Dynamic Energy-Efficient Spectrum Sharing Among Autonomous Users

Yuanzhang Xiao and Mihaela van der Schaar

Abstract

In this paper, we develop a novel design framework for energy-efficient spectrum sharing among autonomous users, who aim to minimize their transmit power levels subject to their minimum throughput requirements. Most existing works proposed *static* spectrum sharing policies, in which the users transmit at constant power levels as long as the system parameters (e.g. direct and cross channel gains) remain unchanged. Since the users transmit simultaneously under static spectrum sharing policies, they need to transmit at high power levels in order to fulfill the minimum throughput requirements due to multi-user interference. To improve the energy efficiency, we construct *dynamic* spectrum sharing policies, in which the users can transmit at time-varying power levels even when the system parameters do not change. Specifically, we focus on a class of dynamic spectrum sharing policies in which the users transmit in a time-division multiple access (TDMA) fashion. Due to the absence of multi-user interference during the transmission, this class of policies can greatly improve the spectrum and energy efficiency compared to existing policies. In addition, the constructed policies have the following desirable properties. First, the policies can be implemented by the users in a decentralized manner. Second, they are deviation-proof, namely the autonomous users will find it in their self-interests to follow the policies. Third, the policies can achieve high energy efficiency while guaranteeing the minimum throughput, even when the users can only observe the interference power imperfectly. Simulation results validate our analytical results and quantify the performance gains enabled by the proposed dynamic spectrum sharing policies.

I. Introduction

In many emerging wireless networks such as cognitive radio networks, there are autonomous users sharing the spectrum [1]. In these networks, the spectrum sharing policies, which specify the users' transmit power levels, are essential to improve the spectrum and energy efficiency [2].

Y. Xiao and M. van der Schaar are with Department of Electrical Engineering, UCLA. Email: {yxiao,mihaela}@ee.ucla.edu.

There are two substantial classes of works on designing spectrum sharing policies. In the first class of works, the users aim to maximize their utilities subject to maximum transmit power constraints [3]–[10]. While many works in this class [7]–[8] define the utility function as the throughput and do not concern energy efficiency, some works [9][10] define the utility function as the throughput to transmit power ratio in order to achieve spectrum and energy efficiency. In the second class of works [11]–[18], the users aim to minimize their transmit power levels subject to minimum throughput requirements. In these works, the energy consumption is minimized, while spectrum efficiency is also achieved because all the users can coexist in the system and have successful transmission. We follow the second approach and develop a design framework for spectrum and energy efficient spectrum sharinghhh.

Since users can use the spectrum as long as they are in the network, they usually coexist in the system for relatively long periods of time [19]–[22]. In general, the optimal spectrum sharing policy should be *dynamic*, namely they should allow users to transmit at different power levels temporally even when the environment (e.g. the number of users, the channel gains) remains unchanged. However, most existing spectrum sharing policies require the users to transmit at *constant* power levels over the time horizon in which they interact[1] [3]–[18]. We call these spectrum sharing policies *static*. The static policies are inefficient in energy consumption in many spectrum sharing scenarios where cross channel gains are strong, because the users need to transmit at high power levels to fulfill the minimum throughput constraints due to multi-user interference. To improve energy efficiency, we study *dynamic* spectrum sharing policies. Specifically, we focus on TDMA spectrum sharing policies, a class of dynamic spectrum sharing policies in which the users transmit in a TDMA fashion. Due to the absence of multi-user interference during the transmission, TDMA policies can achieve high spectrum efficiency that is not achievable under static policies, and greatly improve the energy efficiency of the static policies.

One of the challenges in designing TDMA spectrum sharing policies is to deal with the autonomy of users. Autonomous users may deviate from the prescribed spectrum sharing policy, if by doing so their energy consumption can be reduced while fulfilling the minimum throughput

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

requirement. Hence, the spectrum sharing policy should be *deviation-proof*, which means that a user cannot improve its energy efficiency and still fulfill the throughput requirement. In this way, autonomous users will find it in their self-interest to follow the policy. In the works where the users minimize their transmit power subject to throughput requirements [11]–[18], the optimal static spectrum sharing policy, if it exists, is deviation-proof against other static policies, but not deviation-proof against dynamic policies. In contrast, our proposed spectrum sharing policy is deviation-proof against both static and dynamic policies.

Another challenge in the design is that the users are unable to perfectly observe the behavior of each other, which results in inaccuracy in detecting deviating behavior.. Without perfect monitoring of the deviating behavior, the punishment may be triggered by mistake, resulting in a loss in energy efficiency or a violation of the minimum throughput requirement. The existing deviation-proof dynamic policies in [19]–[22] were designed under the assumption that each user can perfectly monitoring the others' behavior. In these policies, a punishment phase, in which all the users transmit at high power levels, is triggered when deviation happens. The inefficient resource allocation in the punishment phase does not reduce spectrum or energy efficiency under perfect monitoring, because the punishment phase serves as a threat and is never triggered. However, under imperfect monitoring, due to monitoring errors, the punishment phase will be triggered with a positive probability even when no user deviates. Hence, the deviation-proof dynamic policies in [19]–[22] are inefficient under imperfect monitoring. On the contrary, our proposed policy has no performance loss under imperfect monitoring.

In this paper, we design dynamic spectrum sharing policies that are deviation-proof given the imperfect monitoring of the users. We provide a systematic design approach, which first characterizes the set of deviation-proof policies that fulfill the minimum throughput requirements, and then for a given energy efficiency criterion, choose the optimal deviation-proof policy. The proposed policy can be easily implemented by the users in a decentralized manner, and can achieve energy efficiency with throughput requirements fulfilled even under imperfect monitoring of the users.

Table I categorizes existing works based on four criteria: the policy is dynamic or static, whether the policy can be implemented in a decentralized fashion or not, whether the policy is deviation-proof (against static policies or dynamic policies) or not, and whether there is performance loss due to imperfect monitoring. Note that we only consider the performance loss due to imperfect monitoring for the deviation-proof policies.

The rest of the paper is organized as follows. In Section II, we describe the system model for spectrum sharing. Then, in Section III, we give a motivating example to show the performance gain of using dynamic policies and the necessity of constructing deviation-proof policies. In Section IV, we formulate the policy design problem, solve it in three phases, and discuss related design and implementation issues. Simulation results are presented in Section V. Finally, Section VI concludes the paper.

II. SYSTEM MODEL

A. Model For Dynamic Spectrum Sharing

We consider a single-hop wireless ad hoc network, where a fixed set of N users transmit in a single frequency channel (see Fig 1 for an illustrative example of two users). The set of the users is denoted by $\mathcal{N} \triangleq \{1, 2, \ldots, N\}$. Each user has a transmitter and a receiver. The channel gain from user j's transmitter to user i's receiver is g_{ij}. Each user i chooses a 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 $\mathcal{P} = \prod_{i=1}^{N} \mathcal{P}_i$, and the joint power profile of all the users is denoted by $\mathbf{p} = (p_1, \ldots, p_N) \in \mathcal{P}$. Let $\mathbf{p}_{-i}$ 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 $r_i : \mathcal{P} \rightarrow \mathbb{R}$. Since the users cannot jointly decode their signals, each user i treats the interference from the other users as noise, and obtains the throughput

$$r_i(\mathbf{p}) = \log_2 \left(1 + \frac{g_{ii} p_i}{\sum_{j \neq i} g_{ij} p_j + \sigma_i^2} \right). \tag{1}$$

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

As in [17][24]–[27], there is a local spectrum server (LSS) serving as a mediating entity among the users. The LSS has a transmitter and a receiver and can measure the interference temperature (i.e. the total receive power) at its receiver, but it cannot control the actions of the autonomous users. The LSS, aiming to improve the spectrum and energy efficiency, could be a device deployed by some regulatory agency such as Federal Communications Commission (FCC), who uses it for spectrum management in that local geographic area.

The LSS measures the interference temperature (IT) at its receiver imperfectly. The measurement can be written as $\sum_{i \in \mathcal{N}} g_{0i} p_i + \varepsilon$, where g_{0i} is the channel gain from user i's transmitter to the LSS's receiver, and ε is the additive measurement error. We assume that the measurement error has zero mean and a probability distribution function (p.d.f.) f_ε known to the LSS. When the measurement $\sum_{i\in\mathcal{N}} g_{0i}p_i+\varepsilon$ exceeds an IT threshold I set by the LSS, the LSS will broadcast a distress signal to all the users. We write the outcome of measuring the interference temperature as $y \in Y = \{0, 1\}$ with $y = 0$ indicating that the IT threshold is not exceeded, namely

$$y = \begin{cases} 1, & \text{if } \sum_{i\in\mathcal{N}} g_{0i}p_i + \varepsilon > I \\ 0, & \text{otherwise} \end{cases}. \tag{2}$$

We write the conditional probability distribution of the outcome y given the action profile $\mathbf{p}$ as $\rho(y|\mathbf{p})$, which can be calculated as

$$\rho(y = 0|\mathbf{p}) = \int_{x \leq I - \sum_{i\in\mathcal{N}} g_{0i}p_i} f_\varepsilon(x)dx, \text{ and } \rho(y = 1|\mathbf{p}) = 1 - \rho(y = 0|\mathbf{p}). \tag{3}$$

Similar to [7]–[18], we assume that the system parameters, such as the number of users and the channel gains, remain fixed during the considered time horizon. The system is time slotted at $t = 0, 1, \ldots$. We assume as in [7]–[18] that the users are synchronized. At the beginning of time slot t, each user i chooses its action p_i^t, and achieves the throughput $r_i(\mathbf{p}^t)$. At the end of time slot t, the LSS measures the interference temperature $\sum_{i\in\mathcal{N}} g_{0i}p_i^t + \varepsilon^t$, where ε^t is the realization of the error ε at time slot t. If the outcome at time t is $y^t = 1$, the LSS broadcast the distress signal to the users. In other words, the LSS feedbacks the binary information about the interference temperature. However, note that it does not need to broadcast the distress signal every time slot; it broadcast only when the IT threshold is exceeded.

B. Spectrum Sharing Policies

In a general spectrum sharing policy, each user should determine the transmit power level at each time slot t based on all the available information, namely the history of its own transmit powers up to time t, the history of its interference and noise power levels at its receiver up to time t, and the history of the outcomes of the IT measurement by the LSS. However, the computational complexity of such a policy is too high. In this paper, we focus on a class of low-complexity spectrum sharing policies, in which each user i determines the transmit power level p_i^t based only on the history of the outcomes of the IT measurement. The history of outcomes up to time slot $t \geq 1$ is $h^t = \{y^0; \ldots ; y^{t-1}\} \in Y^t$, and that at time slot 0 is $h^0 = \varnothing$. Formally, we consider a class of spectrum sharing policies $\pi = (\pi_1, \ldots, \pi_N)$, in which each user i's strategy π_i is a mapping from the set of all possible histories $\mathcal{H} \triangleq \cup_{t=0}^{\infty} Y^t$ to its action set $\mathcal{P}_i$. In other words, user i's transmit power level at time slot t is determined by $p_i^t = \pi_i(h^t)$, and the users' joint transmit power profile is determined by $\mathbf{p}^t = \pi(h^t)$. We can classify all the spectrum sharing policies into two categories, static and dynamic policies, as follows.

Definition 1: A spectrum sharing policy π is *static* if and only if for all $i \in \mathcal{N}$, for all $t \geq 0$, and for all $h^t \in \mathcal{H}$, we have $\pi_i(h^t) = p_i^{\text{static}}$, where $p_i^{\text{static}} \in \mathcal{P}_i$ is a constant. A spectrum sharing policy is *dynamic* if and only if it is not static.

To further simplify the computational complexity of the spectrum sharing policy, we restrict our attention to a special class of dynamic polices, namely the TDMA policies with constant power levels defined as follows.

Definition 2 (TDMA policies with constant power levels): A spectrum sharing policy π is a TDMA policy with constant power levels if and only if for all $t \geq 0$, only one user transmits, namely

$$\exists i \in \mathcal{N}, s.t.\ \pi_i(h^t) > 0, \text{ and } \pi_j(h^t) = 0,\ \forall j \neq i, \tag{4}$$

and for all $i \in \mathcal{N}$, for all $t \geq 0$, for all $h^t \in \mathcal{H}$, we have $\pi_i(h^t) = p_i^{\text{TDMA}}$, where $p_i^{\text{TDMA}} \in \mathcal{P}_i$ is a constant.

A TDMA policy with constant power levels is completely specified by each user i's transmit power level p_i^{TDMA} when it transmits and by a schedule of which user transmits at each time slot t. Hence, such a policy can be relatively easily constructed by the designer and implemented by the users. Since we focus on this special class of dynamic policies, we refer to "TDMA policy with constant power levels" when we say "dynamic policy" in the rest of the paper.

Remark 1: In the formal definition of a dynamic policy, each user needs to keep track of the history of all the past outcomes to determine whether to transmit at each time slot. However, for the proposed policies shown later in Table III, each user only needs to analytically compute N indices to determine whether to transmit, where N is the number of users. Hence, each user can have a finite memory to implement the proposed policy.

C. Definition of Spectrum and Energy Efficiency

The spectrum and energy efficiency of a spectrum sharing policy is characterized by each user's average throughput and average transmit power. A user's average throughput is defined as the expected discounted average throughput per time slot. Assuming as in [19]–[23] that all the users have the same discount factor $\delta \in [0,1)$, user i's average throughput can be written as

$$R_i(\pi) = \mathbb{E}_{h^0,h^1,\ldots}\left\{(1-\delta)\left[r_i(\pi(h^0)) + \sum_{t=1}^{\infty}\delta^t r_i(\pi(h^t))\right]\right\}. \tag{5}$$

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

$$P_i(\pi) = \mathbb{E}_{h^0,h^1,\ldots}\left\{(1-\delta)\left[\pi_i(h^0) + \sum_{t=1}^{\infty}\delta^t \pi_i(h^t)\right]\right\}. \tag{6}$$

Each user i aims to minimize the power consumption $P_i(\pi)$ 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 power consumption. Hence, we can define deviation-proof policies as follows.

Definition 3: A spectrum sharing policy π is deviation-proof if for all i, we have

$$P_i(\pi) \leq P_i(\pi_i', \pi_{-i}) \text{ for any } \pi_i' \text{ such that } R_i(\pi_i', \pi_{-i}) \geq R_i^{\min}, \tag{7}$$

where π_{-i} is the joint strategy of all the users other than user i.

III. Motivation For Deviation-proof Dynamic Spectrum Sharing Policies

Before formally describing the design framework, we provide a motivating example to show why it is beneficial to study dynamic spectrum sharing policies. Consider a simple network with two symmetric users. For simplicity, we assume the direct channel gains are both 1, and the cross channel gains are both $\alpha > 0$, i.e., $g_{ii} = 1$ and $g_{ij} = \alpha$ for $i = 1,2$ and $j \neq i$. The noise power at each user' receiver is $\sigma_i^2 = 0.05$ for $i = 1,2$. Hence, if one user transmits at $p_i = 1$ without interference, the throughput is 4.39 bits/s/Hz.

If the users adopt the static spectrum sharing policy, their minimum transmit power should be $p_1^{\text{static}} = p_2^{\text{static}} = \frac{1}{1-(2^r-1)\alpha}\cdot(2^r-1)\cdot n$, when their minimum throughput requirement is r. We can see that the average transmit power, $P_i^{\text{static}} = p_i^{\text{static}}, i = 1,2$, increases with the cross interference level α. Moreover, the static policy is infeasible when $\alpha \geq \frac{1}{2^r-1}$, namely when the cross interference level or the minimum throughput requirement is very high.

Now suppose that the users adopt a simple dynamic spectrum sharing policy, in which user 1 transmits at a constant power level p_1^{dynamic} in even time slots $t = 0, 2, \ldots$ and user 2 transmits at a constant power level p_2^{dynamic} in odd time slots $t = 1, 3, \ldots$. The users' throughput are

$$R_1 = (1-\delta) \cdot \sum_{t=0}^{\infty} \delta^{2t} \log_2 \left(1 + p_1^{\text{dynamic}}/n\right) = \frac{1}{1+\delta} \log_2 \left(1 + p_1^{\text{dynamic}}/n\right), \tag{8}$$

$$R_2 = (1-\delta) \cdot \sum_{t=0}^{\infty} \delta^{2t+1} \log_2 \left(1 + p_2^{\text{dynamic}}/n\right) = \frac{\delta}{1+\delta} \log_2 \left(1 + p_2^{\text{dynamic}}/n\right). \tag{9}$$

Given the throughput requirement, we can calculate p_1^{dynamic} and p_2^{dynamic} from the above equations. Hence, to fulfill the same throughput requirement r, the users' average transmit power should be

$$P_1^{\text{dynamic}} = (1-\delta) \cdot \sum_{t=0}^{\infty} \delta^{2t} p_1^{\text{dynamic}} = \frac{1}{1+\delta} \left(2^{r(1+\delta)} - 1\right), \tag{10}$$

$$P_2^{\text{dynamic}} = (1-\delta) \cdot \sum_{t=0}^{\infty} \delta^{2t+1} p_2^{\text{dynamic}} = \frac{\delta}{1+\delta} \left(2^{r(1+\frac{1}{\delta})} - 1\right). \tag{11}$$

Note that, as opposed to the static policy, the average transmit power in the dynamic policy is independent of the cross interference level. Hence, the dynamic policy is better in medium to high interference levels, when the static policy may not even be feasible.

We can compare the energy efficiency (in terms of the total average transmit power) of the static and dynamic policies under some representative parameter values. Suppose that the throughput requirement is $r = 1$, and that the discount factor is 0.9. Then we have

$$\begin{aligned} &P_1^{static} = P_2^{static} = \tfrac{1}{1-\alpha} \cdot n; \\ &P_1^{dynamic} = \tfrac{1}{1+\delta} \cdot \left(2^{1+\delta} - 1\right) \cdot n, P_2^{dynamic} = \tfrac{\delta}{1+\delta} \cdot \left(2^{1+\frac{1}{\delta}} - 1\right) \cdot n, \delta = 0.9; \\ &\text{The dynamic policy is better when } \alpha \geq 0.34. \end{aligned} \tag{12}$$

Similar computation indicates that for a discount factor $\delta = 0.5$, the dynamic policy is better when $\alpha \geq 0.44$. Note that in this example, we just show the energy efficiency of one of the many possible dynamic policies, and have already seen the advantage of dynamic policies. We will construct the optimal dynamic policy in Section IV, whose performance will be evaluated under different system parameters in Section V.

Even if a dynamic policy is already energy-efficient, a user may want to deviate from it to achieve higher energy efficiency when the user has a high throughput requirement or a low cross channel gain from another user's transmitter to its own receiver. If it has a high throughput requirement, the user may need a large transmit power in its time slot even with no multi-user interference. Hence, it may benefit from transmitting at certain power level in another user's time slot in order to achieve certain throughput and to greatly reduce the power level in its own time slot. We give the conditions under which it is beneficial for a user to deviate from a given policy in the following lemma.

Lemma 1: Suppose that under a given dynamic policy, user i transmits at power level p_i at time t and user j transmits at power level p_j at time $t+s$, where $t \geq 0$ and $s \geq -t$. User j will deviate by transmitting in both time slot t and $t+s$ to achieve at least the same throughput with a lower average power, if and only if the following conditions hold:

$$p_j^{t+s} = \frac{1 + g_{jj}/n_j \cdot p_j}{g_{jj}/n_j} \left((1 + \frac{g_{jj}}{n_j} \cdot p_j) \cdot \frac{g_{jj}}{g_{ji} p_i + n_j} \right)^{-\frac{1}{\delta^s + 1}} - \frac{n_j}{g_{jj}} \leq p_j, \tag{13}$$

$$p_j^t = \frac{g_{ji} p_i + n_j}{g_{jj}} \left(\left(\frac{1 + g_{jj}/n_j p_j}{1 + g_{jj}/n_j p_j^{t+s}} \right)^{\delta^s} - 1 \right) \leq P_j^{\max}, \tag{14}$$

and

$$p_j^t + \delta^s p_j^{t+s} \leq \delta^s p_j. \tag{15}$$

Proof: See [28, Appendix A]. ■

From the above lemma, we can see that user j has the incentive to deviate when $g_{ji} p_i$ is small, namely the interference from user i is small if user j transmits in user i's time slot, and when p_j is large, namely user j's required throughput in time slot $t+s$ is high.

We illustrate the results in Lemma 1 in Fig. 2. Consider the simple network with two symmetric users discussed above in this section. Fig. 2 shows the range of minimum throughput and cross interference levels under which it is beneficial for at least one user to deviate from the simple dynamic policy described above. The discount factor is $\delta = 0.5$. We can see that in this setting, at least one user has incentive to deviate under a wide range of parameter values. Hence, it is important to design deviation-proof spectrum sharing policies, considering the users' inability to perfectly monitor the spectrum usage.

IV. A Design Framework For Spectrum and Energy Efficient Policies

In this section, we first formulate the policy design problem for spectrum and energy efficient spectrum sharing and outline our design framework to solve it. Then we show in detail how to solve the design problem for the optimal spectrum sharing policy and how to implement the optimal policy.

A. Formulation of The Design Problem

The goal of the designer is to come up with a deviation-proof spectrum sharing policy that fulfills all the users' minimum throughput requirements and optimizes certain energy efficiency criterion. The energy efficiency criterion can be represented by a function defined on the average power of all the users, $E(P_1(\pi), \ldots, P_N(\pi))$. An example of energy efficiency criterion can be the average power consumption of all the users, i.e. $E(P_1(\pi), \ldots, P_N(\pi)) = \frac{1}{N}\sum_{i\in\mathcal{N}} P_i(\pi)$. To sum up, we can formally define the policy design problem as follows

$$\begin{aligned} \min_{\pi} \quad & E(P_1(\pi), \ldots, P_N(\pi)) \\ s.t. \quad & \pi \text{ is deviation} - \text{proof}, \\ & R_i(\pi) \geq R_i^{\min}, \ \forall i \in \mathcal{N}. \end{aligned} \tag{16}$$

We outline the proposed design framework to solve the policy design problem, which is a procedure consisting of three phases as illustrated in Fig. 3. In Phase I, the LSS exchanges some information with the users following the procedure described in Table II. In Phase II, using the information obtained in Phase I, the LSS solves the policy design problem for the optimal deviation-proof dynamic spectrum sharing policy. Specifically, it first determines the set of feasible operating points that can be achieved by deviation-proof dynamic policies, and then selects the optimal operating point based on the given energy efficiency criterion. Finally in Phase III, the LSS sends the optimal operating point to the users, as an input to each user's decentralized implementation of the optimal deviation-proof policy.

Note that the information exchange in Phase I is necessary for the LSS to determine and for the users to achieve the optimal operating point. A similar information exchange phase is proposed in [19][20]. The information exchange phase can be also considered as a substitute for the convergence process needed by the algorithms in [3]–[5]. In the proposed policy, since the users implement the policy without any information exchange in Phase III, the only information exchange happen in Phase I and at the end of Phase II. The procedure of information exchange in our framework is advantageous in that its duration and the amount of information to exchange are predetermined. On the other hand, the amount of information to exchange in [3]–[5] is proportional to the convergence time of their algorithms, which is generally unbounded.

B. Solving The Policy Design Problem By The LSS

1) Characterizing the set of feasible operating points: The first step in solving the design problem (16) is to characterize the set of feasible operating points that can be achieved by deviation-proof policies.[2] Specifically, an operating point $\bar{\mathbf{r}} = (\bar{r}_1, \ldots, \bar{r}_N)$ is defined as a collection of throughput $\bar{r}_i$ achieved by each user i when it transmits in a dynamic policy. Since we focus on TDMA policies, the transmit power level of each user i can be obtained from the operating point according to

$$p_i^{\mathrm{TDMA}}(\bar{\mathbf{r}}) = \frac{n_i}{g_{ii}} \left(2^{\bar{r}_i} - 1\right). \tag{17}$$

An operating point $\bar{\mathbf{r}}$ is feasible if there exists a deviation-proof policy with power levels $\mathbf{p}^{\mathrm{TDMA}}(\bar{\mathbf{r}})$ that achieves the minimum throughput requirements. As said before, the characterization of the set of feasible operating points requires some information exchange between the LSS and the users in Phase I, which will be described after we state Theorem 1.

Before stating Theorem 1, we define the *benefit from deviation* as follows.

Definition 4 (Benefit from Deviation): We define user j's benefit from deviation from user i's reward maximizing action profile $\tilde{\mathbf{p}}^i$ as

$$b_{ij} = \sup_{p_j \in \mathcal{P}_j, p_j \neq \tilde{p}_j^i} \frac{\rho(y=1|\tilde{\mathbf{p}}^i) - \rho(y=1|p_j, \tilde{\mathbf{p}}_{-j}^i)}{r_j(p_j, \tilde{\mathbf{p}}_{-j}^i)/\bar{r}_j}, \ \forall i \in \mathcal{N}, \ \forall j \neq i, \tag{18}$$

where $\tilde{\mathbf{p}}^i = (p_i^{\mathrm{TDMA}}(\bar{\mathbf{r}}), \mathbf{p}_{-i} = \mathbf{0})$ is the joint power profile when user i transmits.
As we will see in Theorem 1, if the operating point $\bar{\mathbf{r}}$ can be achieved by deviation-proof policies, the benefit from deviation b_{ij} for all i and $j \neq i$ must be strictly smaller than 0. Since the throughput r_j is always larger than 0, $b_{ij} < 0$ is equivalent to $\rho(y=1|p_j, \tilde{\mathbf{p}}_{-j}^i) > \rho(y=1|\tilde{\mathbf{p}}^i)$ for all $p_j \neq \tilde{p}_j^i$, which means that the probability of the outcome $y=1$ that indicates deviation increases when deviation happens. This guarantees that any deviation from $\tilde{\mathbf{p}}^i$ by user j ($\forall j \neq i$) can be statistically identified. We can observe that the benefit from deviation is also related to the throughput user j obtains by deviation, $r_j(p_j, \tilde{\mathbf{p}}_{-j}^i)$. If the throughput obtained by deviation is smaller, the benefit from deviation is smaller.

[2]Note that we are interested in the case when the users are impatient (their discount factor is fixed and smaller than 1), as opposed to the case when the users are patient (their discount factor can be arbitrarily close to 1) in [19][20][23]. A similar characterization with impatient users is provided in [22] under the assumption of *perfect* monitoring. Our result in Theorem 1 is the first analytical characterization for impatient users with imperfect monitoring.

Now we state Theorem 1, which characterizes the set of feasible operating points that can be achieved by deviation-proof policies.

Theorem 1: An operating point $\bar{\mathbf{r}}$ is feasible for the minimum throughput requirements $\{R_i^{\min}\}_{i\in\mathcal{N}}$, and can be achieved by deviation-proof policies, if the following conditions are satisfied:

- Condition 1: benefit from deviation $b_{ij} < 0$ for all i and $j \neq i$.
- Condition 2: for all $i \in \mathcal{N}$, we have

$$\bar{r}_i - r_i(p_i, \tilde{\mathbf{p}}_{-i}^i) + \bar{r}_i \cdot \sum_{j\neq i} \frac{\rho(y=1|\tilde{\mathbf{p}}^i) - \rho(y=1|p_i, \tilde{\mathbf{p}}_{-i}^i)}{-b_{ij}} \geq 0, \ \forall p_i \in \mathcal{P}_i. \tag{19}$$

- Condition 3: the discount factor δ satisfies

$$\delta \geq \underline{\delta} \triangleq \frac{1}{1 + \frac{1-\sum_{i\in\mathcal{N}} \underline{\mu}_i}{N-1+\sum_{i\in\mathcal{N}}\sum_{j\neq i}(-\rho(y_0|\tilde{\mathbf{p}}^i)/b_{ij})}}. \tag{20}$$

- Condition 4: $\sum_{i\in\mathcal{N}} R_i^{\min}/\bar{r}_i = 1$, and $\bar{r}_i \leq R_i^{\min}/\underline{\mu}_i$, where

$$\underline{\mu}_i \triangleq \max_{j\neq i} \frac{1-\rho(y=1|\tilde{\mathbf{p}}^i)}{-b_{ij}}, \tag{21}$$

Proof: See [28, Appendix B]. ■

Theorem 1 first provides the sufficient conditions for the existence of feasible operating points. Condition 1 (respectively, Condition 2) ensures that at the power profile $\tilde{\mathbf{p}}^i$, user j for any $j \neq i$ (respectively, user i) has no incentive to deviate. Condition 4 gives us the lower bound for the discount factor. If any of the above conditions is violated, there is no feasible operating point. When Conditions 1–3 are all satisfied, Condition 4 gives us the set of feasible operating points under given system parameters. We can choose any point satisfying Condition 4 as the deviation-proof operating point.

Information Exchange In Phase I: We describe the information exchange phase for the LSS to identify the set of Pareto optimal equilibrium payoffs. The key quantities needed are $\{\rho(y=1|\tilde{\mathbf{p}}^i)\}_{i=1}^N$, $\{\rho(y=1|p_j, \tilde{\mathbf{p}}_{-j}^i)\}_{i\neq j}$, and $\{b_{ij}\}_{i\neq j}$, which can be obtained at the end of the information exchange phase. We list the information obtained by the LSS and users during this phase in Table II.

2) Selecting the optimal operating point: For a set of throughput requirements $\{R_i^{\min}\}_{i\in\mathcal{N}}$, we write the set of feasible points obtained in Theorem 1 as $\mathcal{B}(\{R_i^{\min}\}_{i\in\mathcal{N}})$. The following proposition formulates the problem of finding the optimal operating point that solves the policy design problem.

Proposition 1: The optimal operating point can be solved by the following optimization problem

$$\begin{aligned} \min_{\bar{\mathbf{r}}} \quad & E(P_1(\bar{\mathbf{r}}), \ldots, P_N(\bar{\mathbf{r}})) \\ s.t. \quad & \bar{\mathbf{r}} \in \mathcal{B}(\{R_i^{\min}\}_{i \in \mathcal{N}}), \end{aligned} \tag{22}$$

where $P_i(\bar{\mathbf{r}}) = \frac{R_i^{\min}}{\bar{r}_i} \cdot p_i^{\text{TDMA}}(\bar{\mathbf{r}})$.

Proof: See [28, Appendix C]. ■

The optimization problem (22) can be solved by the LSS, who has enough computational capability. After obtaining the optimal operating point $\bar{\mathbf{r}}$, the LSS will send it to the users.

C. Constructing The Deviation-Proof Policy

Now we can construct the optimal deviation-proof dynamic policy. The deviation-proof policy can be implemented by each user in a decentralized manner. The algorithm run by user i is described in the algorithm in Table III. Note that no information exchange is required at this phase.

Theorem 2 ensures that if all the users run the algorithm in Table III locally, they will achieve the minimum throughput requirements $\{R_i^{\min}\}_{i \in \mathcal{N}}$, and will have no incentive to deviate.

Theorem 2: For any operating point $\bar{\mathbf{r}} \in \mathcal{B}(\{R_i^{\min}\}_{i \in \mathcal{N}})$, and any discount factor $\delta \geq \underline{\delta}$, the throughput achieved by each user running the algorithm in Table III is $R_i^{\min}$ for each user i, and no user has incentive to deviate.

Proof: See [28, Appendix D]. ■

The intuition of why the algorithm in Table III works is as follows. At each time slot t, each user i calculates the indices of all the users $\alpha_i(t), \forall i \in \mathcal{N}$. The user i^* with the largest index $\alpha_{i^*}(t) = \max_i \alpha_i(t)$ can transmit in time slot t. Normally, if user i^* transmits at the current time slot, its index in the next time slot is very likely to be small, in order to give other users larger opportunities to transmit. However, when the users receive the distress signal that indicates deviation, they calculate the indices in a different way, such that user i^* still has a large index at the next time slot. Hence, a user may not have the incentive to deviate, because it will leads to a smaller opportunity to transmit in the future.

V. 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 $g_{ij} \sim \mathcal{CN}(0, \alpha), \forall i \neq j$, where α is defined as the *cross interference level*. The channel gain from each user to the LSS also satisfies $g_{0i} \sim \mathcal{CN}(0, 1), \forall i$. The interference temperature threshold is $I = 1$ W. The measurement error ε is Gaussian distributed with zeros mean and variance 0.1. The energy efficiency criterion is the average transmit power of each user. The discount factor is 0.9.

We compare the performance of the proposed policy with that of the optimal static policy in [11][17] and the optimal punish-forgive policy in [19]–[22]. In the optimal static policy, each user transmits at a constant power level that is just large enough to fulfill the throughput requirement, given the interference from other users. The optimal static policy is deviation-proof against other static policies. The punish-forgive policies in [20][22] were originally proposed for network utility maximization problems (e.g. maximizing the sum throughput). In our simulation, we adapt the punish-forgive policies to solve the energy efficiency problem in (16). The punish-forgive policies are dynamic policies that have two phases. When the users have not received the distress signal, they transmit at the same power levels as in the proposed policy. When they receive a distress signal that indicates deviation, they are required to switch to a punishment phase of L time slots. In the punishment phase, all the users transmit at the same power levels as in the optimal static policy[3]. In the following simulation results, we always show the performance of the punish-forgive policy with the optimal punishment length. The "grim-trigger" policy used in [19] is a special case of the punish-forgive policy when the punishment length $L = \infty$. Hence, the performance of the punish-forgive policy with the optimal punishment length is better than that of the grim-trigger policy. As discussed before, the punish-forgive policy works well if the users can perfectly monitor the power levels of all the users, because the punishment serves as

[3]Note that in the punish-forgive policies in [20][22], the users transmit at the maximum power levels in the punishment phase, which is the Nash equilibrium. In our setting, transmitting at the power levels in the optimal static policy is the Nash equilibrium.

a threat to deter the users from deviating, and it will never happen in perfect monitoring case if no user deviates. However, when the users have imperfect monitoring ability, the punishment will happen with some positive probability, which decreases all the users' spectrum and energy efficiency.

We first illustrate the users' transmit power levels in different policies in Fig. 4. Consider a simple example of two symmetric users whose minimum throughput requirements are 1 bits/s/Hz. We can see that both users transmit at the same constant power levels in the optimal static policy. In the punish-forgive policy and the proposed policy, both users transmit at lower transmit power levels alternatively before they receive the distress signal at time slot 5. Since a distress signal is broadcast at the time slot in which user 1 is transmitting, it indicates that user 2 may have deviated. In the punish-forgive policy, both users transmit at the power levels in the optimal static policy. On the contrary, in the proposed policy, they still transmit in a TDMA fashion. However, user 1 transmits in the first three time slots after receiving the distress signal, and user 2 has to wait for the opportunity to transmit later.

Then we compare the energy efficiency of the optimal static policy, the optimal punish-forgive policy, and the proposed policy under different cross interference levels in Fig. 5. We consider a network of two users whose minimum throughput requirements are 1 bits/s/Hz. First, notice that the energy efficiency of the proposed policy remains constant under different cross interference levels, while the average transmit power increases with the cross interference level in the other two policies. The proposed policy outperforms the other two policies in medium to high cross interference levels (approximately when $\alpha \geq 0.3$). In the cases of high cross interference levels ($\alpha \geq 1$), there is no static policy that can fulfill the minimum throughput requirements. As a consequence, the punish-forgive policies cannot fulfill the throughput requirements when $\alpha \geq 1$, either.

In Fig. 6, we examine how the performance of these three policies scales with the number of users. The number of users in the network increases, while the minimum throughput requirement for each user remains 1 bits/s/Hz. The cross interference level is $\alpha = 0.2$. We can see that the static and punish-forgive policies are infeasible when there are more than 6 users. In contrast, the proposed policy can accommodate 18 users in the network with each users transmitting at a power level less than 0.8 W.

Fig. 7 shows the joint spectrum and energy efficiency of the three policies. We can see that the optimal static and punish-forgive polices are infeasible when the minimum throughput requirement is larger than 1.6 bits/s/Hz. On the other hand, the proposed policy can achieve a much higher spectrum efficiency (2.5 bits/s/Hz) with a better energy efficiency (0.8 W transmit power). Under the same average transmit power, the proposed policy is always more energy efficient than the other two policies.

VI. Conclusion

In this paper, we studied power control in dynamic spectrum sharing, and proposed a dynamic spectrum sharing policy that allows the users to transmit in a TDMA fashion. The proposed policy achieves higher spectrum efficiency compared to existing static spectrum sharing policies, and is more energy efficient than static policies under the same minimum throughput requirements. The proposed policy is amenable to decentralized implementation and is deviation-proof, in that the users find it in their self-interests to follow the policy. The proposed policy can achieve high spectrum and energy efficiency even under limited and imperfect monitoring, namely the users only observe the binary distress signals that erroneously indicate the violation of the interference temperature threshold. Simulation results validate our analytical results on the policy design and demonstrate the performance gains enabled by the proposed policy.

References

[1] Q. Zhao and B. M. Sadler, "A survey of dynamic spectrum access," *IEEE Signal Process. Mag.*, vol. 24, no. 3, pp. 79–89, May 2007.

[2] M. Chiang, P. Hande, T. Lan, and C. W. Tan, "Power control in wireless cellular networks," *Foundations and Trends in Networking*, vol. 2, no. 4, pp. 381–533, Apr. 2008.

[3] 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.

[4] Y. Xing, C. N. Mathur, M. A. Haleem, R. Chandramouli, and K. P. Subbalakshmi, "Dynamic spectrum access with QoS and interference temperature constraints," *IEEE Trans. Mobile Comput.*, vol. 6, no. 4, pp. 423–433, Apr. 2007.

[5] 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.

[6] 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.

[7] J. Huang, R. A. Berry, and M. L. Honig, "Auction-based spectrum sharing," *Mobile Networks and Applications*, vol. 11, pp. 405–418, 2006.

[8] Y. Xiao, J. Park, and M. van der Schaar, "Intervention in power control games with selfish users," *IEEE J. Sel. Topics Signal Process., Special issue on Game Theory in Signal Processing*, vol. 6, no. 2, pp. 165–179, Apr. 2012.

[9] 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.

[10] F. Meshkati, H. V. Poor, S. C. Schwartz, and N. B. Mandayam, “An energy-efficient approach to power control and receiver design in wireless data networks,” *IEEE Trans. Commun.*, vol. 53, no. 11, pp. 1885–1894, Nov. 2005.

[11] R. D. Yates, “A framework for uplink power control in cellular radio systems,” *IEEE J. Sel. Areas Commun.*, vol. 13, no. 7, pp. 1341–1347, Sep. 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] T. Alpcan, T. Basar, R. Srikant, and E. Altman, “CDMA uplink power control as a noncooperative game,” *Wireless Networks*, vol. 8, pp. 659–670, 2002.

[14] 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.

[15] 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.

[16] 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.

[17] S. Sorooshyari, C. W. Tan, M. Chiang, “Power control for cognitive radio networks: Axioms, algorithms, and analysis,” To appear in *IEEE/ACM Trans. Netw.*, 2011.

[18] 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., Special issue on Game Theory in Signal Processing*, vol. 6, no. 2, pp. 104–116, Apr. 2012.

[19] R. Etkin, A. Parekh, and D. Tse, “Spectrum sharing for unlicensed bands,” *IEEE J. Sel. Areas Commun.*, vol. 25, no. 3, 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] 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.

[24] O. Ileri, D. Samardzija, and N. Mandayam, “Demand responsive pricing and competitive spectrum allocation via a spectrum server,” in *Proc. IEEE DySPAN*, Baltimore, MD, Nov. 2005, pp. 194–202.

[25] T. Kamakaris, M. Buddhikot, and R. Iyer, “A case for coordinated dynamic spectrum access in cellular networks,” in *Proc. IEEE DySPAN*, Baltimore, MD, Nov. 2005, pp. 289–298.

[26] M. Buddhikot and K. Ryan, “Spectrum management in coordinated dynamic spectrum access based cellular networks,” in *Proc. IEEE DySPAN*, Baltimore, MD, Nov. 2005, pp. 299–307.

[27] J. Acharya and R. Yates, “A framework for dynamic spectrum sharing between cognitive radios,” in *Proc. IEEE ICC*, Glasgow, Scotland, Jun. 2007, pp. 5166–5171.

[28] Y. Xiao and M. van der Schaar, "The appendix for: Dynamic Energy-Efficient Spectrum Sharing Among Autonomous Users," Available at: "http://www.ee.ucla.edu/~yxiao".

TABLE I

COMPARISON WITH RELATED WORKS IN DYNAMIC SPECTRUM SHARING.

	[6]	[3]–[5]	[7]–[18]	[19]–[22]	This work
Dynamic	No	No	No	Yes	Yes
Decentralized	No	Yes	Yes	Yes	Yes
Deviation-proof	No	No	Yes (static)	Yes (static, dynamic)	Yes (static, dynamic)
Performance loss (due to imperfect monitoring)	N/A	N/A	Yes	Yes	No

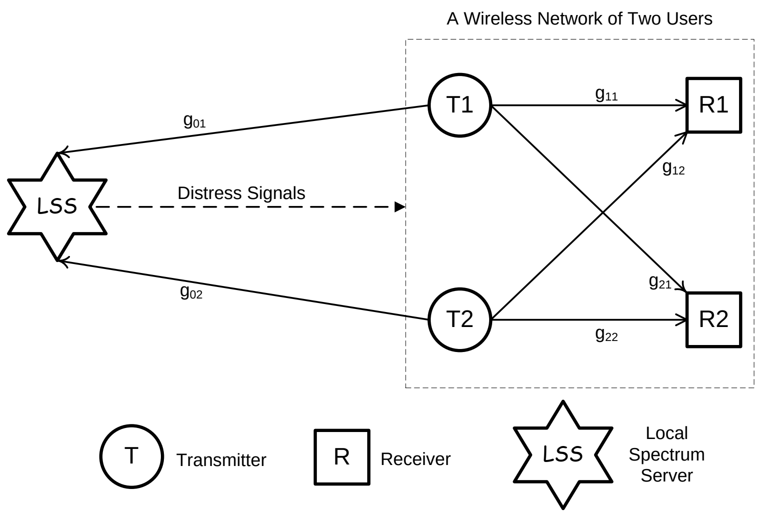

Fig. 1. An example system model with two users. The solid line represents a link for data transmission, and the dashed line indicate a link for control or feedback signals. The channel gains for the corresponding data link are shown in the figure. The local spectrum server (LSS) sends distress signals if the estimated interference power exceeds a threshold.

TABLE II

INFORMATION OBTAINED BY THE LSS AND USERS DURING THE INFORMATION EXCHANGE PHASE.

Events	Information obtained by LSS	Information obtained by user i
Initialization	I (obtained according to $\bar{I}$, $\bar{\Gamma}$)	$\mathcal{P}_i$ (known initially)
Each user i transmits at $P_i^{\max}$	$\{h_{0i}P_i^{\max}\}_{i=1}^N$,$\{\rho(y_0\|\tilde{\mathbf{p}}^i)\}_{i=1}^N$	
LSS broadcasts $\{\rho(y_0\|\tilde{\mathbf{p}}^i)\}_{i=1}^N$, $\min\{1, \frac{\bar{I}}{h_{0i}P_i^{\max}}\}, \forall i$		$\tilde{p}_i^i$, $\{\rho(y_0\|\tilde{\mathbf{p}}^j)\}_{j=1}^N$, $\{h_{ij}\tilde{p}_j^j\}_{j\neq i}$, $\bar{v}_i$
Each user i transmits at $\forall p_i \in \mathcal{P}_i \setminus \{0, P_i^{\max}\}$	$\{\rho(y_0\|\tilde{\mathbf{p}}^j) - \rho(y_0\|p_i, \tilde{\mathbf{p}}_{-i}^j)\}_{j\neq i}$	$u_i(p_j, \tilde{\mathbf{p}}_{-j}^i), \forall j, \forall p_j$
LSS broadcasts $\{\rho(y_0\|\tilde{\mathbf{p}}^j) - \rho(y_0\|p_i, \tilde{\mathbf{p}}_{-i}^j)\}_{j\neq i}$		$\{b_{ki}\}_{k\neq i}$
Each user i broadcasts $\{b_{ki}\}_{k\neq i}$	$b_{ij}, \forall i, \forall j \neq i$	$b_{ij}, \forall i, \forall j \neq i$
Each user i transmits $\bar{v}_i$ to LSS	$\{\bar{v}_i\}_{i=1}^N$	

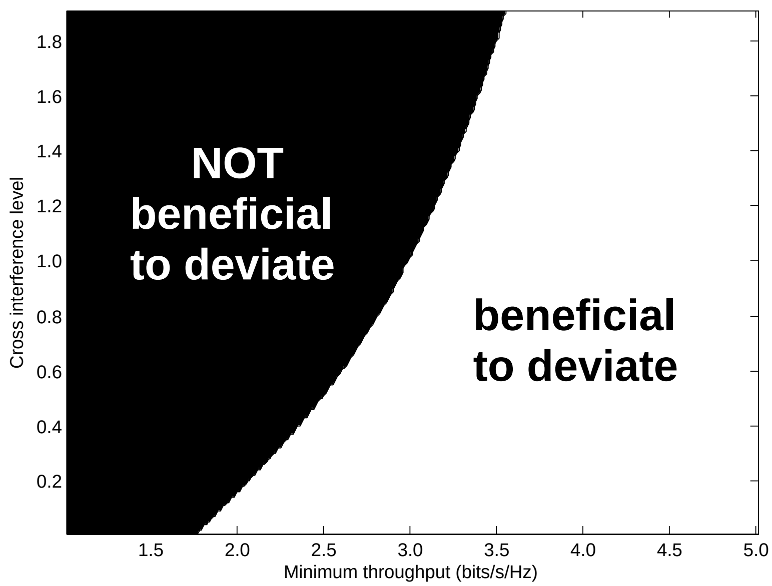

Fig. 2. The system parameters under which it is beneficial for at least one user to deviate.

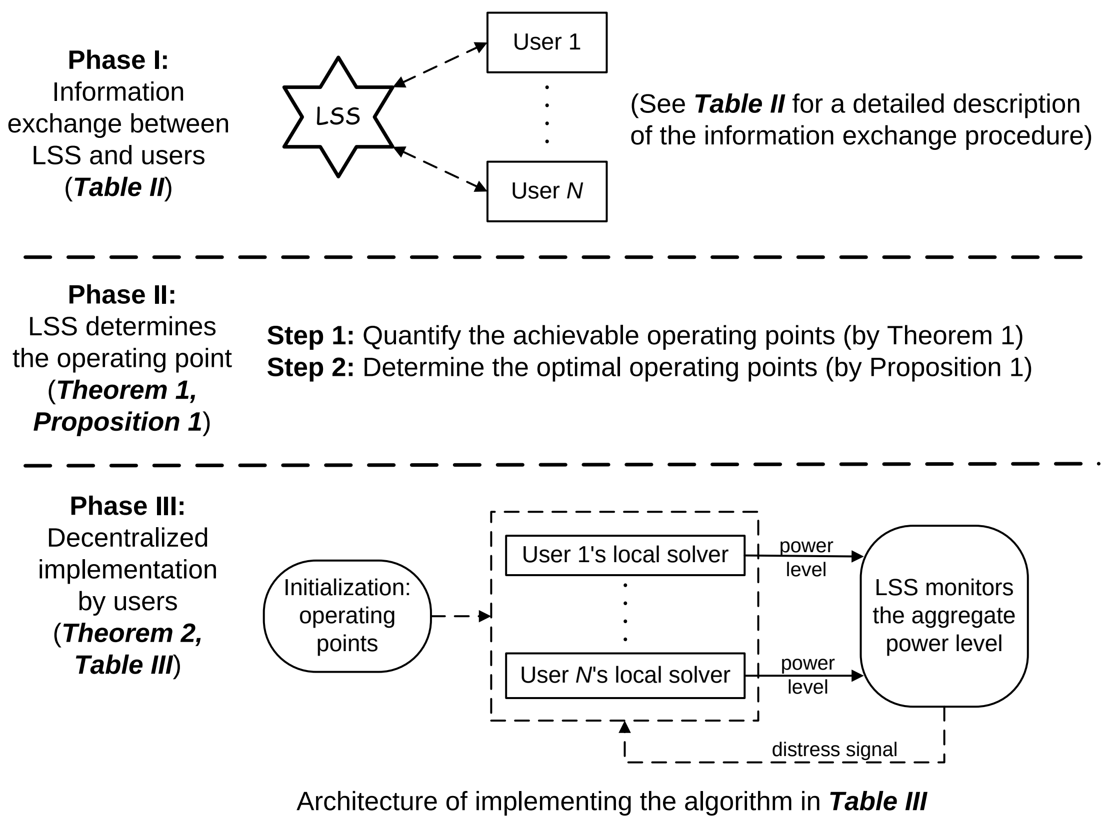

Fig. 3. An illustration of the design framework. In Phase I, the local spectrum server (LSS) and the users exchange information. Based on the information exchanged, LSS determines the optimal operating point in Phase II, and tells the payoff values at the operating points to corresponding users. In Phase III, users implement the spectrum sharing policy in a decentralized way. The key results related to each phase are also listed in the figure.

TABLE III

THE ALGORITHM RUN BY USER i.

Require: The operating point $\bar{\mathbf{r}}$ obtained from the LSS

Initialization: Sets $t = 0$, $r_j(0) = R_j^{\min}$ for all $j \in \mathcal{N}$.

repeat

Calculates the index $\alpha_j(t) = \frac{r_j(t)/\bar{r}_j - \underline{\mu}_j}{1 - r_j(t)/\bar{r}_j + \sum_{k \neq j}(-\rho(y=1|\tilde{\mathbf{p}}^j)/b_{jk})}$ for all $j \in \mathcal{N}$

Finds the largest index $i^* \triangleq \arg\max_{j \in \mathcal{N}} \alpha_j(t)$

if $i = i^*$ **then**

Transmits at the power level $p_i^{\mathrm{TDMA}}(\bar{\mathbf{r}})$

end if

Updates $r_j(t+1)$ for all $j \in \mathcal{N}$

if No Distress Signal Received At Time Slot t **then**

$r_{i^*}(t+1) = \frac{1}{\delta} \cdot r_{i^*}(t) - (\frac{1}{\delta} - 1) \cdot (1 + \sum_{j \neq i^*} \frac{\rho(y=1|\tilde{\mathbf{p}}^{i^*})}{-b_{i^*j}}) \cdot \bar{r}_{i^*}$

$r_j(t+1) = \frac{1}{\delta} \cdot r_j(t) + (\frac{1}{\delta} - 1) \cdot \frac{\rho(y=1|\tilde{\mathbf{p}}^{i^*})}{-b_{i^*j}} \cdot \bar{r}_j$ for all $j \in \mathcal{N}, j \neq i^*$

else

$r_{i^*}(t+1) = \frac{1}{\delta} \cdot r_{i^*}(t) - (\frac{1}{\delta} - 1) \cdot (1 - \sum_{j \neq i^*} \frac{\rho(y=0|\tilde{\mathbf{p}}^{i^*})}{-b_{i^*j}}) \cdot \bar{r}_{i^*}$

$r_j ssss(t+1) = \frac{1}{\delta} \cdot r_j(t) - (\frac{1}{\delta} - 1) \cdot \frac{\rho(y=0|\tilde{\mathbf{p}}^{i^*})}{-b_{i^*j}} \cdot \bar{r}_j$ for all $j \in \mathcal{N}, j \neq i^*$

end if

$t \leftarrow t + 1$

until $\varnothing$

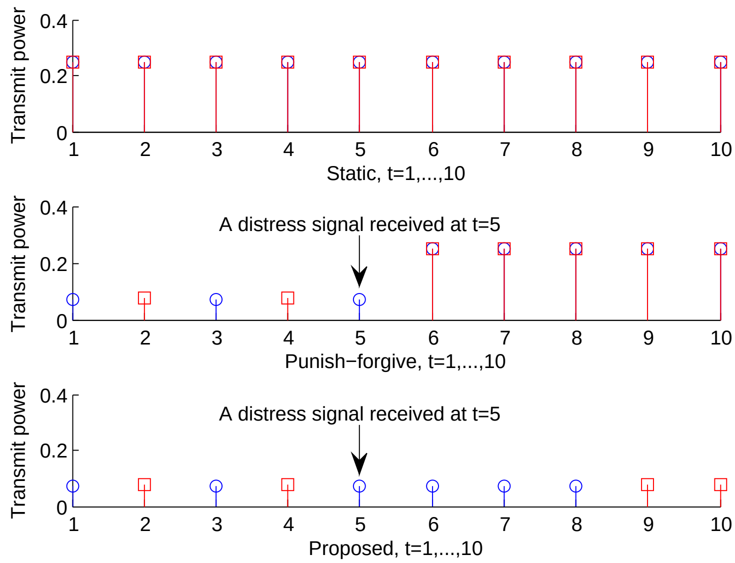

Fig. 4. A snapshot of the two users' transmit power levels under different spectrum sharing policies. User 1's power level is shown as the blue circle, and user 2's power level is shown as the red square.

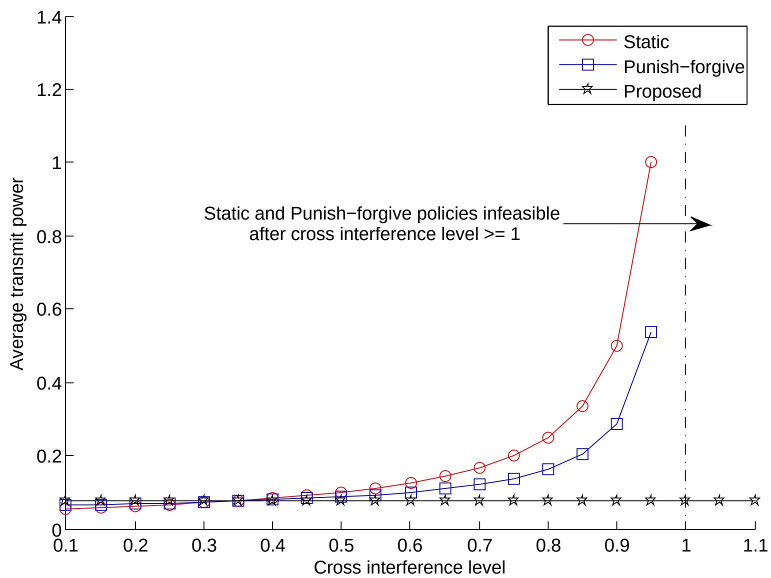

Fig. 5. Energy efficiency of the static, punish-forgive, and proposed policies under different cross interference levels.

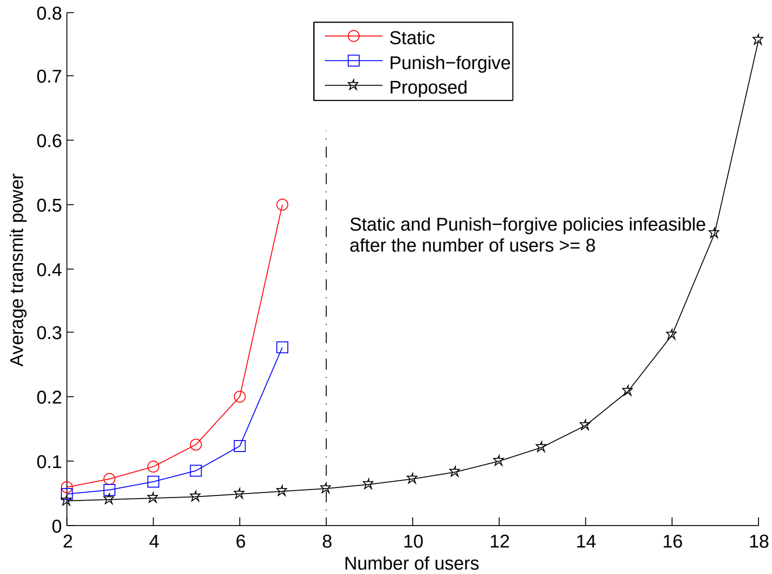

Fig. 6. Energy efficiency of the static, punish-forgive, and proposed policies under different user numbers.

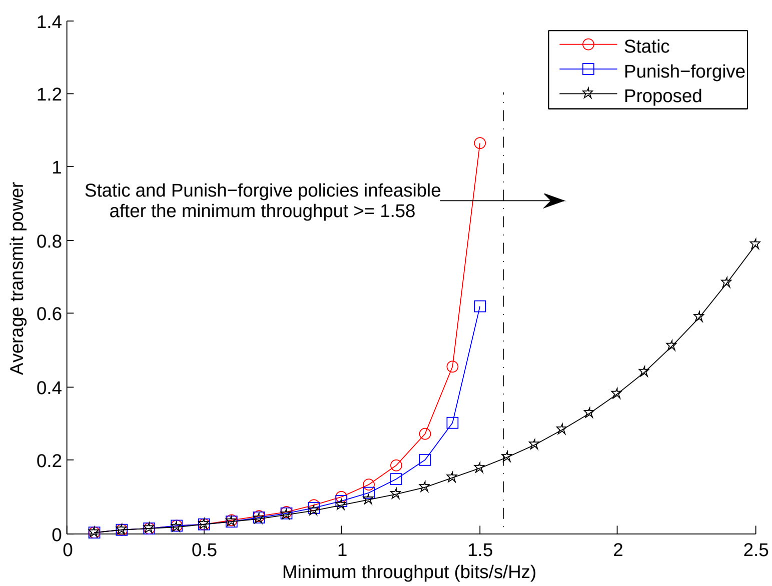

Fig. 7. Energy efficiency of the static, punish-forgive, and proposed policies under different minimum throughput requirements.