and lower weights for SUs.

Given each user i's minimum throughput requirement $R_i^{\min}$, we can formally define the policy design problem as

$$\begin{aligned} \min_{\boldsymbol{\pi}} \quad & E(P_1(\boldsymbol{\pi}), \ldots, P_{M+N}(\boldsymbol{\pi})) \\ s.t. \quad & \boldsymbol{\pi} \text{ is a deviation} - \text{proof TDMA policy}, \\ & R_i(\boldsymbol{\pi}) \geq R_i^{\min}, \ \forall i \in \mathcal{M}\cup\mathcal{N}. \end{aligned} \tag{3}$$

In the above problem formulation, the usual constraints on the interferences caused by SUs to PUs are satisfied by restricting to TDMA policies, in which there is no multi-user interference.

VI. A DESIGN FRAMEWORK FOR SPECTRUM AND ENERGY EFFICIENT POLICIES

We first outline the procedure to solve the policy design problem (3). Then we show in detail how to solve the design problem, and discuss implementation issues. Finally, we adapt the proposed policy to the dynamic entry and exit of users.

A. Outline of The Design Framework

The protocol design problem (3) is difficult to solve directly, because the decision variable $\boldsymbol{\pi}$ is the spectrum sharing policy, which is a mapping from the set of all histories to the set of actions. We first unravel an important property of the optimal TDMA policy, namely each user should adopt the same power level whenever it transmits (see Lemma 1). This greatly reduces the dimension of the decision variable; now we only need to find the *single* transmit power level (or equivalently, the instantaneous throughput) of each user and the transmission schedule. We propose a three-step design framework, illustrated in Fig. 1, to solve the design problem. First, we characterization of the set of feasible instantaneous throughput vectors under which the users can fulfill their throughput requirements (see Theorem 1). Based on this, we then reformulate the original problem (3) into a problem of finding the optimal instantaneous throughput vector, and propose a distributed instantaneous throughput selection (ITS) algorithm to solve the reformulated problem (see Theorem 2). Finally, given the optimal instantaneous throughput vector, we

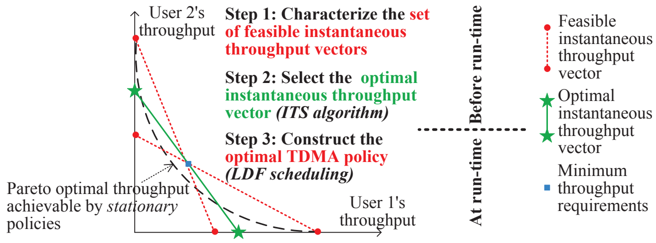


Fig. 1. The design framework to solve the policy design problem.

propose a longest-distance-first (LDF) scheduling algorithm to determine the transmission schedule, which results in the optimal TDMA policy that solves the design problem (3) (see Theorem 3). We illustrate the design framework in Fig. 1.

B. Solving The Policy Design Problem

We first prove a key property of the optimal energy-efficient TDMA protocol: each user should choose the same power level whenever it transmits.

Lemma 1: The optimal solution $\boldsymbol{\pi}^*$ to the design problem (3) must satisfy that each user i chooses the same power level whenever it transmits, namely $\pi_i^*(t_1) = \pi_i^*(t_2)$ for all t_1 and t_2 such that $\pi_i^*(t_1) > 0$ and $\pi_i^*(t_2) > 0$.

Proof: See Appendix A. ■

Lemma 1 greatly simplifies the design problem: now we only need to find a single optimal power level p_i^* for each user i to choose whenever it transmits, instead of solving for its optimal power levels in all its transmissions. In the following, we first find the optimal power levels $\{p_i^*\}_{i\in\mathcal{M}\cup\mathcal{N}}$ during the users' transmissions (which is equivalent to finding each user i's optimal instantaneous throughput, $r_i^* \triangleq \log_2\left(1+\frac{g_{ii}p_i^*}{\sigma_i^2}\right)$). Then given $\{p_i^*\}_{i\in\mathcal{M}\cup\mathcal{N}}$ (or $\{r_i^*\}_{i\in\mathcal{M}\cup\mathcal{N}}$), we find the transmission schedule that achieves the minimum throughput requirements.

1) Step 1 – Characterizing feasible instantaneous throughput vectors: Now we formulate the problem of finding the users' optimal instantaneous throughput $\{r_i^*\}_{i\in\mathcal{M}\cup\mathcal{N}}$. First, the structure of the optimal TDMA protocol discovered in Lemma 1 enables us to establish the following relationship between the average throughput and the average energy consumption:

$$\begin{aligned}\frac{P_i(\pi_i)}{R_i(\boldsymbol{\pi})} &= \frac{(1-\delta)\sum_{t=0}^{\infty}\delta^t \mathbf{1}_{\{\pi_i(t)>0\}}p_i^{\text{tdma}}}{(1-\delta)\sum_{t=0}^{\infty}\delta^t \mathbf{1}_{\{\pi_i(t)>0\}}\log_2\left(1+\frac{g_{ii}p_i^{\text{tdma}}}{\sigma_i^2}\right)} \\ &= \frac{p_i^{\text{tdma}}}{\log_2\left(1+\frac{g_{ii}p_i^{\text{tdma}}}{\sigma_i^2}\right)} = \frac{\sigma_i^2}{g_{ii}}\cdot\frac{2^{r_i^{\text{tdma}}}-1}{r_i^{\text{tdma}}}, \quad (4)\end{aligned}$$

where $\mathbf{1}_{\{\cdot\}}$ is the indicator function, p_i^{tdma} is user i's power level when it transmits in the TDMA protocol, and r_i^{tdma} is the corresponding instantaneous throughput. We can see from (4) that given r_i^{tdma}, the average energy consumption $P_i(\pi_i)$ is proportional to the average throughput $R_i(\boldsymbol{\pi})$. Hence, to minimize the energy consumption, we should let $R_i(\boldsymbol{\pi}) = R_i^{\min}$ for all i. Then based on (4), we can rewrite the objective function $E(P_1(\pi_1),\ldots,P_{M+N}(\pi_{M+N}))$ of the design problem (3) as a function of the instantaneous throughput $\{r_i^{\text{tdma}}\}_{i\in\mathcal{M}\cup\mathcal{N}}$:

$$E\left(\frac{\sigma_1^2}{g_{11}}\frac{2^{r_1^{\text{tdma}}}-1}{r_1^{\text{tdma}}}R_1^{\min},\ldots,\frac{\sigma_{M+N}^2}{g_{M+NM+N}}\frac{2^{r_{M+N}^{\text{tdma}}}-1}{r_{M+N}^{\text{tdma}}}R_{M+N}^{\min}\right).$$

An instantaneous throughput vector $\{r_i^{\text{tdma}}\}_{i\in\mathcal{M}\cup\mathcal{N}}$ is feasible, if there exists a TDMA protocol $\boldsymbol{\pi}$ that has the instantaneous throughput $\{r_i^{\text{tdma}}\}_{i\in\mathcal{M}\cup\mathcal{N}}$ and can achieve the minimum average throughput $\{R_i^{\min}\}_{i\in\mathcal{M}\cup\mathcal{N}}$. Before characterizing the feasible instantaneous throughput vectors, we write $\tilde{\boldsymbol{p}}^i = (p_i^{\text{tdma}}(r_i^{\text{tdma}}), \boldsymbol{p}_{-i} = \mathbf{0})$ as the joint power profile when user i transmits in a TDMA policy. Now we state Theorem 1.

Theorem 1: An instantaneous throughput vector $\{r_i^{\text{tdma}}\}_{i\in\mathcal{M}\cup\mathcal{N}}$ is feasible for the minimum throughput requirements $\{R_i^{\min}\}_{i\in\mathcal{M}\cup\mathcal{N}}$, if the following conditions are satisfied:

- Condition 1: the discount factor δ satisfies $\delta \geq \underline{\delta} \triangleq 1/\left(1+\frac{1-\sum_{i\in\mathcal{M}\cup\mathcal{N}}\underline{\mu}_i}{M+N-1+\sum_{i\in\mathcal{M}\cup\mathcal{N}}\sum_{j\neq i}(-\rho(y=1|\tilde{\boldsymbol{p}}^i)/b_{ij})}\right)$, where $b_{ij} = \sup_{p_j\in\mathcal{P}_j, p_j\neq\tilde{p}_j^i}\frac{\rho(y=1|\tilde{\boldsymbol{p}}^i)-\rho(y=1|p_j,\tilde{\boldsymbol{p}}_{-j}^i)}{r_j(p_j,\tilde{\boldsymbol{p}}_{-j}^i)/\bar{r}_j}$, and $\underline{\mu}_i \triangleq \max_{j\neq i}\frac{1-\rho(y=1|\tilde{\boldsymbol{p}}^i)}{-b_{ij}}$.
- Condition 2: $\sum_i \frac{R_i^{\min}}{r_i^{\text{tdma}}} = 1$, and $r_i^{\text{tdma}} \leq R_i^{\min}/\underline{\mu}_i$.

Proof: See Appendix B. ■

The problem of finding the optimal instantaneous throughput $\{r_i^*\}_{i\in\mathcal{M}\cup\mathcal{N}}$ can then be formulated as

$$\begin{aligned}\min_{\{r_i^{\text{tdma}}\}_{i\in\mathcal{M}\cup\mathcal{N}}} \quad & E\left(\left\{\frac{\sigma_i^2}{g_{ii}}\frac{2^{r_i^{\text{tdma}}}-1}{r_i^{\text{tdma}}}R_i^{\min}\right\}_{i\in\mathcal{M}\cup\mathcal{N}}\right) \\ s.t. \quad & \textstyle\sum_{i\in\mathcal{M}\cup\mathcal{N}} R_i^{\min}/r_i^{\text{tdma}} = 1, \qquad (5) \\ & 0 < r_i^{\text{tdma}} \leq \bar{r}_i \triangleq R_i^{\min}/\underline{\mu}_i, \ \forall i\in\mathcal{M}\cup\mathcal{N}.\end{aligned}$$

2) Step 2 – Select the optimal instantaneous throughput vector: We solve the above optimization problem (5) for the optimal instantaneous throughput vector $\{r_i^*\}_{i\in\mathcal{M}\cup\mathcal{N}}$ using the distributed ITS algorithm, which is proved to converge in logarithmic time in Theorem 2.

The ITS algorithm essentially solves the following equation (derived from the KKT condition) in a distributed fashion:

$$\left.\frac{\partial E}{\partial P_i}\right|_{P_i=\frac{\sigma_i^2 R_i^{\min}}{g_{ii}}\frac{2^{r_i^*}-1}{r_i^*}}\left(2^{r_i^*}-1-r_i^*\cdot 2^{r_i^*}\ln 2\right)\frac{\sigma_i^2}{g_{ii}} = -\lambda, (6)$$

where λ is the Lagrangian multiplier for the constraint $\sum_i R_i^{\min}/r_i^{\text{tdma}} = 1$ in (5), and should be chosen such that $\sum_i R_i^{\min}/r_i^* = 1$. The term $\frac{\partial E}{\partial P_i}$ in (6) is the derivative of the energy efficiency criterion $E(\cdot)$ with respect to user i's average energy consumption. If the energy efficiency criterion is the weighted sum of all the users' energy consumptions, we have $\left.\frac{\partial E}{\partial P_i}\right|_{P_i=\frac{\sigma_i^2 R_i^{\min}}{g_{ii}}\frac{2^{r_i^*}-1}{r_i^*}} = w_i, \ \forall r_i^*$. If the energy efficiency criterion is the weighted proportional fairness $-\sum_{i\in\mathcal{M}\cup\mathcal{N}} w_i \log(P_i)$, we have $\left.\frac{\partial E}{\partial P_i}\right|_{P_i=\frac{\sigma_i^2 R_i^{\min}}{g_{ii}}\frac{2^{r_i^*}-1}{r_i^*}} = -w_i\frac{g_{ii}}{\sigma_i^2 R_i^{\min}}\cdot\frac{r_i^*}{2^{r_i^*}-1}$. Each user i selects the term $\frac{\partial E}{\partial P_i}$ in the ITS algorithm based on the energy efficiency criterion chosen by the protocol designer.

Theorem 2: The problem (5) of finding the optimal instantaneous throughput vector can be converted into a convex optimization problem, whose solution $\{r_i^*\}_{i\in\mathcal{M}\cup\mathcal{N}}$ can be

Algorithm 1 Instantaneous Throughput Selection (ITS) algorithm run by user i.

Require: Minimum throughput requirement $R_i^{\min}$, precision e

Set $\underline{\lambda} = 0$, $\bar{\lambda} = 1$, $\lambda = \bar{\lambda}$.
Solve (6) for r_i^*, set $r_i^* \leftarrow \min\{r_i^*, \bar{r}_i\}$
Broadcast $R_i^{\min}/r_i^*$, and receive $R_j^{\min}/r_j^*$ for all $j \neq i$
while $\sum_{j \in \mathcal{M} \cup \mathcal{N}} R_j^{\min}/r_j^* > 1$ **do**
  $\bar{\lambda} \leftarrow 2 \cdot \bar{\lambda}$, $\lambda \leftarrow \bar{\lambda}$
  Solve (6) for r_i^*, set $r_i^* \leftarrow \min\{r_i^*, \bar{r}_i\}$
  Broadcast $R_i^{\min}/r_i^*$, and receive $R_j^{\min}/r_j^*$ for all $j \neq i$
end while
while $\left|\sum_{j \in \mathcal{M} \cup \mathcal{N}} R_j^{\min}/r_j^* - 1\right| > e$ **do**
  $\lambda \leftarrow \frac{\underline{\lambda} + \bar{\lambda}}{2}$
  Solve (6) for r_i^*, set $r_i^* \leftarrow \min\{r_i^*, \bar{r}_i\}$
  Broadcast $R_i^{\min}/r_i^*$, and receive $R_j^{\min}/r_j^*$ for all $j \neq i$
  if $\sum_{j \in \mathcal{M} \cup \mathcal{N}} R_j^{\min}/r_j^* < 1$ **then**
    $\bar{\lambda} \leftarrow \lambda$
  else
    $\underline{\lambda} \leftarrow \lambda$
  end if
end while
Normalize $r_i^* \leftarrow r_i^* / \left(\sum_{j \in \mathcal{M} \cup \mathcal{N}} R_j^{\min}/r_j^*\right)$

Algorithm 2 The Longest-Distance-First (LDF) scheduling run by user i.

Require: $\{R_j^{\min}/r_j^*\}_{j \in \mathcal{M} \cup \mathcal{N}}$, r_i^*
Initialization: Set $t = 0$, $r'_j(0) = R_j^{\min}/r_j^*$ for all $j \in \mathcal{M} \cup \mathcal{N}$
repeat
  Calculates the distance from the optimal operating point $d_j(t) = \frac{r'_j(t) - \underline{\mu}_j}{1 - r'_j(t)} \rho(y = 1|\tilde{\mathbf{p}}^j), \forall j$
  Find the user with the largest distance $i^* \triangleq \arg\max_{j \in \mathcal{M} \cup \mathcal{N}} d_j(t)$
  if $i = i^*$ **then**
    Transmit at power level $p_i^{\mathrm{tdma}}(r_i^*)$
  end if
  Updates $r'_j(t+1)$ for all $j \in \mathcal{M} \cup \mathcal{N}$ as follows:
  if No Distress Signal Received At Time Slot t **then**
    $r'_{i^*}(t+1) = \frac{1}{\delta} \cdot r'_{i^*}(t) - (\frac{1}{\delta} - 1) \cdot (1 + \sum_{j \neq i^*} \frac{\rho(y=1|\tilde{\mathbf{p}}^{i^*})}{-b_{i^*j}})$
    $r'_j(t+1) = \frac{1}{\delta} \cdot r'_j(t) + (\frac{1}{\delta} - 1) \cdot \frac{\rho(y=1|\tilde{\mathbf{p}}^{i^*})}{-b_{i^*j}}, \forall j \neq i^*$
  else
    $r'_{i^*}(t+1) = \frac{1}{\delta} \cdot r'_{i^*}(t) - (\frac{1}{\delta} - 1) \cdot (1 - \sum_{j \neq i^*} \frac{\rho(y=0|\tilde{\mathbf{p}}^{i^*})}{-b_{i^*j}})$
    $r'_j(t+1) = \frac{1}{\delta} \cdot r'_j(t) - (\frac{1}{\delta} - 1) \cdot \frac{\rho(y=0|\tilde{\mathbf{p}}^{i^*})}{-b_{i^*j}}, \forall j \neq i^*$
  end if
  $t \leftarrow t + 1$
until $\varnothing$

found by each user running the distributed ITS algorithm. The algorithm converges linearly[4] at rate $\frac{1}{2}$.

Proof: See Appendix C. ■

3) Step 3 – Construct the optimal deviation-proof policy: Given the optimal instantaneous throughput vector, each user i runs the longest-distance-first scheduling algorithm in a decentralized manner. On one hand, the transmission schedule can be viewed as a simple "largest-distance-first" scheduling, namely the user farthest away from its throughput requirement transmits. On the other hand, it is nontrivial to define the "distance" from its throughput requirement. As we will prove later, user j's distance from its throughput requirement can be defined as $d_j(t) = \frac{r'_j(t) - \underline{\mu}_j}{1 - r'_j(t) + \sum_{k \neq j}(-\rho(y=1|\tilde{\mathbf{p}}^j)/b_{jk})}$, where $r'_j(t)$ is the future throughput to achieve starting from time slot t normalized by r_j^*. The normalized future throughput $r'_j(t)$ can be also interpreted the future transmission opportunity. If user j transmitted all the time in the future, it would have an average throughput r_j^*. If it transmits in a fraction $r'_j(t)$ of time after time t, it has an average future throughput of $r'_j(t) \cdot r_j^*$.

Theorem 3 proves the desirable properties of the LDF scheduling algorithm.

Theorem 3: If each user $i \in \mathcal{M} \cup \mathcal{N}$ runs the LDF scheduling algorithm, then we have

- each user i can achieve its minimum throughput requirement $R_i^{\min}$ with an energy consumption P_i that minimizes the energy efficiency criterion $E(P_1, \ldots, P_{M+N})$;
- if a user does not follow the algorithm, it will either fail to achieve the minimum throughput requirement, or achieve it with a higher energy consumption;
- the distance between each user i's average throughput at time t and its throughput requirement decreases exponentially with time, namely

$$\left|(1-\delta)\sum_{\tau=0}^{t} \delta^{\tau} \cdot r_i^{\tau} - R_i^{\min}\right| \leq r_i^* \cdot \delta^{t+1}. \quad (7)$$

Proof: See Appendix D. ■

Theorems 2 and 3 establish the convergence results of our proposed scheme. Theorem 2 proves that the process of finding the optimal instantaneous throughput vector converges in logarithmic time, and Theorem 3 proves that the LDF scheduling achieves the minimum throughput requirements in logarithmic time. Hence, the overall convergence speed is fast.

Note that our convergence results are very different from the convergence results in some recent works on power control in cognitive radio [4] and wireless networks [5]. These works [4][5] belong to the stationary spectrum sharing policies, namely they aim to find the optimal fixed power levels of the users that maximize the network utility. The convergence results in [4][5] differ from our results in two important ways. First, since our work studies nonstationary spectrum sharing with time-varying power levels, we need to determine not only the optimal power levels of the users, but also the transmission schedule of the users. We prove that the average throughput obtained by adopting the proposed LDF scheduling converges linearly. Such a result does not appear in [4][5]. Second, the techniques used in proving the convergence to the optimal power levels are different. In [4][5], the algorithms are akin to the celebrated distributed power control algorithm [11], and hence the proofs use and extend the "standard interference function" argument. Such an argument is not used in our

[4]Following [32, Sec. 9.3.1], we define linear convergence as follows. Suppose that the sequence $\{x_k\}$ converges to x. We say that this sequence converges linearly at rate c, if we have $\lim_{k \to \infty} \frac{|x_{k+1} - x|}{|x_k - x|} = c$.

work since there is no interference among the users under the proposed TDMA spectrum sharing policy.

C. Implementation

We discuss the total overhead of information exchange and feedback and the computational complexity of the proposed scheme.

1) Overhead of initial information exchange and feedback: In Table V, we compare the overhead of information exchange and feedback of the proposed framework with the energy efficient spectrum sharing policies proposed in [11]–[17] and [18] for wireless networks and cognitive radio networks, respectively. Before run-time, the information exchange in the proposed framework comes from the ITS algorithm ($(M+N)\cdot\mathcal{O}(\log_2(1/e))$ with e being the performance loss tolerance) and the exchange of b_{ij} for the LDF scheduling. The exchange of b_{ij} is for deviation-proofness. However, in the run time, the feedback overhead of the proposed policy is significantly lower than that of [11]–[18]. Specifically, in [11]–[18], each user i's receiver needs to feedback the interference temperature I_{-i} in *each time slot*. Hence, the total amount of feedback in [11]–[18] grows linearly with time. In conclusion, our proposed framework has a much lower total overhead than [11]–[18].

2) Computational complexity: The implementation of the proposed policy includes the ITS algorithm before run-time and the LDF scheduling at run-time. First, both the ITS algorithm and the LDF scheduling converge fast in logarithmic time as proved in Theorems 2 and 3. Second, each iteration in the ITS algorithm involves solving the equation (6), which can be done efficiently using the Newton method. Each iteration in the LDF scheduling involves computing $M+N$ indices $\{d_j(t)\}_{j\in\mathcal{M}\cup\mathcal{N}}$ and $M+N$ normalized values $\{r'_j(t)\}_{j\in\mathcal{M}\cup\mathcal{N}}$, all of which are determined by analytical expressions. Finally, although the original definition of the policy requires each user to memorize the entire history of distress signals, in the LDF scheduling, each user only needs to know the current distress signal y^t and memorize $M+N$ normalized values $\{r'_j(t)\}_{j\in\mathcal{M}\cup\mathcal{N}}$. In conclusion, the overall computational complexity of each user in implementing the proposed policy is small.

D. Users Entering and Leaving the Network

We adapt the protocol to the scenario where users enter and leave the network. We divide time into *epochs*, where a new epoch begins when users enter or leave. The system starts at epoch 0, and we denote the optimal instantaneous throughput in epoch 0 by $r_i^{(0)}$. When new users enter or existing users leave at t_1, each of them broadcasts a "ENTER" or "EXIT" signal, respectively. Upon receiving such a signal, the users run the ITS algorithm again to determine the optimal instantaneous throughput in epoch 1, $r_i^{(1)}$. Note that for each existing user i, the input to the ITS algorithm is the continuation throughput at t_1, namely $\gamma_i(t_1)$; while for each new user j, the input should be its minimum throughput $R_j^{\min}$. Then they run the LDF scheduling with the new instantaneous throughput, until a new epoch begins when the "ENTER" or "EXIT" signals

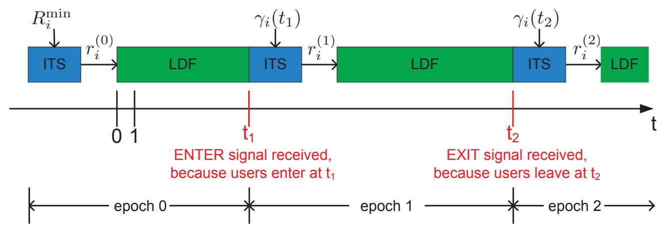


Fig. 2. The proposed protocol implemented by user i when it receives "ENTER" signal at t_1 and "EXIT" signal at t_2.

are broadcast by some users at t_2. We illustrate how to adapt the protocol in Fig. 2.

One nice property of the proposed protocol is that, the convergence of the LDF scheduling is not affected by users coming or leaving.

Theorem 4: In the proposed spectrum sharing protocol, each user's average throughput converges to the minimum throughput requirement in logarithmic time, even with users entering and leaving the network.

Proof: See Appendix E. ■

Note that we can also deal with the changes of system parameters (e.g. the channel gains) in the same way as we deal with the dynamic entry and exit of users. Specifically, whenever a user observes a change in the system parameters, it can broadcast a signal that triggers the users to run the ITS algorithm and the LDF scheduling again. The convergence result in Theorem 4 also applies to this case.

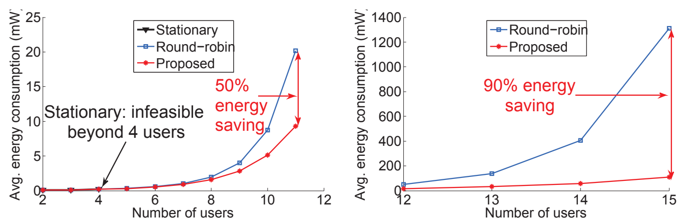
In some works [17] for energy efficient power control in wireless networks, the locally stable asymptotic convergence of the proposed algorithm is proved. The locally stable asymptotic convergence guarantees that slight perturbation from the equilibrium (induced by, for example, an incoming user) will not make the algorithm diverge. However, the convergence result in Theorem 4 are different from that in [17]. Specifically, we study the convergence of not only the transmit power levels, but also the transmission schedule, which is not studied in [17]. More importantly, the influence of dynamic entry and exit of users on the convergence and stability is quite different in our work as compared to [17]. Since our proposed policy is TDMA, there is no interference among the users. Hence, an incoming user will not interfere with the existing users when they transmit. In other words, the influence of incoming users is not through the interference as in [17], but through acquiring the transmission opportunities of the existing users. We show that under such perturbation (in terms of transmission opportunities), the proposed LDF scheduling still converges to the target throughput at the same rate.

VII. Performance Evaluation

In this section, we demonstrate the performance gain of our spectrum sharing policy over existing policies, and validate our theoretical analysis through numerical results. Throughout this section, we use the following system parameters by default unless we change some of them explicitly. The noise powers at all the users' receivers are 0.05 W. For simplicity, we assume that the direct channel gains have the same distribution $g_{ii}\sim\mathcal{CN}(0,1),\forall i$, and the cross channel gains have the same distribution $g_{ij}\sim\mathcal{CN}(0,0.25),\forall i\neq j$. The users have the

TABLE V
COMPARISON OF THE TOTAL OVERHEAD OF INITIAL INFORMATION EXCHANGE AND FEEDBACK.

	Information exchange before run-time	Feedback at run-time
[11]–[17]	N/A	Each user i: I_{-i} *each time slot* Amount: $M+N$ reals *per time slot*
[18]	Spectrum coordinator to each user: degradation of its minimum throughput requirement Amount: $M+N$ reals	Each user i: I_{-i} *each time slot*, distress signal if necessary Amount: $M+N$ reals *per time slot*, distress signal if necessary
Proposed	Each user i broadcasts to other users: $\rho(y=1\|\tilde{\mathbf{p}}^i)$ and $\{b_{ji}\}_{j\neq i}$ once, and $R_i^{\min}/r_i^*$ at each iteration of the ITS algorithm; Each user i's receiver to its transmitter: $\bar{I}_i, \underline{I}_i$ Amount: $(M+N)^2+(M+N)\cdot\mathcal{O}(\log_2(1/e))$ reals	distress signal if necessary Amount: distress signal if necessary



(a) Small numbers of users. (b) Large numbers of users.

Fig. 3. Energy efficiency of the stationary, round-robin, and proposed policies under different numbers of users.

same minimum throughput requirement of 1 bits/s/Hz. The discount factor is 0.95. The interference temperature threshold is $\theta = 1$ W. The measurement error ε is Gaussian distributed with zeros mean and variance 0.1. The energy efficiency criterion is the average energy consumption across users.

A. Comparisons Against Existing Policies

First, assuming that the population is fixed, we compare the proposed policy against the optimal stationary policy in [11]–[18], and the optimal round-robin policy with cycle length $L = M + N$ (i.e. each user gets one slot in a cycle). We compare the energy efficiency of the policies as the number of users increase in Fig. 3. Each data point plotted is the average of 1000 channel realizations. First, we can see that the stationary policy becomes infeasible when the number of users is more than 4. In contrast, the round-robin and proposed policies remain feasible when the number of users increases. Second, the proposed policy achieves significant energy saving compared to the round-robin policy, especially when the number of users is large. Specifically, it achieves 50% and 90% energy saving compared to the round-robin policy when the number of users is 11 and 15, respectively. These are exactly the deployment scenarios where improvements in spectrum and energy efficiency are much needed.

B. Adapting to Users Entering and Leaving the Network

We demonstrate how the proposed policy can seamlessly adapt to the entry and exit of PUs/SUs. We consider a network with 10 PUs and 2 SUs initially. The PUs' minimum throughput requirements range from 0.2 bits/s/Hz to 0.38 bits/s/Hz with 0.02 bits/s/Hz increments, namely PU n has a minimum throughput requirement of $0.2 + (n-1) * 0.02$ bits/s/Hz. The SUs' have the same minimum throughput requirement of 0.1 bits/s/Hz. We show the dynamics of average energy consumptions and throughput of several PUs and all the SUs in Fig. 4 and Fig. 5, respectively.

In the first 100 time slots, we can see that all the users quickly achieve the minimum throughput requirements at around $t = 50$. PUs have different energy consumptions because of their different minimum throughput requirements. The two SUs converge to the same average energy consumption and average throughput. There are SUs leaving ($t = 100$) and entering ($t = 150, 250$), and a PU entering ($t = 200$). We can see that during the entire process, the PUs/SUs that are initially in the system maintain the same throughput and energy consumption. The new PU (PU 11) has a higher energy consumption, because of its higher minimum throughput requirement (0.4 bits/s/Hz), and because of the limited transmission opportunities left for it. SU 3, however, does not need a higher energy consumption because it occupies the time slots originally assigned to SU 2, who left the network at $t = 100$. But SU 4 does need a higher energy consumption, because there are more SUs and less transmission opportunities in the network after $t = 250$.

VIII. CONCLUSION

In this paper, we proposed nonstationary spectrum sharing policies that allow the PUs and SUs to transmit in a TDMA fashion. The proposed policy can achieve high spectrum efficiency that is not achievable by existing policies, and is more energy efficient than existing policies under the same minimum throughput requirements. The proposed policy can achieve high spectrum and energy efficiency even when the users have erroneous and binary feedback of the interference temperature. We extend the policy to the case with users entering and leaving the network, while still maintaining the spectrum and energy efficiency of the existing users. The proposed policy is amenable to decentralized implementation and is deviation-proof. Simulation results demonstrate significant performance gains over state-of-the-art policies. Interesting future research directions include how to design the optimal policy when the feedback is finer than binary and when the users have different delay sensitivities (i.e. different discount factors).

APPENDIX A
PROOF OF LEMMA 1

Suppose that in the optimal TDMA protocol $\boldsymbol{\pi}^*$, there exists a user i and two time slots $t_1 \neq t_2$, such that $0 < \pi_i^*(t_1) < \pi_i^*(t_2)$ (note that we do not assume $t_1 < t_2$ or $t_1 > t_2$). We will find another protocol $\boldsymbol{\pi}'$ that fulfills the same minimum throughput requirements with lower energy consumptions, which contradicts the fact that $\boldsymbol{\pi}^*$ is optimal.

We construct the protocol $\boldsymbol{\pi}'$ as follows. The transmission strategies of the users other than user i remain the same,

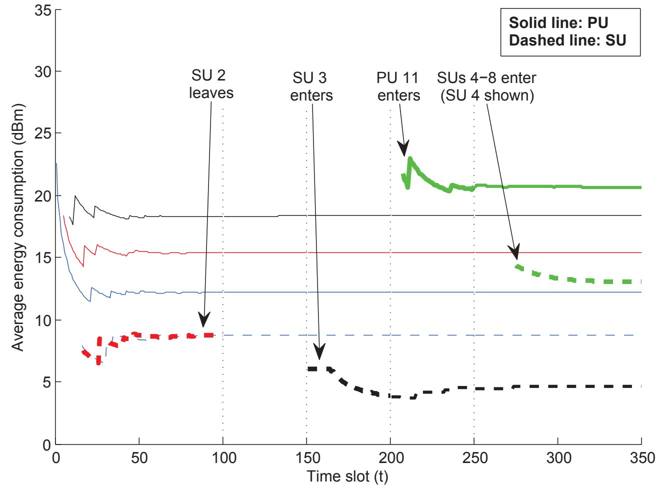


Fig. 4. Dynamics of average energy consumption with users entering and leaving the network. At $t = 0$, there are 10 PUs and 2 SUs. SU 2 leaves at $t = 100$. SU 3 enters at $t = 150$. PU 11 enters at $t = 200$. SUs 4–8 enter at $t = 250$. We only show PUs 1, 5, 9, 11 (solid lines) and SUs 1, 2, 3, 4 (dashed lines) in the figure.

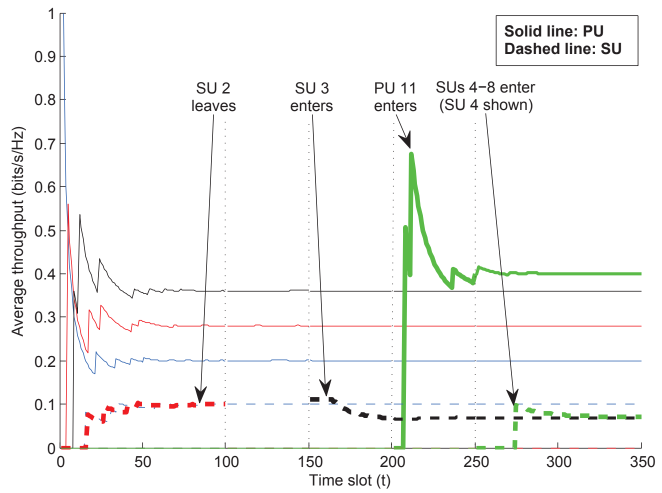


Fig. 5. Dynamics of average throughput under the same dynamics of the entry and exit of users as in Fig. 4.

namely $\boldsymbol{\pi}'_{-i} = \boldsymbol{\pi}^*_{-i}$. For user i, the transmission remains the same for the time slots other than t_1 and t_2, namely $\pi'_i(t) = \pi^*_i(t), \forall t \neq t_1, t_2$. Then we increase user i's power level at t_1 by $\epsilon_1 > 0$, i.e. $\pi'_i(t_1) = \pi^*_i(t_1) + \epsilon_1$, and decrease its power level at t_2 by $\epsilon_2 > 0$, i.e. $\pi'_i(t_2) = \pi^*_i(t_2) - \epsilon_2$. To maintain user i's average throughput, ϵ_1 and ϵ_2 should satisfy

$$\begin{aligned} & \delta^{t_1} \log_2\left(1+\frac{g_{ii}p_i^*(t_1)}{\sigma_i^2}\right)+\delta^{t_2}\log_2\left(1+\frac{g_{ii}p_i^*(t_2)}{\sigma_i^2}\right) \\ = \; & \delta^{t_1}\log_2\left[1+\frac{g_{ii}(p_i^*(t_1)+\epsilon_1)}{\sigma_i^2}\right]+\delta^{t_2}\log_2\left[1+\frac{g_{ii}(p_i^*(t_2)-\epsilon_2)}{\sigma_i^2}\right]. \end{aligned}$$

Given ϵ_1, we can calculate ϵ_2 as a function of ϵ_1, $\epsilon_2(\epsilon_1)$. Then the decrease in average energy consumption by switching to protocol $\boldsymbol{\pi}'$ can be calculated as $\Delta(\epsilon_1) = -\delta^{t_1}\epsilon_1 + \delta^{t_2}\epsilon_2(\epsilon_1)$. Taking the derivative of $\Delta(\epsilon_1)$ with respect to ϵ_1, we have

$$\frac{\partial\Delta}{\partial\epsilon_1} = \delta^{t_1}\left[\frac{\sigma_i^2+g_{ii}p_i^*(t_2)}{\sigma_i^2+g_{ii}(p_i^*(t_1)+\epsilon_1)}\left(\frac{\sigma_i^2+g_{ii}p_i^*(t_1)}{\sigma_i^2+g_{ii}(p_i^*(t_1)+\epsilon_1)}\right)^{\delta^{t_1-t_2}}-1\right].$$

Since $p_i^*(t_2) > p_i^*(t_1)$, we have $\frac{\partial\Delta}{\partial\epsilon_1} > 0$ when $\epsilon_1 = 0$. Since $\frac{\partial\Delta}{\partial\epsilon_1}$ is continuous in ϵ_1 when $\epsilon_1 \geq 0$, we can find a small enough $\zeta > 0$, such that $\frac{\partial\Delta}{\partial\epsilon_1} > 0$ for all $\epsilon_1 \in [0, \zeta]$. Hence, the decrease $\Delta(\epsilon_1)$ in user i's average energy consumption by switching to $\boldsymbol{\pi}'$ is positive for any $\epsilon_1 \in [0, \zeta]$. This contradicts with the fact that $\boldsymbol{\pi}^*$ is optimal, which proves the lemma.

Appendix B
Proof of Theorem 1

Due to space limitation, we present the proof of a simplified version of Theorem 1 in the special case when the users are not self-interested. This proof will illustrate the main idea of the complete proof. Please refer to [30, Appendix B] for the complete proof of Theorem 1.

Specifically, we prove the following lemma on the feasible instantaneous throughput when the users are obedient. The lemma is a special case of Theorem 1 by setting $b_{ij}^+ = -\infty$ for all i, j.

Lemma 2: When the users are obedient, an instantaneous throughput vector $\{r_i^{\text{tdma}}\}_{i\in\mathcal{M}\cup\mathcal{N}}$ is feasible for the minimum throughput requirements $\{R_i^{\min}\}_{i\in\mathcal{M}\cup\mathcal{N}}$, if

- the discount factor δ satisfies $\delta \geq 1 - \frac{1}{M+N}$,
- $\sum_{i\in\mathcal{M}\cup\mathcal{N}} R_i^{\min}/r_i^{\text{tdma}} = 1$.

Proof: As in dynamic programming, we can decompose each user i's discounted average throughput into the current throughput and the *continuation throughput* as follows:

$$\begin{aligned} R_i(\boldsymbol{\pi}) &= (1-\delta)\sum_{t=0}^{\infty}\delta^t\cdot\left(\mathbf{1}_{\{\pi_i(t)>0\}}\cdot r_i^{\text{tdma}}\right) \\ &= (1-\delta)\cdot\underbrace{\left(\mathbf{1}_{\{\pi_i(0)>0\}}\cdot r_i^{\text{tdma}}\right)}_{\text{the current throughput at } t=0} \\ &+ \delta\cdot\underbrace{\left[(1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}\cdot\left(\mathbf{1}_{\{\pi_i(t)>0\}}\cdot r_i^{\text{tdma}}\right)\right]}_{\text{the continuation throughput starting from } t=1}. \end{aligned}$$

We can see that the continuation throughput starting from $t = 1$ is the discounted average throughput as if the system starts from $t = 1$. In general, we can define user i's continuation throughput starting from t as $\gamma_i(t) \triangleq (1-\delta)\sum_{\tau=t}^{\infty}\delta^{\tau-t}\cdot(\mathbf{1}_{\{\pi_i(\tau)>0\}}\cdot r_i^{\text{tdma}})$. Then the decomposition at time t can be written as $\gamma_i(t) = (1-\delta)\cdot(\mathbf{1}_{\{\pi_i(t)>0\}}\cdot r_i^{\text{tdma}}) + \delta\cdot\gamma_i(t+1)$. Write the continuation throughput vector as $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_N)$.

Definition 2 (Self-generating set): A set of throughput vectors $\mathcal{R}$ is a self-generating set, if for any throughput vector $\boldsymbol{\gamma} \in \mathcal{R}$, there exists a $i^* \in \mathcal{N}$ and a continuation throughput vector $\boldsymbol{\gamma}' \in \mathcal{R}$ such that for all $i \in \mathcal{N}$, we have $\gamma_i = (1-\delta)\cdot(\mathbf{1}_{\{i=i^*\}}\cdot r_i^{\text{tdma}}) + \delta\cdot\gamma'_i$.

An important property of the self-generating set, proved in [27], is that any throughput vector in $\mathcal{R}$ can be achieved by a TDMA protocol. This is because for any throughput vector $\boldsymbol{\gamma} \in \mathcal{R}$, we can schedule a user i^* to transmit in the current time slot, and the resulting continuation throughput vector $\boldsymbol{\gamma}'$ starting from the next time slot can be decomposed (by a user to transmit and the following continuation throughput vector) again. We can do the above decomposition iteratively to determine the transmission schedule.

Consider the following set of throughput vectors $\mathcal{R} = \left\{\boldsymbol{\gamma} : \sum_{i\in\mathcal{M}\cup\mathcal{N}}\frac{\gamma_i}{r_i^{\text{tdma}}} = 1, \gamma_i \geq 0, \forall i\right\}$. We derive the condition on the discount factor δ such that $\mathcal{R}$ is self-generating.

For a given vector $\boldsymbol{\gamma} \in \mathcal{R}$, if we let user i to transmit, the continuation throughput vector $\boldsymbol{\gamma}'$ is

$$\gamma_i' = \frac{\gamma_i}{\delta} - \frac{1-\delta}{\delta} \cdot r_i^{\mathrm{tdma}}, \text{ and } \gamma_j' = \frac{\gamma_j}{\delta}, \ \forall j \neq i. \quad (8)$$

To ensure $\boldsymbol{\gamma}' \in \mathcal{R}$, the discount factor must satisfy $\delta \geq 1 - \frac{\gamma_i}{r_i^{\mathrm{tdma}}}$. Hence, to ensure that any $\boldsymbol{\gamma} \in \mathcal{R}$ can be decomposed, the discount factor must satisfy

$$\delta \geq \max_{\boldsymbol{\gamma} \in \mathcal{R}} \min_{i \in \mathcal{M} \cup \mathcal{N}} \left\{1 - \gamma_i / r_i^{\mathrm{tdma}}\right\} = 1 - \frac{1}{M+N}, \quad (9)$$

where the optimal solution is achieved when $\gamma_i = \frac{1}{M+N} r_i^{\mathrm{tdma}}, \forall i$. ■

Appendix C
Proof of Theorem 2

We first convert the optimization problem (5) into a convex optimization problem. Defining $x_i = \frac{1}{r_i^{\mathrm{tdma}}}$, the objective function can be rewritten as

$$E\left(\left\{\frac{\sigma_i^2 R_i^{\min}}{g_{ii}} \cdot (2^{\frac{1}{x_i}} - 1) \cdot x_i\right\}_{i \in \mathcal{M} \cup \mathcal{N}}\right).$$

Based on our assumption, $E(\cdot)$ is convex and increasing in each argument $\frac{\sigma_i^2 R_i^{\min}}{g_{ii}} \cdot (2^{\frac{1}{x_i}} - 1) \cdot x_i$. According to the composition rule [32, Sec. 3.2.4], $E(\cdot)$ is a convex function of $(x_1, \ldots, x_{M+N})$ if $\frac{\sigma_i^2 R_i^{\min}}{g_{ii}} \cdot (2^{\frac{1}{x_i}} - 1) \cdot x_i$ is convex in x_i, which can be proved easily by taking the second derivative of $(2^{\frac{1}{x_i}} - 1) \cdot x_i$ with respect to x_i. Hence, the objective function is a convex function of $(x_1, \ldots, x_{M+N})$. It is not difficult to see that the constraints in (5) can be rewritten as linear constraints $\sum_{i \in \mathcal{M} \cup \mathcal{N}} R_i^{\min} \cdot x_i = 1$ and $x_i \geq \frac{1}{\bar{r}_i}$. As a result, the following optimization problem with decision variables $(x_1, \ldots, x_{M+N})$

$$\begin{aligned} \min_{(x_1, \ldots, x_{M+N})} \quad & E\left(\left\{\frac{\sigma_i^2 R_i^{\min}}{g_{ii}} \cdot (2^{\frac{1}{x_i}} - 1) \cdot x_i\right\}_{i \in \mathcal{M} \cup \mathcal{N}}\right) \\ s.t. \quad & \textstyle\sum_i R_i^{\min} \cdot x_i = 1, \ x_i \geq 1/\bar{r}_i, \ \forall i \ , \end{aligned} \quad (10)$$

is a convex optimization problem.

We solve (10) by looking at the KKT conditions. Write λ as the Lagrangian multiplier of the constraint $\sum_{i \in \mathcal{M} \cup \mathcal{N}} R_i^{\min} \cdot x_i = 1$, and $\mu_i \geq 0$ as the Lagrangian multiplier of the inequality $x_i \geq \frac{1}{\bar{r}_i}$. The optimal $(x_1^*, \ldots, x_{M+N}^*)$ and the optimal λ^* and μ_i^* should satisfy the KKT conditions:

$$\frac{\partial E}{\partial x_i}\Big|_{x_i = x_i^*} - \mu_i^* = -\lambda^* R_i^{\min} \quad (11)$$

with $\mu_i^* = 0$ when $x_i^* > \frac{1}{\bar{r}_i}$, due to the complementary slackness condition. Hence, the problem (10) can be solved by finding the optimal λ^*, such that the solutions $(x_1^*, \ldots, x_{M+N}^*)$ to the equations (11) satisfy the equality $\sum_{i \in \mathcal{N}} R_i^{\min} \cdot x_i = 1$. Equivalently, we can find the optimal λ^* such that the optimal instantaneous throughput $(r_1^*, \ldots, r_{M+N}^*)$ satisfy

$$\frac{\partial E}{\partial x_i^*}\Big|_{x_i^* = \frac{1}{r_i^*}} - \mu_i^* = -\lambda^* R_i^{\min}, \text{ with } \mu_i^* = 0 \text{ if } r_i^* < \bar{r}_i, \quad (12)$$

and $\sum_{i \in \mathcal{M} \cup \mathcal{N}} R_i^{\min} / r_i^* = 1$.

Since the first-order derivative $\frac{\partial E}{\partial x_i^*}$ is monotone in x_i (because the second-order derivative is always positive), we can find the optimal λ^* using the bisection method, which converges linearly with rate $\frac{1}{2}$.

Appendix D
Proof of Theorem 3

Due to space limitation, we present the proof of a simplified version of Theorem 3 in the special case when the users are not self-interested. Please refer to [30, Appendix C] for the complete proof of Theorem 3.

This proof is closely related to the proof of Theorem 1. Recall that for each continuation throughput vector $\boldsymbol{\gamma}(t)$ at time t, if we choose user i to transmit, we can calculate the resulting continuation throughput vector $\boldsymbol{\gamma}(t+1)$ at time $t+1$ as in (8). The proof ofTheorem 1 ensures that as long as we choose the user to transmit at time t based on $i = \arg\min_{j \in \mathcal{M} \cup \mathcal{N}} \{1 - \gamma_j(t)/r_i^*\}$ (see (9)), the continuation throughput vector $\boldsymbol{\gamma}(t+1)$ at time $t+1$ will also be achievable. The LDF scheduling schedules the transmission exactly in this way in each time slot. By setting the continuation throughput at time 0 as $\gamma_i(0) = R_i^{\min}$, each user i can achieve the average throughput $R_i^{\min}$. Since the instantaneous throughput is the optimal one, r_i^*, the energy efficiency criterion is minimized.

Note that $R_i^{\min} = (1-\delta)\sum_{\tau=0}^{\infty} \delta^\tau \cdot (\mathbf{1}_{\pi_i(\tau)>0} \cdot r_i^*) = (1-\delta)\sum_{\tau=0}^{t} \delta^\tau \cdot (\mathbf{1}_{\pi_i(\tau)>0} \cdot r_i^*) + (1-\delta)\sum_{\tau=t+1}^{\infty} \delta^\tau \cdot (\mathbf{1}_{\pi_i(\tau)>0} \cdot r_i^*)$. Since $0 \leq (1-\delta)\sum_{\tau=t+1}^{\infty} \delta^\tau \cdot (\mathbf{1}_{\pi_i(\tau)>0} \cdot r_i^*) \leq (1-\delta)\sum_{\tau=t+1}^{\infty} \delta^\tau \cdot r_i^* = \delta^{t+1} \cdot r_i^*$, we have $|(1-\delta)\sum_{\tau=0}^{t} \delta^\tau \cdot (\mathbf{1}_{\pi_i(\tau)>0} \cdot r_i^*) - R_i^{\min}| \leq r_i^* \cdot \delta^{t+1}$.

Appendix E
Proof of Theorem 4

For a user i, consider the distance between its average throughput at time t and its minimum throughput $R_i^{\min}$. Suppose that each time slot τ is in the k_τth epoch (time slot t is in the ℓth epoch), and that the beginning of the kth epoch is t_k with $t_0 = 0$. Then the distance is

$$\begin{aligned} & \left|(1-\delta)\textstyle\sum_{\tau=0}^{t} \delta^\tau (\mathbf{1}_{\pi_i(\tau)>0} \cdot r_i^{(k_\tau)}) - R_i^{\min}\right| \\ = & \left|\left[(1-\delta)\textstyle\sum_{\tau=t_0}^{t_1-1} \delta^\tau (\mathbf{1}_{\pi_i(\tau)>0} \cdot r_i^{(0)}) - R_i^{\min}\right]\right. \\ + & \left.(1-\delta)\textstyle\sum_{\tau=t_1}^{t} \delta^\tau (\mathbf{1}_{\pi_i(\tau)>0} \cdot r_i^{(k_\tau)})\right| \\ = & \left|(1-\delta)\textstyle\sum_{\tau=t_1}^{t} \delta^\tau (\mathbf{1}_{\pi_i(\tau)>0} \cdot r_i^{(k_\tau)}) - \gamma_i(t_1)\right| \\ = & \left|(1-\delta)\textstyle\sum_{\tau=t_\ell}^{t} \delta^\tau (\mathbf{1}_{\pi_i(\tau)>0} \cdot r_i^{(\ell)}) - \gamma_i(t_\ell)\right| . \end{aligned}$$

Since $\gamma_i(t_\ell)$ is the input to the LDF scheduling at the beginning of the ℓth epoch, from Theorem 3, we have $\left|(1-\delta)\sum_{\tau=t_\ell}^{t} \delta^\tau (\mathbf{1}_{\pi_i(\tau)>0} \cdot r_i^{(\ell)}) - \gamma_i(t_\ell)\right| \leq r_i^{(\ell)} \cdot \delta^{t+1}$. Hence, the distance between the average throughput and the minimum throughput requirement decreases exponentially with time even with users entering and leaving.

References

[1] C. W. Tan and S. H. Low, "Spectrum management in multiuser cognitive wireless networks: optimality and algorithm," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 2, pp. 421–430, Feb. 2011.

[2] J. Huang, R. A. Berry, and M. L. Honig, "Distributed interference compensation for wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 5, pp. 1074–1084, May 2006.

[3] N. Gatsis, A. G. Marques, G. B. Giannakis, "Power control for cooperative dynamic spectrum access networks with diverse QoS constraints," *IEEE Trans. Commun.*, vol. 58, no. 3, pp. 933–944, Mar. 2010.

[4] L. Zheng and C. W. Tan, “Cognitive radio network duality and algorithms for utility maximization,” *IEEE J. Sel. Areas Commun.*, vol. 31, no. 3, pp. 500–513, Mar. 2013.

[5] C. W. Tan, M. Chiang, and R. Srikant, “Fast algorithms and performance bounds for sum rate maximization in wireless networks,” *IEEE/ACM Trans. Netw.*, vol. 21, no. 3, pp. 706–719, Jun. 2013.

[6] J. Huang, R. Berry, and M. Honig, “Auction-based spectrum sharing,” *Mobile Networks and Applications*, vol. 11, pp. 405–418, 2006.

[7] Y. Xiao, J. Park, and M. van der Schaar, “Intervention in power control games with selfish users,” *IEEE J. Sel. Topics Signal Process.*, vol. 6, pp. 165–179, 2012.

[8] C. U. Saraydar, N. B. Mandayam, and D. J. Goodman, “Efficient power control via pricing in wireless data networks,” *IEEE Trans. Commun.*, vol. 50, no. 2, pp. 291–303, Feb. 2002.

[9] G. He, S. Lasaulce, and Y. Hayel, “Stackelberg games for energy-efficient power control in wireless networks,” in *Proc. 2011 IEEE INFOCOM*, pp. 591–595.

[10] R. Xie, F. R. Yu, and H. Ji, “Energy-efficient spectrum sharing and power allocation in cognitive radio femtocell networks,” *Proc. 2012 IEEE INFOCOM*, pp. 1665–1673.

[11] R. Yates, “A framework for uplink power control in cellular radio systems,” *IEEE J. Sel. Areas Commun.*, vol. 13, pp. 1341–1347, 1995.

[12] N. Bambos, S. Chen, and G. Pottie, “Channel access algorithms with active link protection for wireless communication networks with power control,” *IEEE/ACM Trans. Netw.*, vol. 8, no. 5, pp. 583–597, Oct. 2000.

[13] M. Xiao, N. B. Shroff, and E. K. P. Chong, “A utility-based power control scheme in wireless cellular systems,” *IEEE/ACM Trans. Netw.*, vol. 11, no. 2, pp. 210–221, Apr. 2003.

[14] E. Altman and Z. Altman, “S-modular games and power control in wireless networks,” *IEEE Trans. Autom. Control*, vol. 48, no. 5, pp. 839–842, May 2003.

[15] P. Hande, S. Rangan, M. Chiang, and X. Wu, “Distributed uplink power control for optimal SIR assignment in cellular data networks,” *IEEE/ACM Trans. Netw.*, vol. 16, no. 6, pp. 1420–1433, Dec. 2008.

[16] S. M. Perlaza, H. Tembine, S. Lasaulce, and M. Debbah, “Quality-of-service provisioning in decentralized networks: a satisfaction equilibrium approach,” *IEEE J. Sel. Topics Signal Process.*, vol. 6, pp. 104–116, Apr. 2012.

[17] C. W. Tan, D. P. Palomar, and M. Chiang, “Energy-robustness tradeoff in cellular network power control,” *IEEE/ACM Trans. Netw.*, vol. 17, no. 3, pp. 912–925, Jun. 2009.

[18] S. Sorooshyari, C. W. Tan, and M. Chiang, “Power control for cognitive radio networks: axioms, algorithms, and analysis,” *IEEE/ACM Trans. Netw.*, vol. 20, no. 3, pp. 878–891, Jun. 2012.

[19] R. Etkin, A. Parekh, and D. Tse, “Spectrum sharing for unlicensed bands,” *IEEE J. Sel. Areas Commun.*, vol. 25, pp. 517–528, Apr. 2007.

[20] Y. Wu, B. Wang, K. J. R. Liu, and T. C. Clancy, “Repeated open spectrum sharing game with cheat-proof strategies,” *IEEE Trans. Wireless Commun.*, vol. 8, no. 4, pp. 1922–1933, 2009.

[21] M. Le Treust and S. Lasaulce, “A repeated game formulation of energy-efficient decentralized power control,” *IEEE Trans. Wireless Commun.*, vol. 9, no. 9, pp. 2860–2869, Sep. 2010.

[22] Y. Xiao, J. Park, and M. van der Schaar, “Repeated games with intervention: theory and applications in communications,” accepted by *IEEE Trans. Commun.* Available: http://arxiv.org/abs/1111.2456.

[23] Q. Zhao, L. Tong, A. Swami, and Y. Chen, “Decentralized cognitive MAC for opportunistic spectrum access in ad hoc networks: a POMDP framework,” *IEEE J. Sel. Areas Commun.*, vol. 25, pp. 589–600, 2007.

[24] K. Liu and Q. Zhao, “Distributed learning in multi-armed bandit with multiple players,” *IEEE Trans. Signal Process.*, vol. 58, no. 11, pp. 5667–5681, Nov., 2010.

[25] K. Liu, Q. Zhao, and B. Krishnamachari, “Dynamic multichannel access with imperfect channel state detection,” *IEEE Trans. Signal Process.*, vol. 58, no. 5, May 2010.

[26] K. Liu and Q. Zhao, “Cooperative game in dynamic spectrum access with unknown model and imperfect sensing,” *IEEE Trans. Wireless Commun.*, vol. 11, no. 4, Apr. 2012.

[27] D. Abreu, D. Pearce, and E. Stacchetti, “Toward a theory of discounted repeated games with imperfect monitoring,” *Econometrica*, vol. 58, no. 5, pp. 1041–1063, 1990.

[28] D. Fudenberg, D. K. Levine, and E. Maskin, “The folk theorem with imperfect public information,” *Econometrica*, vol. 62, no. 5, pp. 997–1039, Sep. 1994.

[29] Y. Xiao and M. van der Schaar, “Dynamic spectrum sharing among repeatedly interacting selfish users with imperfect monitoring,” *IEEE J. Sel. Areas Commun.*, vol. 30, no. 10, pp. 1890–1899, Nov. 2012.

[30] Y. Xiao and M. van der Schaar, “Energy-efficient nonstationary spectrum sharing,” technical report. Available: http://arxiv.org/abs/1211.4174

[31] B. Makki and T. Eriksson, “On the ergodic achievable rates of spectrum sharing networks with finite backlogged primary users and an interference indicator signal,” *IEEE Trans. Wireless Commun.*, vol. 11, no. 9, pp. 3079–3089, Jul. 2012.

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

[33] E. Altman, *Constrained Markov Decision Processes*. Chapman & Hall/CRC, 1999.

[34] G. J. Mailath and L. Samuelson, *Repeated Games and Reputations: Long-run Relationships*. Oxford University Press, 2006.

Yuanzhang Xiao received the B.E. and M.E. degree in electrical engineering from Tsinghua University, Beijing, China, in 2006 and 2009, respectively. He is currently pursuing the Ph.D. degree in the Electrical Engineering Department at University of California, Los Angeles. His research interests include multi-agent sequential decision making, game theory, and optimization.

Mihaela van der Schaar is Chancellor’s Professor of Electrical Engineering at University of California, Los Angeles. She is an IEEE Fellow. She received an NSF CAREER Award (2004), the Best Paper Award from IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY (2005), the Okawa Foundation Award (2006), the IBM Faculty Award (2005, 2007, 2008), the Most Cited Paper Award from EURASIP: *Image Communications Journal* (2006), the Gamenets Conference Best Paper Award (2011) and the 2011 IEEE Circuits and Systems Society Darlington Award Best Paper Award. For more information about her research visit: http://medianetlab.ee.ucla.edu/