Distributed Spectrum Sharing Policies for Selfish Users with Imperfect Monitoring Ability

Yuanzhang Xiao and Mihaela van der Schaar,
Email: {yxiao, mihaela}@ee.ucla.edu

Abstract—**We develop a novel design framework for distributed spectrum sharing among secondary users (SUs), each one of which adjusts its power level to maximize its own payoff (e.g. throughput) while satisfying the interference temperature constraints imposed by primary users. Since the SUs can coexist in the system for a long time, we propose spectrum sharing policies that allow users to transmit in a time-division multiple-access (TDMA) fashion. In the presence of strong multi-user interference, our proposed TDMA policy outperforms existing spectrum sharing policies that dictate users to transmit at constant power levels simultaneously. Our proposed policy achieves Pareto optimality even when the SUs have limited and imperfect monitoring ability: they only observe whether the IT constraints are violated, and their observation is imperfect due to the erroneous measurements. In addition, our policy is deviation-proof, such that the autonomous users will find it in their self-interests to follow the policy. The policy can be implemented by the users in a distributed manner. Simulation results validate our analytical results and quantify the performance gains enabled by the proposed spectrum sharing policies.**

I. Introduction

Cognitive radios have the potential to significantly improve the spectrum efficiency by enabling the secondary users (SUs) to share the spectrum with primary users (PUs), as long as the PUs' quality of service (QoS), such as throughput, is not affected by the SUs [1]. A common approach to guarantee PUs' QoS requirements is to impose *interference temperature* (IT) constraints [1]–[9]; that is, the SUs cannot create an interference level higher than the IT limit set by the PUs. To fulfill the potential of cognitive radios, we need to design spectrum sharing policies that achieve high spectrum efficiency while maintaining the PUs' IT constraints.

The spectrum sharing policy, which specifies the SUs' transmit power levels, is essential to improve spectrum efficiency and protect the PUs' QoS. Since SUs can use the spectrum as long as they do not degrade the PUs' QoS, they can use the spectrum and coexist in the system for long periods of time. However, most existing spectrum sharing policies require the SUs to transmit at *constant* power levels over the time horizon in which they interact[1] [2]–[9]. These policies with constant power levels are inefficient in many spectrum sharing scenarios where the interference among the SUs is strong. Under strong multi-user interference, increasing one user's power level significantly degrades the other users' QoS, which results in a nonconvex feasible QoS region [14]. In this case of nonconvex feasible QoS region, a spectrum sharing policy with constant power levels is inferior to a policy in which the users transmit in a time-division multiple-access (TDMA) fashion. Hence, we focus on TDMA policies in this paper. Note, however, that in the optimal TDMA policy, the users may *not* transmit in a simple round-robin fashion.

[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.

TABLE I
Comparison With Related Works.

	Power levels	Deviation-proof	Monitoring
[2]–[6]	Constant	No	N/A
[7]–[9]	Constant	Yes	N/A
[10]–[13]	Time-varying	Yes	Perfect
Proposed	Time-varying	Yes	Imperfect

Another important feature neglected in the design of spectrum sharing policies in recent works [2]–[6] is the selfishness of SUs, who aim to maximize their own QoS and may deviate from the prescribed spectrum sharing policy. Hence, the spectrum sharing policy should be *deviation-proof*, such that selfish SUs cannot improve their QoS by deviating from the policy.

Deviation-proof spectrum sharing policies with time-varying power levels were studied in [10]–[13], under the assumption of *perfect* monitoring, namely the assumption that each SU can perfectly monitor the individual transmit power levels of all the other SUs. In the policies in [10]–[13], when a deviation from the prescribed policy by any user is detected, a punishment phase will be triggered. In the punishment phase, all the users transmit at the maximum power levels to create strong interference to each other, resulting in low QoS of all the users as a punishment. Due to the threat of this punishment, all the users will follow the policy in their self-interests. However, since the monitoring can never be perfect, the punishment phase, in which all the users receive low throughput, will be triggered even if no one deviates. Thus, the users' average payoffs cannot be Pareto optimal because of the low payoffs received in the punishment phases. Hence, the policies in [10]–[13] must have performance loss in practice where the monitoring is always imperfect.

In this paper, we design TDMA spectrum sharing policies to achieve Pareto optimal operating points that are not achievable by existing policies with constant power levels [2]–[9]. Our proposed policy is deviation-proof, and can achieve Pareto optimality, even when the SUs have *limited* and *imperfect* monitoring ability. Specifically, their monitoring ability can be limited in that they can only distinguish whether the IT

978-1-4673-5051-8/12/$31.00 ©2012 IEEE

constraint is violated, and their monitoring can be imperfect due to the erroneous measurements of the interference temperature.[2] The proposed policy can be easily implemented in a distributed manner.

Finally, we summarize the comparison of our work with the existing works in dynamic spectrum sharing in Table I. We distinguish our work from existing works in the following categories: the power levels prescribed by the spectrum sharing policy are constant or time-varying, whether the policy is deviation-proof or not, and what are the requirements on the SUs' monitoring ability. The "monitoring" category is only discussed within the works based on repeated games.

The rest of the paper is organized as follows. In Section II, we describe the system model and formulate the policy design problem. Then we solve the policy design problem in Section III. Simulation results are presented in Section IV. Finally, Section V concludes the paper.

II. System Model and Problem Formulation

A. System Model For Dynamic Spectrum Sharing

We consider a system with one PU and N SUs (see Fig 1 for an illustrating example of a system with two SUs). The set of SUs is denoted by $\mathcal{N} \triangleq \{1, 2, \ldots, N\}$. Each SU has a transmitter and a receiver. The channel gain from SU i's transmitter to SU j's receiver is g_{ij}. Each SU i chooses a power level p_i from a finite set $\mathcal{P}_i$. We assume that $0 \in \mathcal{P}_i$, namely SU i can choose not to transmit. The set of joint power profiles is denoted by $\mathcal{P} = \prod_{i\in\mathcal{N}} \mathcal{P}_i$, and the joint power profile of all the SUs is denoted by $\mathbf{p} = (p_1, \ldots, p_N) \in \mathcal{P}$. Let $\mathbf{p}_{-i}$ be the power profile of all the SUs other than SU i. Each SU i's instantaneous payoff is a function of the joint power profile, namely $u_i : \mathcal{P} \to \mathbb{R}_+$. Each SU i's payoff $u_i(\mathbf{p})$ is decreasing in the other SUs' power levels p_j, $\forall j \neq i$. Note that we do *not* assume that $u_i(\mathbf{p})$ is increasing in p_i.[3] But we do assume that $u_i(\mathbf{p}) = 0$ if $p_i = 0$, because a SU's payoff should be zero when it does not transmit. One example of many possible payoff functions is the SU's throughput:

$$u_i(\mathbf{p}) = \log_2\Big(1 + \frac{p_i g_{ii}}{\sum_{j\in\mathcal{N}, j\neq i} p_j g_{ji} + n_i}\Big), \tag{1}$$

where n_i is the noise power at SU i's receiver.

As in [4]–[7], there is a local spectrum server (LSS), a device deployed by the PU or the spectrum manager in the local geographic area, serving as a mediating entity among the SUs. The LSS has a receiver to measure the interference temperature and a transmitter to broadcast signals, but it cannot control the actions of the autonomous SUs. The LSS measures the interference temperature at its receiver imperfectly. The measurement can be written as $\sum_{i\in\mathcal{N}} p_i g_{i0} + \varepsilon$, where g_{i0} is the channel gain from SU 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 f_ε known to the LSS. We assume as in most existing works (e.g. [2]–[9]) that the IT limit $\bar{I}$ set by the PU is known perfectly by the LSS. Although the LSS aims to keep the interference temperature below the IT limit $\bar{I}$, it will set a lower intermediate IT limit $I \leq \bar{I}$ to be conservative because of measurement errors. Hence, the IT constraint imposed on the SUs by the LSS is

$$\textstyle\sum_{i\in\mathcal{N}} p_i g_{i0} \leq I. \tag{2}$$

Even if the actual interference temperature $\sum_{i\in\mathcal{N}} p_i g_{i0}$ does not exceed the intermediate IT limit I, the erroneous measurement $\sum_{i\in\mathcal{N}} p_i g_{i0} + \varepsilon$ may still exceed the IT limit $\bar{I}$ set by the PU. In this case, the LSS will broadcast a distress signal to all the SUs. Given the joint power profile $\mathbf{p}$, this false alarm probability is

$$\Gamma(\mathbf{p}) = \Pr\left(\textstyle\sum_{i\in\mathcal{N}} p_i g_{i0} + \varepsilon > \bar{I} \mid \sum_{i\in\mathcal{N}} p_i g_{i0} \leq I\right), \tag{3}$$

where $\Pr(A)$ is the probability that the event A happens.

A SU's payoff is affected by the multi-user interference, which depends on the cross channel gains among the SUs. When the multi-user interference is weak due to small cross channel gains, power control becomes less important, since one SU's power level does not affect the others' payoffs. Hence, in this paper, we focus on the more interesting scenario when the multi-user interference is strong and power control is essential for efficient interference management. We quantify the strength of multi-user interference as follows. First, we write $\tilde{\mathbf{p}}^i = (\tilde{p}_1^i, \ldots, \tilde{p}_N^i)$ as the joint power profile that maximizes SU i's payoff subject to the IT constraint, namely

$$\tilde{\mathbf{p}}^i = \arg\max\nolimits_{\mathbf{p}\in\mathcal{P}} u_i(\mathbf{p}), \text{ subject to } \textstyle\sum_{i\in\mathcal{N}} p_i g_{i0} \leq I. \tag{4}$$

Since u_i is decreasing in $p_j, \forall j \neq i$, we have $\tilde{p}_j^i = 0$, $\forall j \neq i$. For notational simplicity, we define the maximum payoff achievable by SU i as $\bar{v}_i \triangleq u_i(\tilde{\mathbf{p}}^i)$. Then, we say a spectrum sharing scenario has strong multi-user interference if the following property is satisfied.

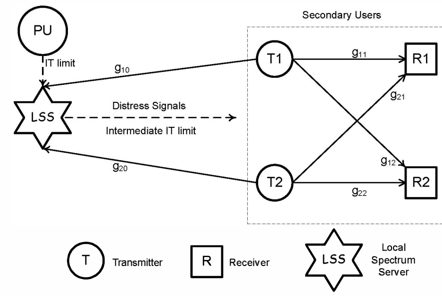

Fig. 1. An example system model with two secondary users. The solid line represents a link for data transmission, and the dashed line indicate a link for control signals. The primary user specifies the interference temperature (IT) limit $\bar{I}$ to the local spectrum server (LSS). The LSS sets the intermediate IT limit I to the secondary users and sends distress signals if the estimated interference temperature exceeds the IT limit $\bar{I}$.

[2]As will be described later in this paper, there is an entity that regulates the interference temperature in the system, who measures the interference temperature imperfectly and feedbacks to the users a binary signal indicating whether the constraints are violated.

[3]In some scenarios with energy efficiency considerations, the payoff is defined as the ratio of throughput to transmit power, which may not monotonically increase with the transmit power.

Definition 1 (Strong Multi-user Interference): A spectrum sharing scenario has strong multi-user interference, if the set of feasible payoffs $\mathcal{V} = \text{conv}\{\mathbf{u}(\mathbf{p}) = (u_1(\mathbf{p}), \ldots, u_N(\mathbf{p})) : \mathbf{p} \in \mathcal{P}, \sum_{i\in\mathcal{N}} p_i g_{i0} \leq I\}$, where $\text{conv}(X)$ is the convex hull of X, has $N+1$ extremal points[4]: $(0,\ldots,0) \in \mathbb{R}^N$, $\mathbf{u}(\tilde{\mathbf{p}}^1), \ldots, \mathbf{u}(\tilde{\mathbf{p}}^N)$.

A spectrum sharing scenario satisfies the above property when the cross channel gains among users are large [14]. According to the definition, the set of feasible payoffs can be written as $\mathcal{V} = \text{conv}\{(0,\ldots,0), \mathbf{u}(\tilde{\mathbf{p}}^1), \ldots, \mathbf{u}(\tilde{\mathbf{p}}^N)\}$. Moreover, its Pareto boundary is part of a hyperplane $\mathcal{B} = \{\mathbf{v} \in \mathcal{V} : \sum_{i=1}^N v_i/\bar{v}_i = 1,\ v_i \geq 0, \forall i\}$, which can be achieved only by SUs transmitting in a TDMA fashion.

B. Deviation-proof Spectrum Sharing Policies

Similar to [2]–[9], we assume that the system parameters, such as the number of SUs and the channel gains, remain fixed during the considered time horizon. The system is time slotted at $t = 0, 1, \ldots$. We assume that the users are synchronized as in [2]–[9]. At the beginning of each time slot t, each SU i chooses its power level p_i^t, and receives a payoff $u_i(\mathbf{p}^t)$. The LSS obtains the measurement $\sum_{i\in\mathcal{N}} p_i^t g_{i0} + \varepsilon^t$, where ε^t is the realization of the error ε at time slot t, and compare the measurement with the IT limit $\bar{I}$. Define $y^t \in Y \triangleq \{y_0, y_1\}$ as the measurement outcome at t, which is determined by

$$y^t = \begin{cases} y_0, & \text{if } \sum_{i\in\mathcal{N}} p_i^t g_{i0} + \varepsilon^t > \bar{I} \\ y_1, & \text{otherwise} \end{cases}. \quad (5)$$

Write the conditional probability distribution of the outcome y^t given the power profile $\mathbf{p}$ as $\rho(y|\mathbf{p})$, which is calculated as

$$\rho(y_1|\mathbf{p}) = \int_{x \leq \bar{I} - \sum_i p_i g_{i\bullet}} f_\varepsilon(x)\, dx, \rho(y_0|\mathbf{p}) = 1 - \rho(y_1|\mathbf{p}).$$

At the end of time slot t, the LSS sends a distress signal if the outcome $y^t = y_0$. Note that the LSS does not send signals when the outcome is y_1, and the SUs know that the outcome is y_1 by default when they do not receive the distress signal.

At each time slot t, each SU i determines its transmit power p_i^t based on the collection of all the past measurement outcomes. Formally, we define the collection of all the past measurement outcomes at time t as $h^t = \{y^0; \ldots; y^{t-1}\} \in Y^t$ for $t \geq 1$ and $h^0 = \varnothing$ for $t = 0$. Then each SU i's strategy σ_i can be formally defined as a mapping $\sigma_i : \sqcup_{t=0}^\infty Y^t \to \mathcal{P}_i$. The spectrum sharing policy is the joint strategy profile of all the SUs, defined as $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_N)$. We write the joint strategy profile of all the SUs other than SU i as $\boldsymbol{\sigma}_{-i}$.

Each SU i aims to maximize its own long-term discounted payoff. Assuming, as in [10]–[13], the same discount factor $\delta \in [0, 1)$ for all the SUs, each SU i's (long-term discounted) payoff is

$$U_i(\boldsymbol{\sigma}) = (1-\delta)\left[u_i(\mathbf{p}^0) + \sum_{t=1}^{\infty} \delta^t \cdot \sum_{y^{t-1}\in Y} \rho(y^{t-1}|\mathbf{p}^{t-1}) u_i(\mathbf{p}^t)\right],$$

[4]The extremal points of a convex set are those that are not convex combinations of other points in the set.

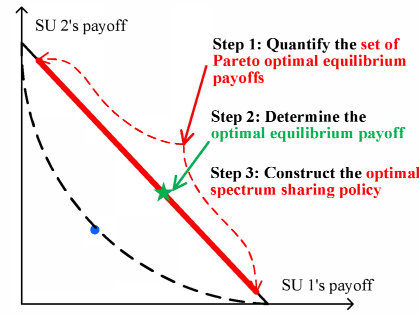

Fig. 2. The procedure of solving the design problem.

where $\mathbf{p}^0$ is determined by $\mathbf{p}^0 = \boldsymbol{\sigma}(\varnothing)$, and $\mathbf{p}^t$ for $t \geq 1$ is determined by $\mathbf{p}^t = \boldsymbol{\sigma}(h^t) = \boldsymbol{\sigma}(h^{t-1}; y^{t-1})$.

We say a spectrum sharing policy is deviation-proof, if no user can obtain a higher payoff by deviating from the policy. The formal definition is given below.

Definition 2 (Deviation-proof Policies): A spectrum sharing policy $\boldsymbol{\sigma}$ is deviation-proof, if for any $i \in \mathcal{N}$,

$$U_i(\boldsymbol{\sigma}) \geq U_i(\sigma_i', \boldsymbol{\sigma}_{-i}),\ \text{for all } \sigma_i'. \quad (6)$$

We define an *equilibrium payoff* as a vector of payoffs $\boldsymbol{v} = (U_1(\boldsymbol{\sigma}), \ldots, U_N(\boldsymbol{\sigma}))$ achieved by a deviation-proof policy $\boldsymbol{\sigma}$.

C. The Policy Design Problem

We want to maximize a welfare function defined on the SUs' payoffs, $W(U_1(\boldsymbol{\sigma}), \ldots, U_N(\boldsymbol{\sigma}))$. This definition of the welfare function is general enough to include the objective functions deployed in many existing works [2]–[13] as special cases. An example welfare function is the average payoff $\sum_{i=1}^N \frac{1}{N} U_i$. It is important to maintain the IT constraint (2). To reduce the cost of sending distress signals, a constraint on the false alarm probability is also imposed as $\Gamma(\mathbf{p}) \leq \bar{\Gamma}$, where $\bar{\Gamma}$ is the maximum false alarm probability allowed. At the maximum of the welfare function, some SUs may have extremely low payoffs. To avoid this, a minimum payoff guarantee $\gamma_i \geq 0$ is imposed for each SU i. To sum up, we can formally define the policy design problem as follows

$$\begin{aligned} \max_{\boldsymbol{\sigma}} \quad & W(U_1(\boldsymbol{\sigma}), \ldots, U_N(\boldsymbol{\sigma})) \quad (7) \\ s.t. \quad & \boldsymbol{\sigma} \text{ is deviation} - \text{proof}, \\ & \sum_{i\in\mathcal{N}} \sigma_i(h^t) \cdot g_{i0} \leq I,\ \forall t,\ \forall h^t \in Y^t, \\ & \Gamma(\boldsymbol{\sigma}(h^t)) \leq \bar{\Gamma},\ \forall t,\ \forall h^t \in Y^t, \\ & U_i(\boldsymbol{\sigma}) \geq \gamma_i,\ \forall i \in \mathcal{N}. \end{aligned}$$

III. Solving The Policy Design Problem

In this section, we solve the policy design problem (7) following the procedure outlined in Fig. 2. We first quantify the set of Pareto optimal equilibrium payoffs (i.e. the Pareto optimal payoffs that can be achieved by deviation-proof policies), then determine the optimal equilibrium payoff based on the welfare function, and finally construct the deviation-proof policy to achieve the optimal equilibrium payoff.

A. Quantify The Set of Pareto Optimal Equilibrium Payoffs

The first step in solving the design problem (7) is to characterize the set of Pareto optimal equilibrium payoffs for the dynamic spectrum sharing system. For the spectrum sharing systems with strong multi-user interference, recall from Definition 1 that the set of feasible payoffs can be written as $\mathcal{V} = \mathrm{conv}\{(0,\ldots,0), \mathbf{u}(\tilde{\mathbf{p}}^1), \ldots, \mathbf{u}(\tilde{\mathbf{p}}^N)\}$, and that its Pareto boundary is $\mathcal{B} = \{\mathbf{v} : \sum_{i=1}^N v_i/\bar{v}_i = 1,\ v_i \geq 0, \forall i\}$. Now we need to determine which portion of the Pareto boundary $\mathcal{B}$ can be achieved as equilibrium payoffs.

We first define the *benefit from deviation* as follows.

Definition 3 (Benefit From Deviation): We define SU j's benefit from deviation from SU i's payoff maximizing power profile $\tilde{\mathbf{p}}^i$ as $b_{ij} = \max_{p_j \in \mathcal{P}_j, p_j \neq \tilde{p}_j^i} \frac{\rho(y_\bullet|\tilde{\mathbf{p}}^i) - \rho(y_\bullet|p_j, \tilde{\mathbf{p}}_{-j}^i)}{u_j(p_j, \tilde{\mathbf{p}}_{-j}^i)/\bar{v}_j}$.

Now we state Theorem 1, which analytically quantifies the set of Pareto optimal equilibrium payoffs.

Theorem 1: We can achieve the following set of Pareto optimal equilibrium payoffs

$$\mathcal{B}_{\underline{\mu}} = \left\{ \boldsymbol{v} : \textstyle\sum_{i=1}^N \frac{v_i}{\bar{v}_i} = 1,\ \frac{v_i}{\bar{v}_i} \geq \underline{\mu}_i, \forall i \in \mathcal{N} \right\}, \quad (8)$$

where $\underline{\mu}_i \triangleq \max_{j \neq i} \frac{1-\rho(y_0|\tilde{\mathbf{p}}^j)}{-b_{ji}}$, if and only if the following conditions are satisfied:

- Condition 1: benefit from deviation $b_{ij} < 0, \forall i, \forall j \neq i$;
- Condition 2: no incentive for SU i, $\forall i \in \mathcal{N}$, to deviate:

$$1 - \frac{u_i(p_i, \tilde{\mathbf{p}}_{-i}^i)}{\bar{v}_i} \geq -\sum_{j \neq i} \frac{\rho(y_\bullet|\tilde{\mathbf{p}}^i) - \rho(y_\bullet|p_i, \tilde{\mathbf{p}}_{-i}^i)}{-b_{ij}}, \forall p_i,$$

- Condition 3: $\delta \geq 1 \Big/ \left\{ 1 + \frac{1 - \sum_i \underline{\mu}_i}{N - 1 + \sum_i \sum_{j \neq i} (-\rho(y_\bullet|\tilde{\mathbf{p}}^i)/b_{ij})} \right\}$.

Proof: See [15, Appendix A]. ■

Theorem 1 provides the sufficient and necessary conditions for the existence of Pareto optimal equilibrium payoffs. Condition 1 (respectively, Condition 2) ensures that at the power profile $\tilde{\mathbf{p}}^i$, SU j for any $j \neq i$ (respectively, SU i) has no incentive to deviate. Condition 3 gives us the requirement of the discount factor. When the conditions are satisfied, Theorem 1 quantifies the set of Pareto optimal equilibrium payoffs $\mathcal{B}_{\underline{\mu}}$. Note that the set of Pareto optimal equilibrium payoffs $\mathcal{B}_{\underline{\mu}}$ is nonempty if and only if $\sum_{i \in \mathcal{N}} \underline{\mu}_i \leq 1$.

B. Determine The Optimal Equilibrium Payoff

Since we have identified the set of Pareto optimal equilibrium payoffs $\mathcal{B}_{\underline{\mu}}$, the problem of find the optimal equilibrium payoff can be written as

$$\max_{\boldsymbol{v}} \quad W(v_1, \ldots, v_N) \ s.t.\ \boldsymbol{v} \in \mathcal{B}_{\underline{\mu}},\ v_i \geq \gamma_i,\ \forall i \in \mathcal{N}. \quad (9)$$

The constraints in the above problem can be further simplified as linear inequalities $v_i \geq \max\{\underline{\mu}_i \cdot \bar{v}_i, \gamma_i\}$, $\forall i \in \mathcal{N}$. Hence, the optimization problem (9) is easy to solve when W is a convex function in $(v_1, \ldots, v_N)$.

TABLE II
THE ALGORITHM RUN BY USER i.

Input: Normalized target payoffs $\{v_i^\star/\bar{v}_i\}_{i \in \mathcal{N}}$
Initialization: Set $t = 0$, $v_j'(0) = v_j^\star/\bar{v}_j$ for all $j \in \mathcal{N}$.
repeat
Calculates the index $\alpha_j(t) = \frac{v_j'(t) - \underline{\mu}_j}{1 - v_j'(t) + \sum_{k \neq j} (-\rho(y_0|\tilde{\mathbf{p}}^j)/b_{jk})}, \forall j$
Finds the largest index $i^* \triangleq \arg\max_{j \in \mathcal{N}} \alpha_j(t)$
if $i = i^*$ **then**
Transmits at the power level $\tilde{p}_i^i$
end if
Updates $v_j'(t+1)$ for all $j \in \mathcal{N}$ as follows:
if No Distress Signal Received At Time Slot t ($y^t = y_1$) **then**
$v_{i^*}'(t+1) = \frac{1}{\delta} \cdot v_{i^*}'(t) - (\frac{1}{\delta} - 1) \cdot (1 + \sum_{j \neq i^*} \frac{\rho(y_\bullet|\tilde{\mathbf{p}}^{i^*})}{-b_{i^*j}})$
$v_j'(t+1) = \frac{1}{\delta} \cdot v_j'(t) + (\frac{1}{\delta} - 1) \cdot \frac{\rho(y_\bullet|\tilde{\mathbf{p}}^{i^*})}{-b_{i^*j}}, \forall j \in \mathcal{N}, j \neq i^*$
else
$v_{i^*}'(t+1) = \frac{1}{\delta} \cdot v_{i^*}'(t) - (\frac{1}{\delta} - 1) \cdot (1 - \sum_{j \neq i^*} \frac{\rho(y_1|\tilde{\mathbf{p}}^{i^*})}{-b_{i^*j}})$
$v_j'(t+1) = \frac{1}{\delta} \cdot v_j'(t) - (\frac{1}{\delta} - 1) \cdot \frac{\rho(y_1|\tilde{\mathbf{p}}^{i^*})}{-b_{i^*j}}, \forall j \in \mathcal{N}, j \neq i^*$
end if
$t \leftarrow t + 1$
until $\varnothing$

C. Construct The Deviation-Proof Policy

Given the optimal payoff $\boldsymbol{v}^\star \in \mathcal{B}_{\underline{\boldsymbol{\mu}}}$, we can construct a deviation-proof policy to achieve it. The deviation-proof policy can be implemented by each SU in a distributed manner. The algorithm run by SU i is described in the algorithm in Table II.

Theorem 2 ensures that if all the SUs run the algorithm in Table II locally, they will achieve the optimal equilibrium payoff $\boldsymbol{v}^\star$, and will have no incentive to deviate.

Theorem 2: If each user i runs the algorithm in Table II, then each user i achieves the optimal payoff $v_i^\star$. The policy implemented by the algorithm is deviation-proof: no user can obtain a higher payoff by deviating from the algorithm.

Proof: See [15, Appendix B]. ■

As we can see from Table II, the computational complexity of implementing the optimal policy is very small. At each period t, each SU only needs to compute N indices $\{\alpha_j(t)\}_{j \in \mathcal{N}}$, and N normalized payoffs $\{v_j'(t)\}_{j \in \mathcal{N}}$, all of which can be calculated analytically. In addition, each SU only needs to know the current measurement outcome and memorize N normalized payoffs $\{v_j'(t)\}_{j \in \mathcal{N}}$. Due to space limitation, we refer interesting readers to [15, Sec. IV-D] for detailed discussions on implementation issues.

IV. Simulation Results

In this section, we demonstrate the performance gain of our spectrum sharing policy over existing policies. Throughout this section, we use the following system parameters. The noise powers at all the SUs' receivers are normalized as 0 dB. The maximum transmit powers of all the SUs are 10 dB. 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 $g_{ij} \sim \mathcal{CN}(0, \beta), \forall i \neq j$, where β is defined as the *cross interference level*. The channel gain from each SU to the LSS also satisfies $g_{i0} \sim \mathcal{CN}(0, 1), \forall i$. The

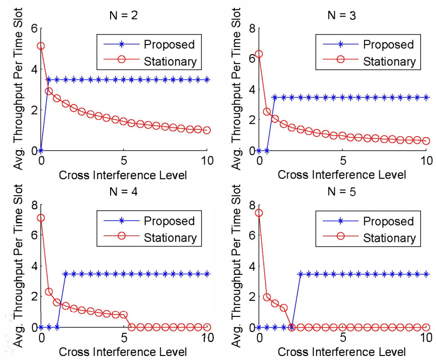

Fig. 3. Performance comparison of the proposed policy and the optimal policy with constant power levels ('stationary' in the legend) under different numbers of users and different cross interference levels.

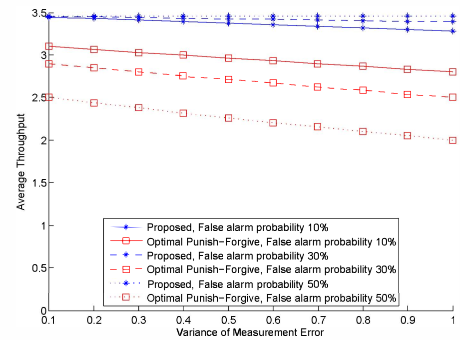

Fig. 4. Performance comparison of the proposed policy and the punish-forgive policy with the optimal punishment length under different error variances and different false alarm probabilities.

IT limit set by the PU is $\bar{I} = 10$ dB. The measurement error ε is Gaussian distributed with zeros mean and variance 0.1. The maximum false alarm probability is $\bar{\Gamma} = 10\%$. The SUs' payoffs are their throughput as in (1). The welfare function is the average throughput, i.e. $W = \sum_{i=1}^{N} \frac{1}{N} U_i$. The minimum payoff guarantee is 10% of the maximum achievable payoff, i.e. $\gamma_i = 0.1 \cdot \bar{v}_i, \forall i$.

In Fig. 3, we compare the average throughput of the proposed policy with that of the optimal policy with constant power levels (referred to as "stationary policy" for simplicity) [3]–[9]. As expected, the proposed policy outperforms the optimal stationary policy in medium to high cross interference levels (approximately when $\beta \geq 1$). When the cross interference level is high ($\beta \geq 2$) and the number of users is large ($N = 5$), the stationary policy fails to meet the minimum payoff guarantees due to strong interference (indicated by zero average throughput in the figure). On the other hand, the desirable feature of the proposed policy is that the average throughput does not decrease with the cross interference level, because SUs transmit in a TDMA fashion. For the same reason, the average throughput does not change with the number of SUs, either.

In Fig. 4, we compare the average throughput of the proposed policy with that of existing policies with time-varying power levels [10]–[13]. Specifically, we consider the "punish-forgive" policy in [10]–[13], which requires SUs to switch to the punishment phase once a deviation is detected. In the punishment phase, all the SUs transmit at the maximum power levels. The performance of punish-forgive polices degrades with the increase of the error variance and the false alarm probability, because of the increasing probability of mistakenly triggered punishments.

V. Conclusion

In this paper, we studied power control in dynamic spectrum sharing among SUs under interference temperature constraints, and proposed a spectrum sharing policy that allows SUs to transmit in a TDMA fashion. The proposed policy can achieve Pareto optimal operating points that are not achievable under existing spectrum sharing policies with constant power levels. The proposed policy is amenable to distributed implementation and is deviation-proof. The proposed policy can achieve Pareto optimality even when the SUs only imperfectly observe whether the interference temperature constraint is violated.

References

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

[2] X. Kang, R. Zhang, Y.-C. Liang, and H. K. Garg, "Optimal power allocation strategies for fading cognitive radio channels with primary user outage constraint," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 2, pp. 374-383, 2011.

[3] 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, 2011.

[4] 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.

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

[6] S. Sharma and D. Teneketzis, "An externalities-based decentralized optimal power allocation algorithm for wireless networks," *IEEE/ACM Trans. Netw.*, vol. 17, no. 6, pp. 1819–1831, 2009.

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

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

[9] S. Sharma and D. Teneketzis, "A game-theoretic approach to decentralized optimal power allocation for cellular networks," *Telecommunication Systems*, pp. 1–16, 2010.

[10] R. Etkin, A. Parekh, and D. Tse, "Spectrum sharing for unlicensed bands," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 3, pp. 517–528, 2007.

[11] 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.

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

[13] Y. Xiao, J. Park, and M. van der Schaar, "Repeated games with intervention: Theory and applications in communications," *IEEE Trans. Commun.*, vol. 60, no. 10, pp. 3123–3132, 2012. Also available at: http://arxiv.org/abs/1111.2456.

[14] S. Stańczak and H. Boche, "On the convexity of feasible QoS regions," *IEEE Trans. Inf. Theory*, vol. 53, no. 2, 2007.

[15] 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, 2012. Also available at: http://arxiv.org/abs/1201.3328.