Interference-Based Pricing for Opportunistic Multi-Carrier Cognitive Radio Systems

Salvatore D'Oro, *Student Member, IEEE*, Panayotis Mertikopoulos, *Member, IEEE*, Aris L. Moustakas, *Senior Member, IEEE*, Sergio Palazzo, *Senior Member, IEEE*

***Abstract*—Cognitive radio (CR) systems allow opportunistic, secondary users (SUs) to access portions of the spectrum that are unused by the network's licensed primary users (PUs), provided that the induced interference does not compromise the PUs' performance guarantees. To account for interference constraints of this type, we consider flexible spectrum access pricing schemes that charge SUs based on the interference that they cause to the system's PUs, and we examine how SUs can react so as to maximize their achievable transmission rate in this setting. We show that the resulting non-cooperative game admits a unique Nash equilibrium under very mild assumptions on the pricing mechanism employed by the network operator, and under both static and ergodic (fast-fading) channel conditions. In addition, we derive a dynamic power allocation policy that converges to equilibrium within a few iterations (even for large numbers of users), and which relies only on local – and possibly imperfect –signal-to-interference-and-noise ratio (SINR) measurements; importantly, the proposed algorithm retains its convergence properties even in the ergodic channel regime, despite its inherent stochasticity. Our theoretical analysis is complemented by extensive numerical simulations which illustrate the performance, robustness and scalability properties of the proposed pricing scheme under realistic network conditions.**

***Index Terms*—Cognitive radio; multi-carrier systems; interference temperature; pricing; exponential learning.**

I. Introduction

Greatly raising the bar from previous generation upgrades, current design specifications for 5th generation (5G) wireless systems target a massive increase in network capacity, fiber-like connection speeds (well into the Gb/s range), and an immersive overall user experience with zero effective latency and response times [1, 2]. As such, the ICT industry is faced with a formidable challenge: these ambitious design goals necessitate the massive deployment of new wireless interfaces but the required overhaul is limited by the inherent constraints of upgrading an entrenched (and often ageing) wireless infrastructure.

This research was supported in part by the European Commission in the framework of the FP7 Network of Excellence in Wireless COMmunications NEWCOM# (contract no. 318306), and by the French National Research Agency project NETLEARN (ANR–13–INFR–004). Part of this work was presented at WiOpt 2014: the 12th International Symposium on Modelling and Optimization in Mobile, Ad Hoc, and Wireless Networks, Hammamet, Tunisia, May 2014.

S. D'Oro and S. Palazzo are with the CNIT Research Unit, University of Catania, Italy; P. Mertikopoulos is with the French National Center for Scientific Research (CNRS) and the Laboratoire d'Informatique de Grenoble, France; A. L. Moustakas is with the Department of Physics, University of Athens, the Institute of Accelerating Systems and Applications (IASA), Athens, Greece, and the École Supérieure d'Électricité (Supélec), Gif-sur-Yvette, France, supported by the Digiteo Senior Chair "ASAPGONE".

Chief among these concerns is the projected spectrum crunch: if not properly managed, the existing radio spectrum will not be able to accommodate the soaring demand for wireless broadband and the ever-growing volume of data traffic [3]. Worse, studies by the US Federal Communications Commission (FCC) and the National Telecommunications and Information Administration (NTIA) have shown that this vital commodity is effectively squandered through underutilization and inefficient use: for instance, only 15% to 85% of the licensed radio spectrum is used on average, leaving ample spectral voids that could be exploited via efficient spectrum management techniques [3, 4]. Accordingly, in this often unregulated context, the emerging paradigm of cognitive radio (CR) has attracted considerable interest as a promising way out of the spectrum gridlock [5–8].

At its most basic level, cognitive radio comprises a two-level hierarchy between wireless users induced by spectrum licencing: the network's licensed, primary users (PUs) have purchased spectrum rights from the network operator (often in the form of contractual quality of service (QoS) guarantees), but they allow unlicenced secondary users (SUs) to access the spectrum provided that the induced co-channel interference (CCI) remains below a certain threshold [5, 7]. Put differently, by sensing the wireless medium and exploiting any spectrum holes that might appear, the network's cognitive SUs essentially free-ride on the PUs' licensed spectrum and try to communicate under the constraints imposed by the PUs (though, of course, without any QoS guarantees). Thus, by opening up the unused part of the spectrum to opportunistic user access, overall spectrum utilization is increased without needing to deploy more (and more expensive) wireless interfaces [6, 9].

Of course, given that PUs utilize the spectrum in a dynamic (and unpredictable) manner, sensing spectrum holes can be a very challenging task [10]. To that end, considerable efforts have been made to develop *spectrum sensing* techniques that enable the optimal allocation of the available spectrum to SUs – e.g. via low-complexity energy detectors [11] or statistical estimation [12]. However, this often requires coordination among SUs, resulting in increased overall complexity and overhead [11]; if SUs do not cooperate, most of the proposed techniques cannot be properly exploited – for a comprehensive account, see e.g. [10, 12, 13] and references therein.

For this reason, throughput maximization in opportunistic, non-cooperative CR environments calls for flexible transmit policies with minimal information exchange between SUs, PUs, and access points/base stations. In particular, a major challenge involves safeguarding the performance guarantees

that the network's licensed PUs have already paid for: if SUs are allowed to transmit without some power/interference control mechanism in place, then the PUs' quality of service requirements may be violated, thus invalidating the fundamental operational premise of CR systems. To that end, the authors of [14–17] investigated the role of pricing as an effective mechanism to control interference and they provided an energy-efficient formulation of the problem where users seek to maximize their transmission rate while keeping their transmit power in check. To reach a stable equilibrium state in this setting, several distributed approaches have been proposed, based chiefly on reaction functions [14], Gauss-Seidel and Jacobi update algorithms [16], or learning [17, 18]; however, these works do not distinguish between licenced and unlicensed users, so their results do not immediately apply to CR networks.

In CR systems, PU requirements are often treated as interference temperature (IT) constraints [19] and the theoretical analysis of the resulting system aims to characterize (and provide the means to converge achieve) the network's optimum/equilibrium states [20–24]. These constraints are often enforced indirectly via exogenous pricing mechanisms that charge SUs based on the aggregate interference that they cause to the network's PUs (and, of course, PUs are reimbursed commensurately). In this context, the authors of [20] introduced a spectrum-trading mechanism based on a market-equilibrium approach [25] and they provided an algorithm allowing SUs to estimate spectrum prices and adjust their spectrum demands accordingly. More recently, the authors of [22, 23] introduced and analyzed a game-theoretic formulation of CR interference channels where SUs are charged proportionally to the aggregate interference caused. A game-theoretic account of the impact of IT constraints on system performance is also studied in [26] where the authors derive cost-aware optimal power allocation policies by relaxing the problem's hard IT constraints and incorporating an exponential cost in the SUs' utility functions. Finally, by exploiting the innate hierarchy between primary and secondary users, the authors of [27] provided a Stackelberg game formulation where the system's PU acts as the leader and seeks to maximize the revenue generated by discriminatory spectrum access pricing mechanisms imposed to SUs (the game's followers).

That being said, the above works focus almost exclusively on wireless systems with static channel conditions where the benefits of interference control mechanisms are relatively easy to evaluate; by contrast, very little is known in the case where the channels vary with time (e.g., due to user mobility) or when the users' measurements are not accurate. In the presence of (fast) fading, channel gains are typically assumed to follow a stationary ergodic process, so the users' throughput and induced interference depend crucially on the channel statistics. In this stochastic framework, the authors of [18] studied the problem of ergodic rate maximization in multi-carrier (MC) systems and derived an efficient power allocation algorithm that allows users to attain the system's capacity; however, no distinction was made between licensed and unlicensed users, so the results of [18] do not readily translate to a CR setting. More recently, [28] provided an efficient online learning algorithm for unilateral rate optimization in dynamic multi-carrier multiple-input and multiple-output (MIMO) cognitive radio systems, but, again, without taking into account any IT constraints imposed by the network's primary users.

In this paper, we consider the problem of cost-efficient throughput maximization in multi-carrier cognitive radio networks where SUs are charged based on the interference that they cause to the network's PUs (either on an aggregate or a per-user basis). Our system model is presented in Section II where we consider a general game-theoretic formulation that is flexible enough to account for both aggregate (flat-rate), temperature-based, and per-user pricing schemes. In the case of static channels (Section III), we show that the resulting game admits a unique Nash equilibrium almost surely, provided that the SUs' pricing schemes satisfy some fairly mild requirements (for instance, that a user's transmission cost increases with his radiated power). On the other hand, in the case of fast-fading channels (which we study in Section IV), we show that the game under study admits a unique Nash equilibrium always, without any further caveats.

Moreover, extending the exponential learning techniques of [18], we also derive a dynamic power allocation policy that converges to Nash equilibrium in a few iterations, even for large numbers of users and/or subcarriers per user. In particular, the proposed learning algorithm has the following desirable attributes:

1) *Distributedness:* user updates are based on local information and signal measurements.
2) *Robustness:* measurements and observations may be subject to random errors and noise.
3) *Statelessness:* users do not need to know the state (or topology) of the system.
4) *Flexibility:* the algorithm can be deployed "as-is" in both static and ergodic environments.

Importantly, the proposed power allocation policy is complementary to the underlying spectrum sensing mechanism used to identify PU activity and spectrum holes; as such, it can be applied to a wide range of power allocation problems and different spectrum sensing approaches.

Our analysis is supplemented in Section V by extensive numerical simulations where we illustrate the throughput and power gains of the proposed approach under realistic conditions.

II. System Model

Consider a set $\mathcal{K} = \{1, \dots, K\}$ of (unlicensed) secondary users (SUs) that seek to connect to a common receiver over a set $\mathcal{S} = \{1, \dots, S\}$ of non-interfering subcarriers (typically in the frequency domain if an orthogonal frequency division multiplexing (OFDM) scheme is employed). Focusing on the uplink case, the aggregate received signal y_s over the s-th subcarrier will then be:

$$y_s = \sum\nolimits_{k \in \mathcal{K}} h_{ks} x_{ks} + z_s, \tag{1}$$

where:

- $x_{ks} \in \mathbb{C}$ denotes the transmitted signal of user $k \in \mathcal{K}$ over the s-th subcarrier.

- $h_{ks} \in \mathbb{C}$ is the corresponding channel coefficient.
- $z_s \in \mathbb{C}$ denotes the aggregate interference-plus-noise received from all sources not in $\mathcal{K}$ (including the aggregate PU transmission on subcarrier s plus ambient and other peripheral interference effects); throughout this paper (and by performing a suitable change of basis if necessary), we will model z_s as a Gaussian variable $z_s \sim \mathcal{CN}(0, \sigma_s^2)$ for some positive $\sigma_s > 0$.

In this context, the average transmit power of user k on subcarrier s will be

$$p_{ks} = \mathbb{E}\left[|x_{ks}|^2\right], \tag{2}$$

where the expectation is taken over the (Gaussian) codebook of user k; furthermore, each user's *total* transmit power $p_k = \mathbb{E}[\mathbf{x}_k^\dagger \mathbf{x}_k] = \sum_s p_{ks}$ must satisfy the power constraint

$$p_k = \sum\nolimits_{s\in\mathcal{S}} p_{ks} \leq P_k, \tag{3}$$

where $P_k > 0$ denotes the maximum transmit power of user $k \in \mathcal{K}$. In this way, the set of admissible power allocation vectors for user k is the S-dimensional polytope

$$\mathcal{X}_k = \left\{\mathbf{p}_k \in \mathbb{R}^{\mathcal{S}} : p_{ks} \geq 0 \text{ and } \textstyle\sum_{s\in\mathcal{S}} p_{ks} \leq P_k\right\}, \tag{4}$$

and the system's *state space* (i.e., the space of all admissible power allocation profiles $\mathbf{p} = (\mathbf{p}_1, \dots, \mathbf{p}_K)$) will be the product $\mathcal{X} = \prod_k \mathcal{X}_k$.

In this multi-carrier (MC) framework, each user's achievable transmission rate depends on his individual signal-to-interference-and-noise ratio (SINR)

$$\gamma_{ks}(\mathbf{p}) = \frac{g_{ks} p_{ks}}{\sigma_s^2 + \sum_{\ell\neq k} g_{\ell s} p_{\ell s}}, \tag{5}$$

where $g_{ks} = |h_{ks}|^2$ denotes the channel gain of user k over the s-th subcarrier. Thus, in the single user decoding (SUD) regime (where interference by all other users is treated as additive noise), the maximum transmission rate for user k (achievable with random Gaussian codes) will be:

$$\begin{aligned} R_k(\mathbf{p}) &= \sum_{s\in\mathcal{S}} \log\left(1 + \gamma_{ks}(\mathbf{p})\right) \\ &= \sum_{s\in\mathcal{S}} \left[\log\left(\sigma_s^2 + w_s(\mathbf{p})\right) - \log\left(\sigma_s^2 + \sum\nolimits_{\ell\neq k} g_{\ell s} p_{\ell s}\right)\right] \end{aligned} \tag{6}$$

where

$$w_s(\mathbf{p}) = \sum\nolimits_k g_{ks} p_{ks}, \quad s = 1, \dots, S, \tag{7}$$

denotes the *aggregate* SU interference level per subcarrier (for convenience we will also write $\mathbf{w} = (w_1, \dots, w_S)$ for the SUs' aggregate interference profile over all subcarriers $s \in \mathcal{S}$).

In the absence of other considerations, the unilateral objective of each SU would be the maximization of his individual transmission rate $R_k(\mathbf{p})$ subject to the total power constraint (3). In our CR setting however, the network operator needs to ensure that the system's PUs meet the QoS guarantees that they have already paid for – typically in the form of minimum rate requirements or maximum interference tolerance per subcarrier. Thus, to achieve this, we will consider a general spectrum access pricing scheme whereby SUs are charged according to the individual and aggregate interference that they cause to the network's PUs. Formally, this can be captured by the general cost model:

$$C_k(\mathbf{p}) = \pi_0(\mathbf{w}(\mathbf{p})) + \pi_k(\mathbf{p}_k), \tag{8}$$

where:

- $\pi_0 \colon \mathbb{R}_+^{\mathcal{S}} \to \mathbb{R}_+$ is a *flat spectrum access price* that is calculated in terms of the aggregate SU interference level w_s per subcarrier $s \in \mathcal{S}$.
- $\pi_k \colon \mathcal{X}_k \to \mathbb{R}_+$ is a *user-specific price* which is charged to user $k \in \mathcal{K}$ based on his *individual* radiated power profile $\mathbf{p}_k \in \mathcal{X}_k$.

In tune with standard economic considerations on diminishing returns [25], the only assumptions that we will make for the price functions π_0 and π_k are that:

(A1) Every price function π is non-decreasing and convex in each of its arguments.

(A2) Every price function π is Lipschitz continuous.

In particular, the convexity assumption (A2) acts as an interference control mechanism: by charging higher spectrum access prices for the same increase in interference when the network operates in a high-interference state, SUs are driven to transmit at lower powers. Additionally, the pricing scheme (8) is flexible enough to account for very diverse pricing paradigms: if $\pi_0 \equiv 0$, every SUs is charged on an equitable, based only on the individual interference induced to the network's PUs;[1] otherwise, if $\pi_k \equiv 0$, the pricing model (8) allows the network operator to reimburse infractions to the PUs' contractual QoS guarantees by imposing an aggregate "sanction" to the network's SUs (who were responsible for causing the violation in the first place).

The specifics of the pricing functions π_0 and π_k are negotiated between network users and operators based on their needs and means, so they can vary widely depending on the context – see e.g. [14, 17, 22, 26, 27]. For concreteness, we provide below some representative examples of pricing models (which we explore further in Section V):[2]

Model 1. Let $I_s^{\max}$ denote the PUs' interference tolerance on subcarrier s. Then, in the spirit of [22], we define the linear pricing (LP) flat-rate model as:

$$\pi_0^{\mathrm{LP}}(\mathbf{w}) = \lambda_0 \sum\nolimits_{s\in\mathcal{S}} w_s / I_s^{\max}, \tag{LP}$$

where the pricing parameter λ_0 represents the price paid by the network's SUs when saturating the PUs' interference tolerance. In words, SUs are charged a flat rate which is proportional to the degree of saturation of the PUs' interference tolerance level.

Model 2. With notation as above, the violation pricing (VP) flat-rate model is defined as:

$$\pi_0^{\mathrm{VP}}(\mathbf{w}) = \lambda_0 \sum\nolimits_{s\in\mathcal{S}} \left[w_s / I_s^{\max} - 1\right]_+ \tag{VP}$$

where $\lambda_0 > 0$ is a sensitivity parameter and $[x]_+ \equiv \max\{x, 0\}$. In this model, SUs are only charged when the PUs' interference tolerance is

[1] Likewise, π_k could also account for the actual cost incurred by the user to recharge the battery of his wireless device as in [17].

[2] For simplicity, we focus on the flat-rate case; the corresponding user-specific price functions π_k are defined similarly.

actually violated, and the steepness of the sanction is controlled by the pricing parameter λ_0; as such, in the large λ_0 limit, (VP) treats the PUs' requirements as a hard constraint with very sharp violation costs.

In light of all this, the *utility* of user k is defined as:

$$u_k(\mathbf{p}) = R_k(\mathbf{p}) - C_k(\mathbf{p}), \tag{9}$$

i.e., $u_k(\mathbf{p})$ is simply the user's achieved transmission rate minus the cost for accessing the system. In turn, this leads to a non-cooperative game $\mathcal{G} \equiv \mathcal{G}(\mathcal{K}, \mathcal{X}, u)$ defined as follows:

1) The game's *players* are the system's secondary users $k \in \mathcal{K} = \{1, \dots, K\}$.
2) The *action set* of each player/user is the set of feasible power allocation profiles $\mathcal{X}_k = \{\mathbf{p}_k \in \mathbb{R}^{\mathcal{S}} : p_{ks} \geq 0 \text{ and } \sum_{s\in\mathcal{S}} p_{ks} \leq P_k\}$.
3) Each player's *utility function* $u_k \colon \mathcal{X} \equiv \prod_k \mathcal{X}_k \to \mathbb{R}$ is given by (9).

In this context, a power allocation profile $p^* \in \mathcal{X}$ is at *Nash equilibrium (NE)* when

$$u_k(\mathbf{p}_k^*; \mathbf{p}_{-k}^*) \geq u_k(\mathbf{p}_k; \mathbf{p}_{-k}^*) \quad \forall \mathbf{p}_k \in \mathcal{X}_k \text{ and } \forall k \in \mathcal{K}, \tag{10}$$

i.e., when *each user's chosen power profile* $p_k^* \in \mathcal{X}_k$ *is individually optimal given the power profile of his opponents* (so no user has a unilateral incentive to deviate). Accordingly, our goal in the rest of the paper will be to characterize the Nash equilibria of $\mathcal{G}$ and to provide distributed optimization methods allowing selfish (and myopic) SUs to converge to equilibrium.

III. Equilibrium Analysis, Learning and Convergence

In this section, we focus on the characterization of the NE of the game $\mathcal{G}$ and on how players can attain such a state by means of a simple, adaptive learning process.

A. Equilibrium structure and characterization

A key property of $\mathcal{G}$ is that it admits a *potential function* [29]:

Proposition 1. *Let* $\mathbf{w}$ *be the aggregate SU interference level defined as in* (7)*. Then, the function*

$$V(\mathbf{p}) = \sum\nolimits_s \log\left(\sigma_s^2 + w_s\right) - \pi_0(\mathbf{w}) - \sum\nolimits_k \pi_k(\mathbf{p}_k) \tag{11}$$

is an exact potential for $\mathcal{G}$*; specifically:*

$$u_k(\mathbf{p}_k; \mathbf{p}_{-k}) - u_k(\mathbf{p}_k'; \mathbf{p}_{-k}) = V(\mathbf{p}_k; \mathbf{p}_{-k}) - V(\mathbf{p}_k'; \mathbf{p}_{-k}) \tag{12}$$

for all $\mathbf{p}_k, \mathbf{p}_k' \in \mathcal{X}_k$ *and for all* $\mathbf{p}_{-k} \in \mathcal{X}_{-k} \equiv \prod_{\ell\neq k} \mathcal{X}_k$.

Proof: By verification. ■

Since the price functions π_0 and π_k are convex, the potential function V is itself concave (though not necessarily strictly so; see below). By Proposition 1, it then follows that maximizers of V are NE of $\mathcal{G}$ (so the Nash set of $\mathcal{G}$ is nonempty); furthermore, with V concave in $\mathbf{p}$ and u_k concave in $\mathbf{p}_k$, every NE of $\mathcal{G}$ is also a maximizer of V. In this way, finding the equilibria of $\mathcal{G}$ boils down to the nonlinear optimization problem:

$$\begin{aligned} &\text{maximize} && V(\mathbf{p}), \\ &\text{subject to} && p_{ks} \geq 0, \quad \sum\nolimits_s p_{ks} \leq P_k. \end{aligned} \tag{13}$$

Thanks to this formulation, we obtain the following equilibrium uniqueness result for $\mathcal{G}$:

Theorem 1. *Assume that:*

(C1) *Each user-specific price function* π_k *is strictly increasing in each of its arguments.*
or:
(C2) *The flat spectrum access price function* π_0 *is either gentle enough or steep enough :* $0 \leq \frac{\partial \pi_0}{\partial w_s} < \left(\sigma_s^2 + \sum_k g_{ks} P_k\right)^{-1}$ *or* $\frac{\partial \pi_0}{\partial w_s} > 1/\sigma_s^2$ *for all* w_s *and for all subcarriers* $s \in \mathcal{S}$.

Then, the game $\mathcal{G}$ *admits a unique Nash equilibrium for almost all realizations of the channel gain coefficients* g_{ks}*; otherwise, if both* (C1) *and* (C2) *fail to hold, the Nash set of* $\mathcal{G}$ *is a convex polytope of dimension at most* $S(K-1)$.

Proof: See Appendix A. ■

Remark 1. The "almost all" part of the statement of Theorem 1 should be interpreted with respect to Lebesgue measure – i.e., uniqueness holds except for a set of price functions and channel gain coefficients of Lebesgue measure zero. In particular, if channel gains are drawn at the outset of the game following some fixed, continuous probability distribution (e.g., induced by the SUs' spatial distribution), then this means that $\mathcal{G}$ admits a unique equilibrium with probability 1.

B. Exponential learning and convergence to equilibrium

The equilibrium characterization of Theorem 1 guarantees a very robust solution set (a convex polytope); in fact, as we just saw, the game's equilibrium set is a singleton under fairly mild conditions for the users' price functions. Regardless, since it is not clear how to compute the solution of the problem (10) in a distributed way, our goal in this section will be to provide a simple learning mechanism that allows users to reach a Nash equilibrium.

Our proposed algorithm will rely on the *marginal utilities*:

$$\mathbf{v}_k(\mathbf{p}) = \nabla_k u_k(\mathbf{p}) \tag{14}$$

where ∇_k denotes differentiation with respect to the power profile $\mathbf{p}_k$ of user k. In particular, writing $\mathbf{v}_k = (v_{k,1}, \dots, v_{k,S})$, some easy algebra yields the component-wise expression

$$v_{ks}(\mathbf{p}) = \frac{\partial u_{ks}}{\partial p_{ks}} = g_{ks}\left(\frac{1}{\sigma_s^2 + w_s} - \frac{\partial \pi_0}{\partial w_s}\right) - \frac{\partial \pi_k}{\partial p_{ks}}, \tag{15}$$

which shows that $v_{ks}(\mathbf{p})$ can be calculated by each individual user knowing only their SINR per subcarrier (which is measured locally) and the functional form of the price functions π_0 and π_k (which are agreed upon by the network's SUs and the PU and are thus also known locally). Indeed, Eq. (5) shows that the aggregate interference level on subcarrier s can be calculated by user k as:

$$\begin{aligned} w_s(\mathbf{p}) &= \sum\nolimits_k g_{ks} p_{ks} = g_{ks} p_{ks} + \sum\nolimits_{\ell\neq k} g_{\ell s} p_{\ell s} \\ &= g_{ks} p_{ks} + \frac{g_{ks} p_{ks}}{\gamma_{ks}(\mathbf{p})} = g_{ks} p_{ks} \frac{1 + \gamma_{ks}(\mathbf{p})}{\gamma_{ks}(\mathbf{p})}, \end{aligned} \tag{16}$$

Algorithm 1 Exponential Learning for Rate Maximization under Pricing

Parameter: step size δ_n.
Initialize: $n \leftarrow 0$; scores $y_{ks} \leftarrow 0$ for all $k \in \mathcal{K}$, $s \in \mathcal{S}$.
Repeat
  $n \leftarrow n + 1$;
  for each user $k \in \mathcal{K}$ **do simultaneously**
    for each subcarrier $s \in \mathcal{S}$ **do simultaneously**
      set transmit power $p_{ks} \leftarrow P_k \dfrac{\exp(y_{ks})}{1 + \sum_r \exp(y_{kr})}$;
      get estimate $\hat{v}_{ks}$ of $v_{ks}(\mathbf{p})$;
      update scores: $y_{ks} \leftarrow y_{ks} + \delta_n \hat{v}_{ks}$;
until termination criterion is reached.

i.e., requiring only local SINR measurements and the knowledge of the user's channel (which can in turn be obtained through the exchange of pilot signals). As a result, the marginal utility vectors v_k can be calculated in a completely distributed fashion with locally available information.

In view of the above, a natural learning process would be for the users to track their marginal utility vectors with the hope of converging to a Nash equilibrium; however, given the problem's power and positivity constraints, this method may quickly lead to inadmissible power profiles that do not lie in $\mathcal{X}$ – in which case convergence is also out of the question. To account for this, we will employ an interior point method which increases power on subcarriers that seem to be performing well, without ever shutting off a particular channel completely.

Formally, consider the *exponential regularization map* $\mathbf{G}\colon \mathbb{R}^{\mathcal{S}} \to \mathbb{R}^{\mathcal{S}}_{+}$ given by

$$\mathbf{G}(\mathbf{v}) = \frac{1}{1 + \sum_s \exp(v_s)} \left(\exp(v_1), \dots, \exp(v_S)\right). \tag{17}$$

This map has the property that it assigns positive weight (power) to all subcarriers and exponentially more weight to the subcarriers $s \in \mathcal{S}$ with the highest marginal utilities v_s. Furthermore, if all marginal utilities are relatively low (indicating high transmission costs), all assigned weights will also be low in order to decrease the user's cost. With this in mind, our proposed *exponential learning* algorithm will rely on the recursion:

$$\begin{aligned} \mathbf{y}_k(n+1) &= \mathbf{y}_k(n) + \delta_n \mathbf{v}_k(\mathbf{p}(n)), \\ \mathbf{p}_k(n+1) &= \mathbf{G}(\mathbf{y}_k(n)), \end{aligned} \tag{XL}$$

where δ_n is a variable step-size sequence whose role will be discussed below (for an algorithmic implementation, see Algorithm 1).

From an implementation point of view, Algorithm 1 has the following desirable properties:

(P1) It is *distributed:* users only need local or publicly available information in order to run it.
(P2) It is *stateless:* users do not need to know the state of the system (e.g., its topology).
(P3) It is *reinforcing:* users tend to allocate more power to cost-efficient subcarriers.

Note that if transmissions are allowed only within spectrum holes, the subcarriers in which PUs transmission activities have been sensed can be removed from those to be considered in the problem. On the contrary, if both PUs and SUs can simultaneously transmit on the same subcarrier, Algorithm 1 guarantees that the aggregate interference to PUs does not exceed the maximum interference level $I^{\max}$. It follows that the proposed approach does not depend on a specific spectrum sensing mechanism. Instead, Algorithm 1 is complementary to the spectrum sensing phase, i.e. it can be exploited to maximize network performance under different spectrum sensing schemes.

TABLE I
PER-USER COMPUTATIONAL COMPLEXITY COMPARISON

Algorithm	Complexity
Proposed	$\mathcal{O}(S)$
IWF [31]	$\mathcal{O}(S \log S)$
PIC [11, 32]	$\mathcal{O}(S \log S + \mu S)$

We then obtain:

Theorem 2. *Let δ_n be a variable step-size sequence such that $\sum_n \delta_n = \infty$ and $\sum_{j=1}^n \delta_j^2 / \sum_{j=1}^n \delta_j \to 0$. Then, Algorithm 1 converges to Nash equilibrium in the cost-efficient rate maximization game $\mathcal{G}$.*

Proof: See Appendix B. ■

Remark 2. The condition $\sum_{j=1}^n \delta_j^2 / \sum_{j=1}^n \delta_j \to 0$ requires the use of a decreasing step-size δ_n (which slows down the algorithm), but the rate of decay of δ_n can be arbitrarily slow – in stark contrast to the more stringent requirement $\sum_j \delta_j^2 < \infty$ that is common in the theory of stochastic approximation [30]. As such, Algorithm 1 can be used with an effectively constant (very slowly varying) step-size, and still converge to equilibrium; we explore this issue in detail in Section V.

Remark 3. As can be readily seen, the per-user computational complexity of each iteration of Algorithm 1 is linear in the number of the available subcarriers S (and it does not depend on K). By comparison, the corresponding complexity of iterative water-filling (IWF) [31] or power interference-constrained (PIC) methods [11, 32] is $\mathcal{O}(S \log S)$, so Algorithm 1 is quite light computationally (cf. Table I).

C. Learning in the presence of noise and uncertainty

An important assumption underlying the analysis of the previous section is that the system's SUs have access to perfect channel state information (CSI) and SINR measurements by which to calculate their marginal utility vectors $\mathbf{v}_k(\mathbf{p}(n))$ at each stage $n = 1, 2, \dots$ of their learning process. In practice however, factors such as noise uncertainty, pilot contamination, sparse feedback and imperfect channel sampling could have a deleterious effect on the algorithm's performance. As such, our goal in this section will be to analyze the convergence properties of Algorithm 1 in the presence of observation noise and uncertainty.

To formalize this, assume that the system's SUs are only able to observe a noisy estimate $\hat{\mathbf{v}}_k(n) = \mathbf{v}_k(\mathbf{p}(n)) + \mathbf{z}_k(n)$ of their true marginal utility vectors $\mathbf{v}_k(\mathbf{p}(n))$ at stage n, with the noise process $\mathbf{z}(n) = (\mathbf{z}_1(n), \dots, \mathbf{z}_N(n))$ satisfying the very general statistical hypotheses:

(H1) *Unbiasedness:*

$$\mathbb{E}\left[\mathbf{z}(n+1) \mid \mathbf{p}(n)\right] = 0. \tag{H1}$$

(H2) *Finite mean square error (MSE):*

$$\mathbb{E}\left[\, \|\mathbf{z}(n+1)\|^2 \mid \mathbf{p}(n)\right] < \infty. \tag{H2}$$

The statistical hypotheses above allow us to account for a very wide range of error processes (including uniform, Gaussian and/or log-normal errors) that do not even have to be independent and identically distributed (i.i.d.). Remarkably, despite these very mild assumptions, we have:

Theorem 3. *Assume that Algorithm 1 is run with a variable step-size sequence δ_n such that $\sum_n \delta_n^2 < \sum_{n=1}^{\infty} \delta_n = \infty$ and with noisy measurements satisfying Hypotheses* (H1) *and* (H2) *above. Then, the SUs' transmit policy $\mathbf{p}(n)$ converges to Nash equilibrium in the pricing-driven rate maximization game $\mathcal{G}$ almost surely – i.e. with probability* 1.

Proof: See Appendix B. ■

We thus see that the system's SUs converge to equilibrium even with very poor measurements. The only restriction with respect to the noiseless regime is that step-sizes must be chosen less aggressively (for instance, the assumptions of Theorem 3 rule out step-sizes of the form $\delta_n \propto n^{-\alpha}$ for $\alpha \in (0, 1/2)$, even though such sequences are allowed in Theorem 2). That said, even this restriction can be relaxed if additional hypotheses are imposed on the noise process $\mathbf{z}(n)$ (for instance, that $\sup_n \mathbf{z}(n) < \infty$ almost surely); we will revisit this issue in Section V.

IV. Fast-Fading and User Mobility

Our analysis so far has focused on static channels, corresponding to wireless users with little or no mobility. In this section, we investigate the case of mobile users where the channel gain coefficients evolve over time following a stationary ergodic process. In this fast-fading regime, the users' achievable rate is given by the ergodic average [33]:

$$\bar{R}_k(\mathbf{p}) = \mathbb{E}_g\, R_k(\mathbf{p}) = \sum_{s\in\mathcal{S}} \mathbb{E}_g \log\left(1 + \frac{g_{ks}p_{ks}}{\sigma_s^2 + \sum_{\ell\neq k} g_{\ell s}p_{\ell s}}\right), \tag{18}$$

leading to the corresponding average utility functions:

$$\bar{u}_k(\mathbf{p}) = \bar{R}_k(\mathbf{p}) - \mathbb{E}_g[C_k(\mathbf{p})] = \bar{R}_k(\mathbf{p}) - \mathbb{E}_g\left[\pi_0(\mathbf{w}) + \pi_k(\mathbf{p}_k)\right], \tag{19}$$

where the expectation $\mathbb{E}_g[\cdot]$ is taken with respect to the law of the channel gain coefficients $g_{ks} = |h_{ks}|^2$. We thus obtain the following game-theoretic formulation in the presence of fast fading:

$$\begin{aligned} &\text{maximize} \quad \bar{u}_k(\mathbf{p}_k; \mathbf{p}_{-k}) \quad \text{(unilaterally for all } k \in \mathcal{K}\text{)}, \\ &\text{subject to} \quad \mathbf{p}_k \in \mathcal{X}_k. \end{aligned} \tag{20}$$

As in the static regime, we then get the following characterization of Nash equilibria:

Proposition 2. *With notation as above, let*

$$\bar{V}(\mathbf{p}) = \sum\nolimits_s \mathbb{E}_g \log\left(\sigma_s^2 + w_s\right) - \mathbb{E}_g\left[\pi_0(\mathbf{w}) + \sum\nolimits_k \pi_k(\mathbf{p}_k)\right]. \tag{21}$$

Then, $\bar{V}(\mathbf{p})$ is an exact potential for the ergodic rate maximization game $\bar{\mathcal{G}} \equiv \bar{\mathcal{G}}(\mathcal{K}, \mathcal{X}, \bar{u})$. In particular, if the channels' law is atom-free (i.e., it is absolutely continuous with respect to Lebesgue measure), $\bar{V}$ is strictly concave and $\bar{\mathcal{G}}$ admits a unique Nash equilibrium.

Proof: See Appendix C. ■

TABLE II
Simulation Setting

Parameter	Value
Carrier frequency	$f_c = 2.4\,\text{GHz}$
Channel bandwidth	$B = 10.93\,\text{KHz}$
Noise spectral density	$\sigma_s = -173\,\text{dBm/Hz}$
Maximum transmitting power of SUs	$P_k = 21.03\,\text{dBm}$
Edge of the simulated square area	$L = 200\,\text{m}$
Transmitting power of the PU	$P^{\text{PU}} = 30\,\text{dBm}$
Distance of the PU from the receiver	$d = 50\,\text{m}$

Proposition 2 shows that the inherent stochasticity in the users' channels actually helps in guaranteeing a very robust solution set for the cost-efficient throughput maximization problem (20) (see also [18] for a related result in the context of rate control). On the other hand, the expectation over the users' channels is typically hard to carry out (especially beyond the Gaussian i.i.d. regime), so it is not clear how to calculate the ergodic marginal utilities $\bar{\mathbf{v}}_k(\mathbf{p}) = \nabla_k \bar{u}_k(\mathbf{p}) = \mathbb{E}_g[\mathbf{v}(\mathbf{p})]$. Thus, instead of trying to reach a Nash equilibrium by employing a variant of Alg. 1 run with the users' ergodic marginal utilities (whose calculation requires considerable computation capabilities and a good deal of knowledge on the channels' statistics), we will consider the same sequence of events as in the case of static channels:

1) At every update period $n = 1, 2, \dots$, each user $k \in \mathcal{K}$ calculates his instantaneous marginal utility vector $\mathbf{v}_k(n)$ following (15):

$$\hat{v}_{ks}(n) = \frac{1}{p_{ks}(n)} \frac{\gamma_{ks}(n)}{1 + \gamma_{ks}(n)} - \left.\frac{\partial C_k}{p_{ks}}\right|_{\mathbf{p}(n)} \tag{22}$$

2) Users update their powers following the recursion step of Alg. 1, and the process repeats.

Remarkably, despite the inherent stochasticity, we have:

Theorem 4. *Assume that the variance of the users' channel gain coefficients is finite. If Alg. 1 is run with step-sizes δ_n such that $\sum_n \delta_n = \infty$ and $\sum_{j=1}^n \delta_j^2 / \sum_{j=1}^n \delta_j \to 0$, then the users' power profiles converge to Nash equilibrium in the cost-efficient ergodic rate maximization game $\bar{\mathcal{G}}$ (a.s.).*

Proof: See Appendix C. ■

Thanks to Theorem 4, we see that Algorithm 1 can be applied "as-is" in both static and fast-fading environments; we will explore this issue in more detail in the following section.

V. Numerical Results

To evaluate the performance of the proposed cost-efficient power allocation framework for throughput maximization in cognitive radio networks, we have performed extensive numerical simulations over a wide range of system parameters. In what follows, we provide a selection of the most representative cases.

Throughout this section, and unless explicitly mentioned otherwise, we consider a population of $K = 10$ SUs uniformly

distributed over a square area and $S = 10$ non-interfering subcarriers with channel gain coefficients g_{ks} drawn according to the path-loss model for Jakes fading proposed in [34]; the other relevant simulation parameters are summarized in Table II. For simplicity, we also assume that σ_s and P_k are equal for all $s \in \mathcal{S}$ and all $k \in \mathcal{K}$; finally, we assume that PUs have the same interference tolerance $I_s^{\max}$ over all subcarriers $s \in \mathcal{S}$.

To begin with, we evaluate the impact of interference pricing on the SUs' behavior by introducing the *violation index*

$$\Psi_s = w_s / I^{\max}, \tag{23}$$

i.e., the amount of interference generated by SUs on the s-th subcarrier relative to the PUs' tolerance. Obviously, $\Psi_s \leq 1$ means that the system's interference temperature (IT) requirements are not violated, whereas $\Psi_s > 1$ indicates a violation of the PUs' contractual QoS guarantees that may have to be reimbursed by the network's SUs. Accordingly, in Fig. 1, we plot the system's average violation index $\Psi = 1/|\mathcal{S}| \sum_{s \in \mathcal{S}} \Psi_s$ as a function of the pricing parameter λ_0 for different values of the maximum interference tolerance level $I^{\max}$ under the flat-rate pricing scheme $\pi_0(\mathbf{w})$. As can be seen, if the PUs' maximum interference tolerance level is low (i.e., $I^{\max}$ is small), SUs violate the resulting interference temperature constraint only if the value of the price parameter λ_0 is also low. Thus, the PUs' QoS guarantees are violated only in the "soft pricing" regime where the pricing parameter λ_0 is not high enough to safeguard the PUs' low interference tolerance. On the other hand, if the cost incurred due to violations is high enough, no violations are performed: our simulations show that under both the LP and VP models, there exists a threshold value of λ_0 such that the violation index at the game's NE is always less than one, i.e., the interference generated by SUs on each subcarrier is never higher than the PUs' IT constraints.

That being said, increasing the flat-rate pricing parameter λ_0 can lead to significantly different SU behavior with respect to the PUs' interference tolerance level.[3] In fact, under the LP pricing model, SU interference disincentives can become excessive: Fig. 1 shows that transmission costs for high λ_0 are so high (even for low interference levels) that SUs prefer to shut down and stop transmitting altogether. On the other hand, under the VP model, λ_0 affects the outcome of the game *only* if the PUs' maximum interference tolerance is low: increasing λ_0 beyond a certain value does not lead SUs to shut down and does not impact their sum-rate at equilibrium, precisely because SUs are charged only if they cause excessive interference to the system's primary users.

To illustrate the system's transient phase when users employ Algorithm 1 to optimize their utility, Fig. 2 shows the aggregate interference on a given subcarrier when the interference constraint is set to $I^{\max} = -70\,\text{dBm}$ and users are charged based on the VP flat-rate model. We see there that the PU's interference constraint is violated only during the first few iterations of the learning process: when the interference in a given subcarrier exceeds the PUs' tolerance, the SUs experience a sharp drop in their marginal utilities (15) because of the incurred cost $\pi_0^{\text{VP}}(\mathbf{w})$, so Algorithm 1 prompts them to reduce their radiated power in the next iteration in order to avoid further violations. In this way, SU violations are quickly reduced and the users' learning process converges to a violation-free Nash equilibrium of the cost-efficient throughput maximization game.

In Fig. 3, we evaluate the impact of pricing and power constraints on the system's performance at Nash equilibrium for different pricing models. Under the VP model, the SUs' sum-rate at equilibrium is affected by the cost parameter λ_0 only when λ_0 is small: the reason for this is that SUs do not violate the PUs' interference temperature constraints for high λ_0 (cf. Fig. 1), so their transmit power and sum-rate at equilibrium remains (almost) constant for high λ_0. On the other hand, as in the case of Fig. 1, Fig. 3 shows that the LP model (solid lines) is strongly affected by the pricing parameter λ_0, for all λ_0 values: since increasing λ_0 in the LP model increases transmission costs across the board, each SU is pushed to reduce his individual transmit power in order to reduce the induced mutual interference in the network commensurately. It is worth noting however that increasing transmission costs is *not* always detrimental to SUs under the LP model: as shown in Fig. 3, there is a pricing parameter region where the overall interference on a given channel decreases when λ_0 is increased, thus enabling users to achieve higher data rates (due to the decreased interference on the channel). Nonetheless, in the presence of much higher transmission costs, the radiated power of SUs is too low to carry any significant amount of information, thus leading to a decrease in achievable throughput.

In Fig. 3, we also show the impact of different system configurations on the achievable SU performance by plotting the users' average sum-rate at equilibrium for different values of the system's *congestion index*, i.e., the ratio K/S between the number of SUs accessing the system and the number of available subcarriers. As expected, networks with low and moderate congestion (i.e., $K/S = 0.5, 1$) exhibit better performance than highly congested networks (i.e., $K/S = 1.5$): when there is a higher number of SUs trying to access the network, the mutual interference also increases, thus causing considerable losses in throughput and leading SUs to shut down instead of incurring high transmission costs for moderate-to-low gains in throughput.

In Fig. 4 we illustrate how the SUs' sum-rate at equilibrium varies as a function of the PUs' interference tolerance $I^{\max}$ for different pricing schemes (linear vs. violation pricing and flat-rate vs. per-user pricing). Obviously, when SU transmission comes at no cost (the $\lambda_0 = 0$ case), the value of $I^{\max}$ does not impact the outcome of the game. On the other hand, when $\lambda_0 > 0$, the SUs' average sum-rate increases as the PUs' interference tolerance increases up to a critical value $I_c^{\max}$ where the SUs' sum-rate achieves its maximum value. For any tolerance level $I^{\max} > I_c^{\max}$, the SUs' average sum-rate starts decreasing and eventually converges to a well-defined limit value as $I^{\max} \to \infty$, corresponding to the case where the PU is allowing free access to the leased part of the spectrum. This occurrence is similar to what we have already discussed in

[3] Recall here that, under VP, the system's SUs are not charged when their aggregate interference w_s is lower than $I^{\max}$, and are (steeply) fined otherwise; by contrast, the LP model charges users even when the system's IT constraints are not violated.

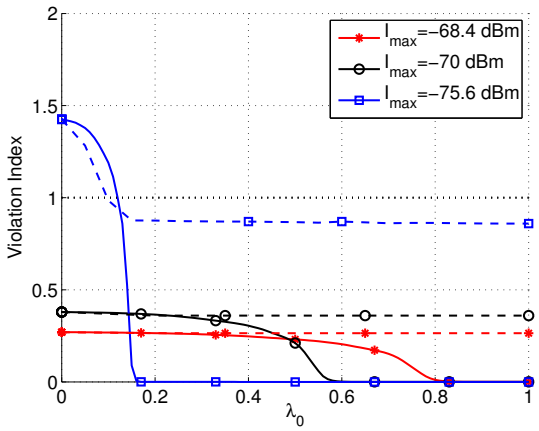


Fig. 1. Violation index as a function of λ_0 for different values of the maximum interference temperature level $I^{\max}$ under the flat-rate pricing schemes (LP: solid lines; VP: dashed lines).

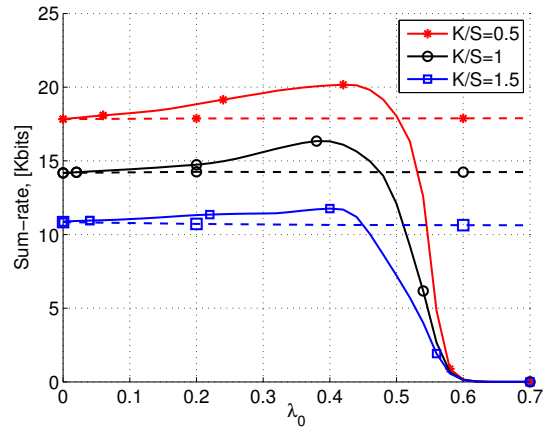


Fig. 3. Average sum-rate as a function of different pricing models, system configurations and values of the pricing parameter λ_0 (LP: solid lines; and VP: dashed lines).

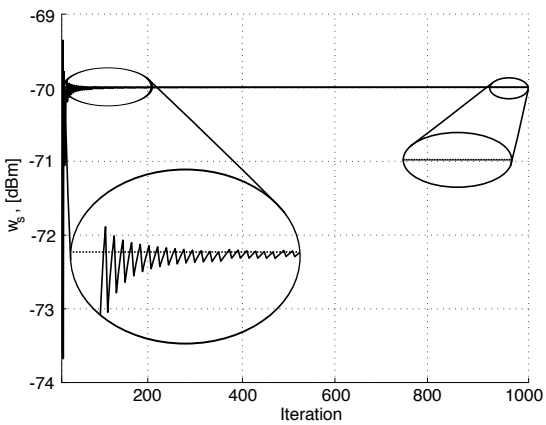


Fig. 2. Impact of the interference constraint on the evolution of the learning process under the LP model, ($I_s^{\max} = -70\,\text{dBm}$).

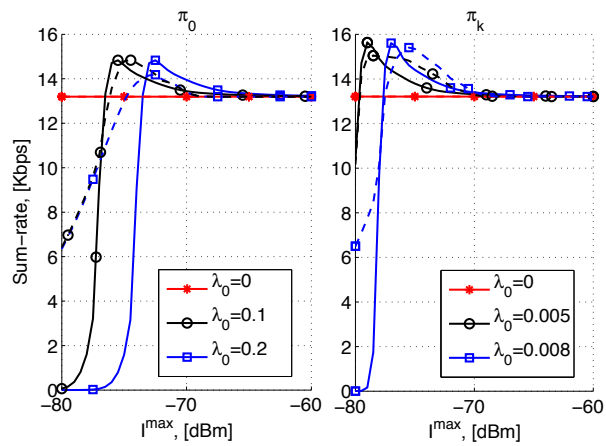


Fig. 4. Average sum-rate as a function of the maximum interference $I^{\max}$ at the PU for different pricing schemes and values of the pricing parameters λ_0 under the LP model (LP: solid lines; VP: dashed lines).

Fig. 3 and stems from the fact that low prices (small λ_0) and/or high interference tolerance (large $I^{\max}$) do not provide a strong disincentive for SUs to reduce their power level; as a result, the mutual interference across SUs also increases and leads to a decrease in the achievable performance of the secondary network. Importantly, when $I^{\max}$ is relatively low, the LP and VP models exhibit different behaviors, illustrated by the fact that the SUs' sum-rate at equilibrium differs. By contrast, (LP) and (VP) both tend to zero as $I^{\max} \to \infty$, so their behavior for very large $I^{\max}$ is similar and the system converges to the same sum-rate value.

The observed sum-rate maximum for intermediate values of $I^{\max}$ can be explained as follows: in the interference-intolerant regime (small $I^{\max}$), users hardly transmit at all because of the PUs' strict QoS requirements; on the other hand, in the "open network" regime (large $I^{\max}$), each user selfishly transmits at maximum power in order to maximize his individual throughput (since there is no cost balancing factor), thus increasing interference and reducing the users' sum-rate (in a manner similar to the classical prisoner's dilemma). As a result, the SUs' sum-rate is maximized for an intermediate value of $I^{\max}$ where SUs have to control their power in order to avoid being charged for IT violations: in other words, a proper choice of $I^{\max}$ (or, equivalently, λ_0) allows SUs to achieve a state which is both unilaterally stable and Pareto efficient (in the above sense).

Finally, in Fig. 4 we also investigate the difference between flat-rate pricing (π_0) and per-user pricing (π_k) models. Both models exhibit similar properties, but for noticeably different values of λ_0: specifically, to achieve the same sum-rate under per-user pricing, lower values of λ_0 should be considered, because users are much more sensitive to the value of λ_0 in the per-user paradigm.

In Fig. 5 we illustrate the transmission rate and revenue achieved by the PU as a function of the pricing parameter λ_0 for different values of $I^{\max}$ under the LP and VP schemes. Specifically, the PU's sum-rate is calculated as

$$R_{\text{PU}}(\mathbf{w}) = \sum_{s\in\mathcal{S}} \log\left(1 + \gamma_s^{\text{PU}}(w_s)\right), \tag{24}$$

where $\gamma_s^{\text{PU}}(w_s) = g^{\text{PU}} P^{\text{PU}}/w_s$ is the PU's SINR on the s-th subcarrier, and g^{PU} and P^{PU} denote the PU's channel gain and

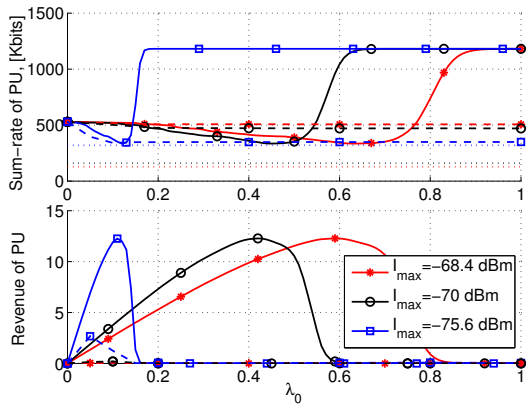


Fig. 5. Sum rate and revenue of the PU and total transmit power of the SUs as a function of λ_0 for different values of the maximum interference level $I^{\max}$ under different pricing schemes (LP: solid lines; VP: dashed lines; Minimum data rate: dotted lines).

transmit power, respectively; by the same token, the revenue of the PU is simply $K\pi_0 + \sum_k \pi_k$, i.e., the sum of the charges paid by the SUs. For comparison purposes, we have fixed three different values of the parameter $I^{\max}$ according to different PU data rate requirements (−68.3, −70 and −75.6 dBm).

Regarding the LP model, Fig. 5 shows that a high pricing parameter λ_0 brings *no* revenue to the PU because it acts as a severe transmission disincentive to the SUs (cf. Fig. 3, where we saw that SUs shut down beyond a certain threshold value λ_0^*). Because of this behavior, there exists a critical value λ_0^c for the pricing parameter that maximizes the PUs' revenue: the calculation of this critical value lies beyond the scope of this paper, but it is evident that λ_0^c increases when the maximum tolerable interference $I^{\max}$ imposed by PUs also increases. On the other hand, the PUs' revenue under the VP model is almost always zero (or close to zero): the reason for this is that the VP model acts as a soft barrier (which hardens in the large λ_0 limit), so users tend to respect the PUs' requirements and thus incur no transmission-related penalties. In other words, if the PU's QoS requirements are not too sharp, the LP model acts as a good source for revenue; otherwise, if the PU's rate requirements are tight, the VP model guarantees that SUs will respect them but does not generate any income. Also, note that under both the LP and VP models, the rate of the PU is always equal or higher than his minimum required data rate (dotted lines). This is an important result that shows that pricing regulates the SUs' behavior indirectly (based on the PU's QoS requirements and revenue targets), simply by fine-tuning the exact pricing model and its parameters (e.g., λ_0).

Figs. 6(a)–6(c) compare the performance of the proposed power allocation scheme to the benchmark case of uniform power allocation – i.e., when SUs transmit at full power and allocate their power uniformly over the available subcarriers, irrespective of the PU's requirements. For some values of λ_0, the SUs' sum-rate under uniform power allocation is higher than the one achieved by the proposed approach, but this comes at the expense of violating the PU's minimum QoS requirements (which constitutes a contractual breach from the PU's perspective); on the contrary, our approach always respects the PU's contractual requirements (since the λ_0 pricing parameter is negotiated with the PU), while guaranteeing high throughput to the SUs. This is seen in Fig. 6(b): the PU's throughput exceeds the throughput achieved when SUs employ a uniform power allocation policy, except when the PU has no significant QoS requirements ($I^{\max} \to \infty$), in which case the SUs exploit all the available spectrum and the PU's rate is reduced. Furthermore, in Fig. 6(c) we illustrate the normalized revenue of the proposed approach w.r.t. the revenues generated by uniform power allocation policies. Note that the income generated by the proposed approach is up to 3× higher than the income generated by SUs that are not cost-/energy-aware and transmit naïvely at full power, using a uniform power allocation policy.[4] Thus, by fine-tuning his pricing scheme, the PU not only achieves his QoS requirements, but also increases his monetary revenue against cost-aware SUs.

In Figs. 7 and 8, we investigate the length of the system's off-equilibrium phase and the convergence rate of the proposed distributed learning scheme (Algorithm 1). By Theorem 2, the iterations of Algorithm 1 converge to Nash equilibrium when using a step-size sequence δ_n such that $\sum_{j=1}^n \delta_j^2 / \sum_{j=1}^n \delta_j \to 0$ as $n \to \infty$. As discussed in [17], a rapidly decreasing step-size sequence slows down the algorithm, so we examine here the usage of a fixed step size to accelerate convergence. This choice makes the algorithm run faster; on the other hand, a fixed step-size may lead to unwanted oscillations around the equilibrium point, thus interfering with the algorithm's end-state. To account for this, we employ an adaptive search-then-converge (STC) approach [35]: we start with a large, constant step-size which is then decreased as soon as oscillations are detected.[5] By means of this approach, Algorithm 1 is very aggressive during the first non-oscillating iterations and it becomes more conservative (thus guaranteeing convergence) once oscillations are noticed.

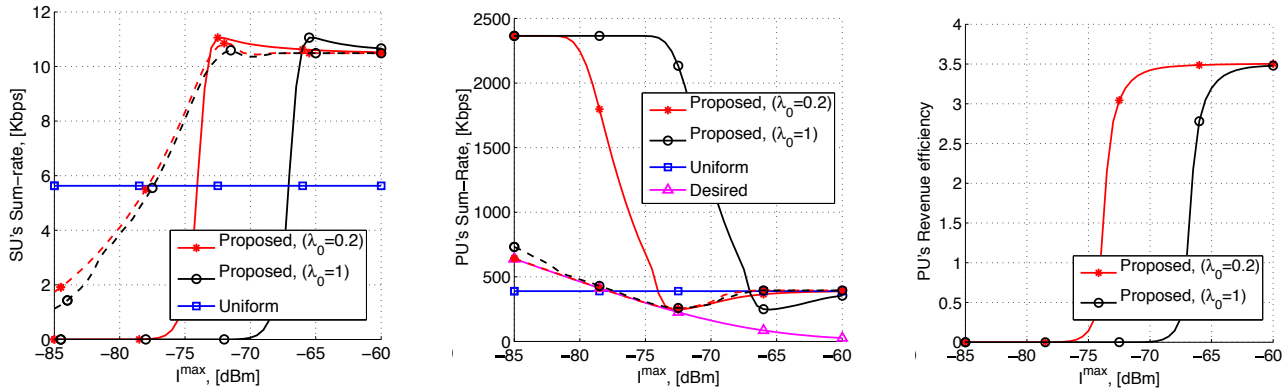
To assess the method's efficiency, we plotted the system's equilibration level (EQL) defined as:

$$\mathsf{EQL}(n) = \frac{V_n - V_{\min}}{V_{\max} - V_{\min}} \tag{25}$$

where $V_n \equiv V(\mathbf{p}(n))$ is the potential (11) of the game at the n-th iteration of the algorithm, and $V_{\min}$ ($V_{\max}$) is the minimum (maximum) value of V; obviously, an EQL value of 1 means that the system is at Nash equilibrium. Accordingly, in Fig. 7, we show the evolution of the EQL and the system's sum-rate at each iteration for different step-size rules and interference pricing models. As expected, a conservative step-size of the form $\delta_n = n^{-\beta}$, $1/2 < \beta < 1$, leads to relatively slow convergence (of the order of several tens of iterations or worse). On the other hand, the use of STC and fixed-step methods greatly accelerates the users' learning rate: after only a few STC iterations the system's EQL exceeds 90%, and the algorithm's convergence is accelerated even further by increasing the STC step-size.

[4]Recall here that the VP model does not generate any revenue so, to reduce clutter, the corresponding curves are not shown.

[5]Note that such a step-size schedule still satisfies the summability postulates of Theorem 2.



(a) (b) (c)

Fig. 6. Comparison between the proposed and uniform power allocation approaches: a) Sum-rate of SUs; b) Sum-rate of the PU; c) Normalized revenue of the proposed approach w.r.t. the uniform power allocation policy (LP: solid lines with star and circle markers; VP: dashed lines with star and circle markers).

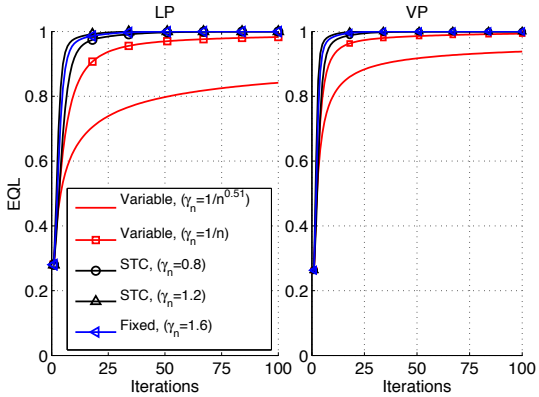


Fig. 7. Equilibration level, EQL(n), for different step-size rules under and flat-rate interference pricing models.

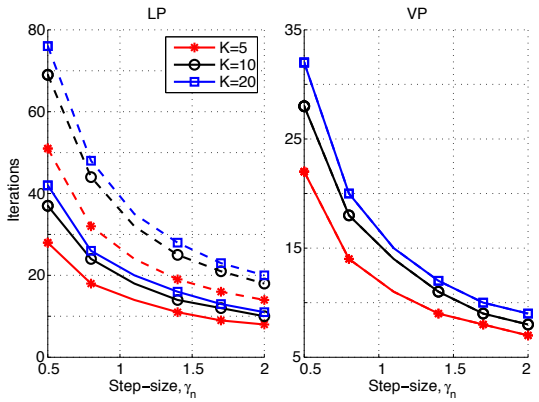


Fig. 8. Scalability of the proposed learning scheme as a function of the step-size δ for different values of the number K of SUs and pricing schemes ($\lambda_0 = 0.1$: solid lines; $\lambda_0 = 0.5$ dashed lines).

To investigate the scalability of the proposed learning scheme, we also examine the algorithm's convergence speed for different numbers of SUs. In Fig. 8 we show the number of iterations needed to reach an EQL of 95%: importantly, by increasing the value of the algorithm's step-size, it is possible to reduce the system's transient phase to a few iterations, even for large numbers of users. Moreover, we also note that the algorithm's convergence speed in the LP model depends on the pricing parameter λ_0 (it decreases with λ_0), whereas this is no longer the case under the VP model. The reason for this is again that the VP model acts as a "barrier" which is only activated when the PUs' interference tolerance is violated.

In Fig. 9 we plot the EQL of the system under intermittent PU activity which leads to the creation of spectrum holes at random intervals. In particular, we focus on the LP model and we assume that the PU disconnects from a given subcarrier at $n = \{30, 80, 140\}$ (all subcarriers are assumed to be occupied at $n = 0$). When a spectrum hole is detected (i.e. no transmission activities are reported on the sensed subcarrier), SUs are no longer constrained to keep their transmissions below a given interference level (though, of course, the SUs throughput is still adversely affected by very high interference levels); as a result, the game changes in a dynamic way that cannot be predicted in advance by the network's SUs. To adapt in this dynamic environment, we thus consider two separate cases: *a*) at each sensed vacancy, the algorithm's step-size is re-initialized but the current power allocation profile is used as an initial condition (Case A); *b*) the algorithm reboots (Case B). In both cases, Fig. 9 highlights that the proposed algorithm reaches a high EQL and converges to the new Nash equilibrium of the game within a few iterations, thus exhibiting high adaptability to dynamic changes in the users' environment.

Fig. 10 shows the performance of Algorithm 1 when CSI and SINR measurements are affected by noisy estimations as in Section III-C. More specifically, we plot the average EQL over $N = 10^3$ different realizations of the channel gain coefficients. In our simulations, we consider the case where the error process is a zero-mean Gaussian vector with finite variance and standard deviation equal to the 50% of the magnitude of the marginal utilities (i.e. the relative error level is at 50%, representing highly noisy measurements). As shown in Fig. 10, Algorithm 1 converges to the NE of the game within a few iterations even in the case of highly inaccurate measurements.

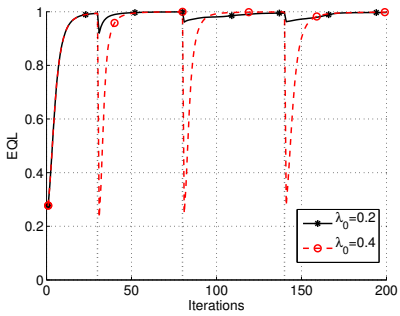


Fig. 9. Equilibration level, EQL(n), under dynamic spectrum changes.

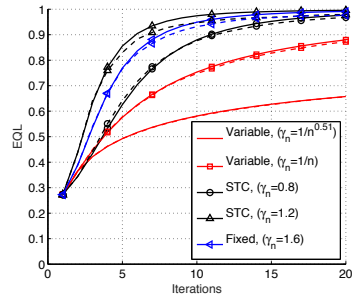


Fig. 10. Equilibration level, EQL(n), under noisy measurements (Perfect measurements: solid lines; Noisy measurements: dashed lines).

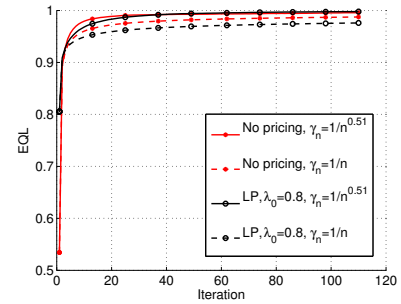


Fig. 11. Equilibration level, EQL(n), for different values of the pricing parameter λ_0 and step-size rules under the fast-fading regime.

Finally, to investigate the impact of mobility and channel fading on the users' learning process, we consider a system with three SUs ($K = 3$) and three i.i.d. Gaussian fast-fading orthogonal subcarriers ($S = 3$). In Fig. 11, we plot the system's EQL with respect to the ergodic potential (21) under the LP model as a function of different price settings and step-sizes. Remarkably, even in this stochastic setting, Algorithm 1 still converges to the game's NE in a few iterations and, as before, the algorithm's convergence rate is improved by choosing more aggressive step-size sequences.

VI. Conclusions

In this paper, we considered a game-theoretic formulation for throughput maximization in multi-carrier CR networks where SUs are charged based on the interference that they cause to the system's PUs. We showed that the resulting game admits a unique Nash equilibrium under fairly mild conditions (and for both static and ergodic channels), and we derived a fully distributed learning algorithm that converges to equilibrium using only local SINR and – possibly *imperfect* – channel measurements (and, again, under both static and fast-fading channel conditions).

Our analysis shows that the choice of the exact pricing scheme has a strong impact on the network's achievable performance (for both licensed and unlicensed users): in the "soft-pricing" regime, the PUs' requirements are violated in exchange for monetary reimbursement; by contrast, higher prices safeguard the PUs' requirements, but (somewhat surprisingly) generate no revenue to the PUs. Moreover, thanks to the algorithm's rapid convergence, the system's transient (off-equilibrium) phase is minimized, so SUs avoid being unduly charged for low throughput levels.

An open question that remains is the behavior of the system under arbitrarily time-varying channel conditions corresponding to more general fading models. Additionally, it is also important to consider dynamic pricing schemes set by the PUs in anticipation of their future needs (or to study the performance of the proposed algorithm in a fully dynamic setting where there is no target state to converge to). A natural way to address such questions would be by studying the users' so-called "regret", i.e. the difference between the users' achieved utility over a given horizon compared to what they would have obtained by choosing the best fixed transmit policy in hindsight (i.e. assuming that the users could somehow anticipate the system's evolution in advance). We intend to explore these directions in future work.

Appendix
Technical Proofs

A. Equilibrium analysis

Proof of Theorem 1: We will first show that the game's potential V is strictly concave under assumption (A1) (i.e., if π_k is strictly increasing in each of its arguments). To that end, let $V_0 = \sum_s \log(\sigma_s^2 + w_s) - \pi_0$, $V_+ = -\sum_k \pi_k$ and differentiate $V = V_0 + V_+$ to obtain $\frac{\partial V}{\partial p_{ks}} = \frac{\partial V_0}{\partial p_{ks}} + \frac{\partial V_+}{\partial p_{ks}} = \frac{\partial V_0}{\partial w_s} g_{ks} - \frac{\partial \pi_k}{\partial p_{ks}}$. Hence:

$$\frac{\partial^2 V}{\partial p_{ks}\, \partial p_{\ell s'}} = -g_{ks} g_{\ell s'} A^0_{ss'} - \delta_{k\ell} B^k_{ss'}, \tag{26}$$

where $A^0_{ss'} = -\frac{\partial^2 V_0}{\partial w_s\, \partial w_{s'}}$ and $B^k_{ss'} = \frac{\partial^2 \pi_k}{\partial p_{ks}\, \partial p_{ks'}}$.

Since V_0 is strictly concave in $\mathbf{w}$ (as the sum of a strictly concave function and a concave function), it follows that $\{A^0_{ss'}\}$ is positive-definite. Accordingly, since $A^0_{ss'}$ does not depend on k, any zero eigenvector $\mathbf{z} \in \mathbb{R}^{KS}$ of the $KS \times KS$ matrix $g_{ks} g_{\ell s'} A^0_{ss'}$ must satisfy:

$$\sum\nolimits_k g_{ks} z_{ks} = 0 \quad \text{for all } s \in \mathcal{S}. \tag{27}$$

The degeneracy condition (27) reflects the fact that if $\mathbf{w}(\mathbf{p}') = \sum_k g_{ks} p'_{ks} = \sum_k g_{ks} p_{ks} = \mathbf{w}(\mathbf{p})$ for two power profiles $\mathbf{p}, \mathbf{p}' \in \mathcal{X}$, then $V_0(\mathbf{p}) = V_0(\mathbf{p}')$; Eq. (27) shows in addition that V_0 admits no other directions along which it is constant. From this, it follows that the kernel Z of Hess(V) is at most S-dimensional; since $\arg\max V$ lies in an affine subspace of $\mathbb{R}^{KS}$ that is parallel to Z, we conclude that the Nash set of $\mathcal{G}$ is a convex polytope of dimension at most $KS - S$, as claimed.

Assume now that $\mathbf{p}^*$ is a Nash equilibrium of $\mathcal{G}$. If there exists a subcarrier $s \in \mathcal{S}$ such that $p^*_{ks} = 0$ for all $k \in \mathcal{K}$, then any profile with $p_{ks} = P_k$ for all $k \in \mathcal{K}$ cannot be Nash – and vice versa. Thus, without loss of generality (and after relabeling indices if necessary), we may assume that there exists a subcarrier $s \in \mathcal{S}$ such that $p^*_{ks} < p^*_{\ell s}$ for two users $k, \ell \in \mathcal{K}$. With this in mind, assume that every user-specific price function π_k is increasing in each of its arguments and consider the tangent vector $\mathbf{z} \in \mathbb{R}^{KS}$ with $z_{ks} = g_{\ell s}$, $z_{\ell s} = -g_{ks}$, and $z_{k's'} = 0$ otherwise. By (27), it follows that $f(t) = V(\mathbf{p}^* + t\mathbf{z})$

is constant for all sufficiently small $t \geq 0$ (note that $\mathbf{p}^* + t\mathbf{z} \in \mathcal{X}$ for small $t \geq 0$). However, by differentiating, we obtain:

$$\begin{aligned}\frac{df}{dt} &= \frac{d}{dt}\left[V_0(\mathbf{p}^* + tz) - \textstyle\sum_{k'} \pi_{k'}(\mathbf{p}^*_{k'} + t\mathbf{z}_{k'})\right] \\ &= -\frac{\partial \pi_k}{\partial p_{ks}} z_{ks} - \frac{\partial \pi_\ell}{\partial p_{\ell s}} z_{\ell s} = g_{ks}\frac{\partial \pi_\ell}{\partial p_{\ell s}} - g_{\ell s}\frac{\partial \pi_k}{\partial p_{ks}},\end{aligned} \tag{28}$$

so we must have

$$g_{ks}\left.\frac{\partial \pi_\ell}{\partial p_{\ell s}}\right|_{\mathbf{p}^*+t\mathbf{z}} = g_{\ell s}\left.\frac{\partial \pi_k}{\partial p_{ks}}\right|_{\mathbf{p}^*+t\mathbf{z}} \quad \text{for all sufficiently small } t \geq 0. \tag{29}$$

With π_k, π_ℓ strictly increasing, this only holds if π_k (resp. π_ℓ) is linear in p_{ks} (resp. p_{ks}) and the channel gain coefficients g_{ks}, $g_{\ell s}$ have the required ratio. This last condition is a (Lebesgue) measure zero event, so our assertion follows.

Otherwise, assume that (A2) holds, implying in particular that $\frac{\partial V_0}{\partial w_s} = (\sigma_s^2 + w_s)^{-1} - \frac{\partial \pi_0}{\partial w_s}$ does not change sign. Then, in view of the above, it suffices to prove uniqueness in the special case where the price functions π_k are constant in a neighborhood of $\mathbf{p}^*$. In this case, the first order KKT conditions for (13) give:

$$a) \quad r_s g_{ks} - \lambda_k \leq 0, \tag{30a}$$

$$b) \quad p_{ks}\left[r_s g_{ks} - \lambda_k\right] = 0, \tag{30b}$$

where λ_k is the Lagrange multiplier of the power constraint $\sum_s p_{ks} \leq P_k$ and $r_s = \left(\frac{1}{\sigma_s^2 + w_s} - \frac{\partial \pi_0}{\partial w_s}\right)^{-1}$. We thus get:

$$\frac{g_{ks}}{g_{ks'}} = \frac{r_s}{r'_s} \quad \text{for all } s, s' \in \operatorname{supp}(\mathbf{p}^*_k). \tag{31}$$

By using a graph-theoretic method introduced in [36], we may deduce that the following hold except on a set of (Lebesgue) measure zero. Indeed, no two users $k, \ell \in \mathcal{K}$ can be using the same two subcarriers s, s' at equilibrium: if this were the case, we would have $g_{ks}/g_{ks'} = g_{\ell s}/g_{\ell s'}$, a measure zero event. Additionally, there is at most $S - 1$ instances of users employing more than one subcarrier. To see this, assume that user k_j employs subcarriers s_j, s'_j, with $j = 1, \dots, N$, $N \geq S$. Then, by the pigeonhole principle, there exists a subset of pairs (s_j, s'_j) that forms a cycle of length $L \geq N$ in the graph with vertex set $\mathcal{S}$. Hence, by relabeling indices if necessary, we obtain the cycle relation:

$$\frac{g_{k_1,s_1}}{g_{k_2,s_2}}\frac{g_{k_2,s_2}}{g_{k_3,s_3}}\cdots\frac{g_{k_{L-1},s_{L-1}}}{g_{k_L,s_L}} = \frac{r_{s_1}}{r_{s_2}}\frac{r_{s_2}}{r_{s_3}}\cdots\frac{r_{s_{L-1}}}{r_{s_L}} = 1, \tag{32}$$

where we used the fact that $s_1 = s_L$. This is a measure zero event, so our assertion follows.

The above shows that $\mathbf{p}^*$ lies in the interior of a face of $\mathcal{X}$ with dimension at most $S - 1$. Since the Nash set of $\mathcal{G}$ is a convex polytope of dimension $KS - S$, we conclude that any Nash equilibrium lies at the intersection of a g-independent $(S-1)$-dimensional and a g-dependent $(KS - S)$-dimensional subspace of $\mathbb{R}^{KS}$. However, since $KS - S + S - 1 < KS$, the intersection of these subspaces is trivial on a set of full measure with respect to the choice of g, implying that there exists a unique Nash equilibrium. ■

B. Convergence of exponential learning

The first step of our proof will be to show that the iterates of Algorithm 1 track the "mean-field" dynamics:

$$\begin{aligned}\dot{\mathbf{y}}_k &= \mathbf{v}_k(\mathbf{p}), \\ p_{ks} &= P_k \frac{\exp(y_{ks})}{1 + \sum_{s' \in \mathcal{S}} \exp(y_{ks'})}.\end{aligned} \tag{33}$$

Theorems 2 and 3 will then follow by showing that the dynamics (33) converge to the maximum set of the game's potential (and, hence, to Nash equilibrium) for any itial condition $\mathbf{y}(0)$.

For simplicity, in the rest of this appendix, we will work with a single user with maximum transmit power $P = 1$; the general case can be treated similarly. Also, since Theorem 2 is essentially a special case of Theorem 3 for the noise process $\mathbf{z}_k \equiv 0$ (which obviously satisfies both (H1) and (H2)), we will prove the two theorems simultaneously.

With this in mind, let $\mathcal{D} = \{\mathbf{p} \in \mathbb{R}^{\mathcal{S}}_+ : 0 \leq \sum_s p_s \leq 1\}$ denote the standard S-dimensional "corner-of-cube", and consider the entropy-like function:

$$h(\mathbf{p}) = \sum\nolimits_s p_s \log p_s + \left(1 - \sum\nolimits_s p_s\right)\log\left(1 - \sum\nolimits_s p_s\right). \tag{34}$$

A key element of our proof will be the associated *Bregman divergence* [37, 38]:

$$D_h(\mathbf{p}^*, \mathbf{p}) = \textstyle\sum_s p^*_s \log \frac{p^*_s}{p_s} + (1 - \sum_s p^*_s)\log\frac{1 - \sum_s p^*_s}{1 - \sum_s p_s}, \tag{35}$$

with the continuity convention $0 \log 0 = 0$. Employing the above, we then get:

Proposition 3. *Every solution orbit* $\mathbf{p}(t)$ *of the dynamics* (33) *converges to a Nash equilibrium of* $\mathcal{G}$.

Proof: Let $\mathbf{p}^*$ be a Nash equilibrium of $\mathcal{G}$, and let $H(t) = D_h(\mathbf{p}^*, \mathbf{p}(t))$. We then have:

$$H = h(\mathbf{p}^*) + \log\left(1 + \sum\nolimits_s e^{y_s}\right) - \sum\nolimits_s p^*_s y_s, \tag{36}$$

and hence:

$$\begin{aligned}\dot{H} &= \frac{\sum_s \dot{y}_s e^{y_s}}{1 + \sum_s e^{y_s}} - \sum\nolimits_s p^*_s \dot{y}_s = \sum\nolimits_s p_s \dot{y}_s - \sum\nolimits_s p^*_s \dot{y}_s \\ &= \sum\nolimits_s (p_s - p^*_s) v_s = \langle \mathbf{p} - \mathbf{p}^* | \mathbf{v} \rangle.\end{aligned} \tag{37}$$

By concavity of V and the fact that $\mathbf{v} = \nabla_{\mathbf{p}} u = \nabla_{\mathbf{p}} V$, it follows that $\langle \mathbf{p} - \mathbf{p}^* | \mathbf{v} \rangle \geq 0$ with equality holding if and only if $\mathbf{p}$ is a maximizer of V (and, hence, a Nash equilibrium of $\mathcal{G}$).

To show that $\mathbf{p}(t)$ converges to a Nash equilibrium of $\mathcal{G}$, assume that $\mathbf{p}^*$ is an ω-limit of $\mathbf{p}(t)$, i.e., $\mathbf{p}(t_n) \to \mathbf{p}^*$ for some increasing sequence $t_n \to \infty$ (that $\mathbf{p}(t)$ admits at least one ω-limit follows from the fact that $\mathcal{D}$ is compact). This implies that $H(t_n) \to 0$, and since $\dot{H} \geq 0$, we also get $\lim_{t\to\infty} H(t) = 0$, so $\mathbf{p}(t) \to \mathbf{p}^*$ by the definition of the Bregman divergence. ■

Proof of Theorems 2 and 3: We will first show that (XL) comprises a stochastic approximation of the dynamics (33) in the sense of [30]. Indeed, it is easy to see that the exponential regularization map (17) is Lipschitz; moreover, since $\mathcal{D}$ is compact and the game's potential function is smooth on $\mathcal{D}$, it follows that the composite map $\mathbf{y} \mapsto \mathbf{v}(\mathbf{p}(\mathbf{y}))$ is also Lipschitz. As a result, Propositions 4.1 and 4.2 of [30] show that (XL) is an asymptotic pseudotrajectory (APT) of (33).

Now, let $\mathcal{D}^*$ denote the Nash set of $\mathcal{G}$ and assume that $\mathbf{p}(n)$ remains a bounded distance away from $\mathcal{D}^*$. Also, fix some $\mathbf{p}^* \in \mathcal{D}^*$ and let $D_n = D_h(p^*, \mathbf{p}(n))$; then, (37) gives:

$$\begin{aligned} D_{n+1} =& D_h(\mathbf{p}^*, \mathbf{p}(\mathbf{y}(n) + \delta_n \hat{\mathbf{v}}(n))) \\ \leq& D_n - \delta_n \langle \mathbf{v}(n) | \mathbf{p}^* - \mathbf{p}(n) \rangle - \delta_n \xi_n + \tfrac{1}{2} M \delta_n^2 \|\hat{\mathbf{v}}(n)\|^2 , \end{aligned} \quad (38)$$

where $\xi_n = \langle \mathbf{z}(n) | \mathbf{p}^* - \mathbf{p}(n) \rangle$ and $M > 0$ is a positive constant (that such a constant exists is a consequence of the fact that $\mathrm{Hess}(h) \succcurlyeq mI$ for some $m > 0$ [39]).

Our original assumption for $\mathbf{p}(n)$ implies that D_n is bounded away from zero; moreover, with V concave, we also get $\langle \mathbf{v}(n) | (\mathbf{p}(n) - \mathbf{p}^*) \rangle \leq -m$ for some $m > 0$. Thus, (38) yields:

$$D_{n+1} \leq D_0 - t_n \left(m - \sum\nolimits_{j=1}^n w_{j,n} \xi_j \right) + \frac{1}{2} \sum\nolimits_{j=1}^n \delta_j^2 \|\hat{\mathbf{v}}(j)\|^2 , \quad (39)$$

where $t_n = \sum_{j=1}^n \delta_j$ and $w_{j,n} = \delta_j / t_n$. By the strong law of large numbers for martingale differences [40, Theorem 2.18], we have $n^{-1} \sum_{j=1}^n \xi_j \to 0$ (a.s.); hence, with $\delta_{n+1}/\delta_n \leq 1$, Hardy's summability criterion [41, p. 58] applied to the weight sequence $w_{j,n} = \delta_j / t_n$ yields $\sum_{j=1}^n w_{j,n} \xi_j \to 0$ (a.s.). Finally, since δ_n is square-summable and $\delta_n \mathbf{z}(n)$ is a martingale difference with finite variance, it follows that $\sum_{n=1}^\infty \delta_n^2 \|\hat{\mathbf{v}}(n)\|^2 < \infty$ (a.s.) by Theorem 6 in [42]. Combining all of the above, we see that the RHS of (39) tends to $-\infty$ (a.s.); this contradicts the fact that $D_n \leq 0$, so we conclude that $\mathbf{p}(n)$ visits a compact neighborhood of $\mathcal{D}^*$ infinitely often. Since $\mathcal{D}^*$ attracts any initial condition $\mathbf{y}(0)$ under the continuous-time dynamics (33), Theorem 6.10 in [30] shows that $\mathbf{p}(n) \to \mathcal{D}^*$, as claimed.

For the more general sequences δ_n of Theorem 2 (the deterministic case), simply note that $\sum_{j=1}^n \delta_j^2 \|\hat{\mathbf{v}}(j)\| \leq L^2 \sum_{j=1}^n \delta_j^2$ for some positive constant $L > 0$ (recall that the game's strategy space is compact). Our claim then follows by noting that $\mathbb{E}[\xi_j] = 0$ and $t_n^{-1} \sum_{j=1}^n \delta_j^2 \to 0$ as $n \to \infty$. ■

C. The fast-fading case

In this appendix, we prove the uniqueness of NE in the ergodic game $\bar{\mathcal{G}}$ (Prop. 2) and the convergence of Algorithm 1 in the presence of fast fading.

Proof of Proposition 2: That $\bar{V}$ is an exact potential for $\bar{\mathcal{G}}$ follows by verification. For strict concavity, let $\mathrm{Hess}_g(V)$ denote the Hessian of the static potential function V for a given g. Then, the dominated convergence theorem allows us to interchange differentiation and integration, so we obtain $\mathrm{Hess}(\bar{V}) = \mathbb{E}_g[\mathrm{Hess}_g(V)]$. Thus, for all $\mathbf{z} \in \mathbb{R}^{KS}$, we will have:

$$\mathbf{z}^\dagger \cdot \mathrm{Hess}(\bar{V}) \cdot \mathbf{z} = \mathbb{E}_g \left[\mathbf{z}^\dagger \cdot \mathrm{Hess}_g(V) \cdot \mathbf{z} \right] \geq 0. \quad (40)$$

From the proof of Theorem 1, we know that $\mathbf{z}^\dagger \cdot \mathrm{Hess}_g(V) \cdot \mathbf{z}$ only if $\sum_k g_{ks} z_{ks} = 0$ for all $s \in \mathcal{S}$; however, since this is a measure zero event, we will have $\mathbf{z}^\dagger \cdot \mathrm{Hess}_g(V) \cdot \mathbf{z} > 0$ on a set of positive measure. This shows that $\mathbf{z}^\dagger \cdot \mathrm{Hess}(\bar{V}) \cdot \mathbf{z} > 0$ for all $\mathbf{z} \in \mathbb{R}^{KS}$, i.e. $\bar{V}$ is strictly concave. We conclude that $\bar{\mathcal{G}}$ admits a unique equilibrium, as claimed. ■

Proof of Theorem 4: The same reasoning as in the proof of Theorem 2 shows that the iterates of Algorithm 1 run with the players' instantaneous utilities calculated as in (22) comprise a stochastic approximation (asymptotic pseudotrajectory) of the mean dynamics:

$$\begin{aligned} \dot{\mathbf{y}}_k &= \bar{\mathbf{v}}_k(\mathbf{p}), \\ p_{ks} &= P_k \frac{\exp(y_{ks})}{1 + \sum_{s' \in \mathcal{S}} \exp(y_{ks'})}. \end{aligned} \quad (41)$$

The same reasoning as in the proof of Theorem 2 shows that (41) converges to the unique Nash equilibrium of $\bar{\mathcal{G}}$; as such, it suffices to show that any APT of (41) induced by Alg. 1 converges to equilibrium. To that end, as in (38), we get:

$$\begin{aligned} D_{n+1} =& D_h(\mathbf{p}^*, \mathbf{p}(n+1)) \\ \leq& D_n - \delta_n \langle \hat{\mathbf{v}}(n) | \mathbf{p}^* - \mathbf{p}(n) \rangle + \tfrac{1}{2} M \delta_n^2 \|\hat{\mathbf{v}}(n)\|^2 , \end{aligned} \quad (42)$$

where $\mathbf{p}^*$ is the (unique) NE of $\bar{\mathcal{G}}$ and $M > 0$ is a positive constant. The rest of the proof is as in Theorem 3. ■

References

[1] J. G. Andrews, S. Buzzi, W. Choi, S. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, June 2014.

[2] Qualcomm, "The 1000x data challenge." [Online]. Available: http://www.qualcomm.com/1000x

[3] FCC Spectrum Policy Task Force, "Report of the spectrum efficiency working group," Federal Communications Comission, Tech. Rep., November 2002.

[4] K. V. Schinasi, "Spectrum management: Better knowledge needed to take advantage of technologies that may improve spectrum efficiency," United States General Accounting Office, Tech. Rep., May 2004.

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

[6] 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.

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

[8] A. Goldsmith, S. A. Jafar, I. Maric, and S. Srinivasa, "Breaking spectrum gridlock with cognitive radios: An information theoretic perspective," *Proc. IEEE*, vol. 97, no. 5, pp. 894–914, 2009.

[9] O. Simeone, I. Stanojev, S. Savazzi, and Y. Bar-Ness, "Spectrum leasing to cooperating ad hoc secondary networks," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 1, pp. 203–213, January 2008.

[10] I. F. Akyildiz, B. F. Lo, and R. Balakrishnan, "Cooperative spectrum sensing in cognitive radio networks: A survey," *Physical Communication*, vol. 4, no. 1, pp. 40–62, 2011.

[11] S. Dikmese, S. Srinivasan, M. Shaat, F. Bader, and M. Renfors, "Spectrum sensing and resource allocation for multicarrier cognitive radio systems under interference and power constraints," *EURASIP Journal on Advances in Signal Processing*, vol. 2014, no. 1, pp. 1–12, 2014.

[12] T. Yucek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *Communications Surveys Tutorials, IEEE*, vol. 11, no. 1, pp. 116–130, First 2009.

[13] E. Axell, G. Leus, E. G. Larsson, and H. V. Poor, "Spectrum sensing for cognitive radio: State-of-the-art and recent advances," *Signal Processing Magazine, IEEE*, vol. 29, no. 3, pp. 101–116, 2012.

[14] T. Alpcan, T. Başar, R. Srikant, and E. Altman, "CDMA uplink power control as a noncooperative game," *Wireless Networks*, vol. 8, pp. 659–670, 2002.

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

[16] G. Scutari, S. Barbarossa, and D. Palomar, "Potential games: A framework for vector power control problems with coupled constraints," in *Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings. 2006 IEEE International Conference on*, vol. 4, May 2006, pp. IV–IV.

[17] S. D'Oro, P. Mertikopoulos, A. L. Moustakas, and S. Palazzo, "Adaptive transmit policies for cost-efficient power allocation in multi-carrier systems," in *WiOpt '14: Proceedings of the 12th International Symposium and Workshops on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks*, 2014.

[18] P. Mertikopoulos, E. V. Belmega, A. L. Moustakas, and S. Lasaulce, "Distributed learning policies for power allocation in multiple access

channels," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 1, pp. 96–106, January 2012.

[19] FCC Spectrum Policy Task Force, "Establishment of interference temperature metric to quantify and manage interference and to expand available unlicensed operation in certain fixed mobile and satellite frequency bands," Federal Communications Comission, Tech. Rep. FCC NPRM ET Docket 03-237, 2003.

[20] D. Niyato and E. Hossain, "Spectrum trading in cognitive radio networks: A market-equilibrium-based approach," *IEEE Wireless Commun. Mag.*, vol. 15, no. 6, pp. 71–80, December 2008.

[21] B. Wang, Y. Wu, and K. Liu, "Game theory for cognitive radio networks: An overview," *Computer networks*, vol. 54, no. 14, pp. 2537–2561, 2010.

[22] J.-S. Pang, G. Scutari, D. P. Palomar, and F. Facchinei, "Design of cognitive radio systems under temperature-interference constraints: A variational inequality approach," *IEEE Trans. Signal Process.*, vol. 58, no. 6, pp. 3251–3271, June 2010.

[23] Y. Yang, G. Scutari, P. Song, and D. P. Palomar, "Robust MIMO cognitive radio under interference temperature constraints," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 11, pp. 2465–2483, November 2013.

[24] Y. Xu and X. Zhao, "Robust power control for multiuser underlay cognitive radio networks under QoS constraints and interference temperature constraints," *Wireless Personal Communications*, vol. 75, no. 4, pp. 2383–2397, 2014.

[25] A. Mas-Colell, M. D. Whinston, and J. R. Green, *Microeconomic Theory*. Oxford University Press, 1995.

[26] W. Wang, Y. Cui, T. Peng, and W. Wang, "Noncooperative power control game with exponential pricing for cognitive radio network," in *VTC '07: Proceedings of the 2007 IEEE Vehicular Technology Conference*, April 2007, pp. 3125–3129.

[27] Z. Wang, L. Jiang, and C. He, "Optimal price-based power control algorithm in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 11, pp. 5909–5920, November 2014.

[28] P. Mertikopoulos and E. V. Belmega, "Transmit without regrets: online optimization in MIMO–OFDM cognitive radio systems," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 11, pp. 1987–1999, November 2014.

[29] D. Monderer and L. S. Shapley, "Potential games," *Games and Economic Behavior*, vol. 14, no. 1, pp. 124 – 143, 1996.

[30] M. Benaïm, "Dynamics of stochastic approximation algorithms," in *Séminaire de Probabilités XXXIII*, ser. Lecture Notes in Mathematics, J. Azéma, M. Émery, M. Ledoux, and M. Yor, Eds. Springer Berlin Heidelberg, 1999, vol. 1709, pp. 1–68.

[31] J.-S. Pang, G. Scutari, D. Palomar, and F. Facchinei, "Design of cognitive radio systems under temperature-interference constraints: A variational inequality approach," *Signal Processing, IEEE Transactions on*, vol. 58, no. 6, pp. 3251–3271, June 2010.

[32] M. Shaat and F. Bader, "Computationally efficient power allocation algorithm in multicarrier-based cognitive radio networks: Ofdm and fbmc systems," *EURASIP J. Adv. Signal Process*, vol. 2010, pp. 5:1–5:11, Jan. 2010. [Online]. Available: http://dx.doi.org/10.1155/2010/528378

[33] A. J. Goldsmith and P. P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Trans. Inf. Theory*, vol. 43, no. 6, pp. 1986–1992, 1997.

[34] G. Calcev, D. Chizhik, B. Goransson, S. Howard, H. Huang, A. Kogiantis, A. Molisch, A. Moustakas, D. Reed, and H. Xu, "A wideband spatial channel model for system-wide simulations," *IEEE Transactions on Vehicular Technology*, vol. 56, no. 2, pp. 389–403, March 2007.

[35] C. Darken and J. Moody, "Note on learning rate schedules for stochastic optimization," DTIC Document, Tech. Rep., 1992.

[36] P. Mertikopoulos, E. V. Belmega, A. L. Moustakas, and S. Lasaulce, "Dynamic power allocation games in parallel multiple access channels," in *ValueTools '11: Proceedings of the 5th International Conference on Performance Evaluation Methodologies and Tools*, 2011.

[37] L. M. Bregman, "The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming," *USSR Computational Mathematics and Mathematical Physics*, vol. 7, no. 3, pp. 200–217, 1967.

[38] K. C. Kiwiel, "Proximal minimization methods with generalized Bregman functions," *SIAM Journal on Control and Optimization*, vol. 35, pp. 1142–1168, 1997.

[39] Y. Nesterov, "Primal-dual subgradient methods for convex problems," *Mathematical Programming*, vol. 120, no. 1, pp. 221–259, 2009.

[40] P. Hall and C. C. Heyde, *Martingale Limit Theory and Its Application*, ser. Probability and Mathematical Statistics. New York: Academic Press, 1980.

[41] G. H. Hardy, *Divergent Series*. Oxford University Press, 1949.

[42] Y. S. Chow, "Convergence of sums of squares of martingale differences," *The Annals of Mathematical Statistics*, vol. 39, no. 1, 1968.

Salvatore D'Oro (S'12) received the B.S. degree in Computer Engineering and the M.S. degree in Telecommunications Engineering degree both at the University of Catania. In 2013 and 2015, he was a Visiting Researcher at Université Paris-Sud 11, Paris, France and at Ohio State University, Ohio, USA. He is currently a PhD Student at the University of Catania and his main interests are next generation communication networks, optimization, security and game theory. In 2013, he served on the Technical Program Committee (TPC) of the 20th European Wireless Conference (EW2014).

Panayotis Mertikopoulos (M'11) received the B.S. degree in physics with a major in astrophysics from the University of Athens in 2003. He received the M.Sc. and M.Phil. degrees in mathematics from Brown University in 2005 and 2006, respectively, and the Ph.D. degree in physics from the University of Athens in 2010. During 2010–2011, he was a post-doctoral researcher at the Economics Department of Ecole Polytechnique, Paris, France. Since 2011, he has been a CNRS Researcher at the Laboratoire d'Informatique de Grenoble, Grenoble, France, and has a joint research affiliation with Inria, France. He has been an Embeirikeion Foundation Fellow since 2003 and was the recipient of the best paper award at NetGCoop'12. He was a track co-chair of IFORS 2014, TPC co-chair of WiOpt 2014, and general co-chair of AlgoGT 2013. His main research interests lie in dynamical systems, optimization, game theory, and their applications to wireless networks.

Aris Moustakas (SM'04) received the B.S. degree in physics from Caltech and M.S. and Ph.D. degrees in theoretical condensed matter physics from Harvard University, respectively. He joined Bell Labs, Lucent Technologies, NJ in 1998, first in the Physical Sciences Division and then also in the Wireless Advanced Technology Laboratory. He is currently a faculty member at the Physics Dept. of the National Capodistrian University of Athens, Greece. During 2013–2014 he holds a DIGITEO Senior Chair in Orsay, France. Dr. Moustakas served as Associate Editor for the IEEE Transactions on Information Theory between 2009 and 2012. His main research interests lie in the areas of multiple antenna systems and the applications of game theory and statistical physics to communications and networks.

Sergio Palazzo (M'92−SM'99) received the degree in electrical engineering from the University of Catania, Catania, Italy, in 1977. Since 1987, he has been with the University of Catania, where is now a Professor of telecommunications networks. His current research interests include mobile systems, wireless and satellite IP networks, and protocols for the next generation of the Internet. He has been the General Chair of some ACM conferences, including MobiHoc 2006 and MobiOpp 2010, and currently is a member of the MobiHoc Steering Committee. He has also been the TPC Co-Chair of some other conferences, including IFIP Networking 2011, IWCMC 2013, and EuropeanWireless 2014. He currently serves on the Editorial Board of Ad Hoc Networks. In the recent past, he also was an Editor of the IEEE Wireless Communications Magazine, IEEE/ACM Transactions on Netrowking, IEEE Transactions on Mobile Computing, Computer Networks, and Wireless Communications and Mobile Computing.