

Design of Cognitive Radio Systems Under Temperature-Interference Constraints: A Variational Inequality Approach

Jong-Shi Pang, Gesualdo Scutari, *Member, IEEE*, Daniel P. Palomar, *Senior Member, IEEE*, and Francisco Facchinei

Abstract—The concept of cognitive radio (CR) has recently received great attention from the research community as a promising paradigm to achieve efficient use of the frequency resource by allowing the coexistence of licensed (primary) and unlicensed (secondary) users in the same bandwidth. In this paper, we propose a novel Nash equilibrium (NE) problem to model concurrent communications of cognitive secondary users who compete against each other to maximize their information rate. The formulation contains constraints on the transmit power (and possibly spectral masks) as well as aggregate interference tolerable at the primary users' receivers. The coupling among the strategies of the players due to the interference constraints presents a new challenge for the analysis of this class of Nash games that cannot be addressed using the game theoretical models proposed in the literature. For this purpose, we need the framework given by the more advanced theory of finite-dimensional variational inequalities (VI). This provides us with all the mathematical tools necessary to analyze the proposed NE problem (e.g., existence and uniqueness of the solution) and to devise alternative distributed algorithms along with their convergence properties.

Index Terms—Cognitive radio, distributed algorithms, game theory, Gaussian interference channel, Nash equilibrium, pricing games, temperature-interference constraints, variational inequality theory.

I. INTRODUCTION AND MOTIVATION

IN recent years, increasing demand of wireless services has made the radio spectrum a very scarce and precious resource. Moreover, most current wireless networks characterized

Manuscript received June 19, 2009; accepted January 07, 2010. Date of publication February 17, 2010; date of current version May 14, 2010. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Erik G. Larsson. The work of J.-S. Pang is based on research supported by the U.S.A. National Science Foundation grants DMS 0754374 and CMI 0802022 and by the Air Force Office of Sponsored Research. The work of G. Scutari and D. P. Palomar was supported by the NSFC/RGC N_HKUST604/08 research grant. The work of F. Facchinei was supported by the Italian project MIUR-PRIN 20079PLL7 Nonlinear Optimization, Variational Inequalities and Equilibrium Problems. Part of this material was presented at the 35th IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Dallas, TX, March 14–19, 2010.

J.-S. Pang and G. Scutari are with the Department of Industrial and Enterprise Systems Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA (e-mail: jspang@uiuc.edu; gscutari@uiuc.edu).

D. P. Palomar is with the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Hong Kong (e-mail: palomar@ust.hk).

F. Facchinei is with the Department of Computer and System Sciences Antonio Ruberti, University of Rome “La Sapienza,” Rome, Italy (e-mail: facchinei@dis.uniroma1.it).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TSP.2010.2043138

by fixed spectrum assignment policies are known to be very inefficient considering that licensed bandwidth demands are highly varying along the time and/or space dimensions. Indeed, according to the Federal Communications Commission (FCC), only 15% to 85% of the licensed spectrum is utilized on average [1]. Cognitive Radio (CR) originated as a possible solution to this problem [2] obtained by endowing the radio nodes with “cognitive capabilities,” e.g., the ability to sense the electromagnetic environment, make short term predictions, and react consequently by adapting transmission parameters (e.g., operating spectrum, modulation, and transmission power) in order to optimize the usage of the available resources [3]–[5].

The widely accepted debated position proposed for implementing the spectrum sharing idea of CR calls for a hierarchical access structure, distinguishing between *primary* users, or legacy spectrum holders, and *secondary* users, who access the licensed spectrum dynamically, under the constraint of not inducing intolerable Quality of Service (QoS) degradations on the primary users [3]–[5]. To deal with such constraints, both deterministic and probabilistic interference constraints have been suggested in the literature (see, e.g., [3] and [4]). In this paper, we focus on deterministic interference constraints and consider a generalization of the original *interference temperature-limit* concept introduced by the FCC Spectrum Policy Task Force [1]: The primary users, according to their own QoS requirements, impose a threshold on the maximum level of *per-carrier* and *total* (i.e., over the whole bandwidth) aggregate interference generated by the secondary users. The system design consists then in finding the most appropriate transmission strategy (according to some prescribed optimality criterion) for the competing cognitive users that share a given portion of spectrum with the primary users, under local transmit power and interference constraints.

One approach to devise such a system design would be using global optimization techniques, under the framework of network utility maximization (NUM) (see, e.g., [6] and [7]). Recent results in [8] have shown that the NUM problem based on the maximization of the information rates over frequency-selective SISO interference channels is an NP-hard problem, under different choices of the system utility function. Several attempts have been pursued in the literature to deal with the nonconvexity of such a problem. Some works proposed suboptimal or close-to-optimal algorithms based on duality theory [9]–[11] or Nash bargaining optimality criterion (see [12] and references therein) to compute, under technical conditions and/or simplifying assumptions on the users' transmission strategies (e.g.,

TDM/FDM strategies), the largest achievable rate region of the system. But these works lack any mechanism to control the amount of aggregate interference generated by the transmitters; which makes them not applicable to CR systems. On the other hand, in [13]–[15], the authors explicitly took into account the temperature-interference constraint (in terms of the total aggregate interference only), but under some simplifying *ad hoc* assumptions—high or low SINR regime [14], flat-fading CDMA interference channels [13] or multiple-access channels [15]—reducing the original nonconvex spectrum management problem to a simpler (scalar) power control convex problem [15] or a Geometric Programming [7], [13], [14]. On top of that, the algorithms proposed in most of the above papers [8]–[11], [13], [15] are centralized and computationally expensive. This raises some practical issues that are insurmountable in the CR context. For example, these algorithms need a central node having full knowledge of all the channels and interference structure at every receiver; which poses serious implementation problems in terms of scalability and amount of signaling to be exchanged among the nodes. Thus, it seems natural to concentrate on decentralized strategies, where the cognitive users are able to self-enforce the negotiated agreements on the usage of the available spectrum without the intervention of a centralized authority. The philosophy underlying this approach is a competitive optimality criterion, as every user aims for the transmission strategy that unilaterally maximizes his own payoff function. This form of equilibrium is, in fact, the well-known concept of Nash equilibrium (NE) in game theory (see, e.g., [16]).

Because of the inherently competitive nature of the interference channel, it is not surprising indeed that game theory has been already adopted to solve distributively many power control problems over either flat-fading or frequency-selective (in practice, multicarrier) Gaussian interference channels. An early application of game theory in the latter context is [17], where the information rates of the users were maximized with respect to the power allocation (under transmit power constraints) in a two-user DSL system. Extensions of the basic problem to *ad hoc* frequency-selective networks, including possibly spectral mask constraints, were given in [18]–[25]. The state-of-the-art algorithm is the asynchronous iterative waterfilling algorithm (IWFA) [24]. Results in the cited papers however have been recognized not to be applicable to CR systems because they do not provide any decentralized mechanism to control the amount of *aggregate* interference generated by the secondary users at the primary receivers. Observe that the use of the spectral masks in the IWFA in [18], [21], [23], and [24] does not provide a satisfactory solution to this problem, since spectral masks limit the power spectral density (PSD) at every *secondary transmitter* rather than the aggregate interference at the *primary receivers*. This is either too conservative (it can severely constrain the data rate of the secondary links, especially when there are no primary systems close to the secondary users) or does not result in the required interference temperature-limit constraints when the number of interfering systems grows (recall that, in [18], [21], [23], and [24], the spectral masks are fixed and chosen *a priori*).

In this paper, we fill this gap using a new framework based on VI. We propose a novel NE problem where the secondary users compete with each other to maximize the information rate

on their own link, given local constraints on the transmit power and (possibly) spectral mask, and global constraints on the maximum per-carrier and total aggregate interference tolerable by the primary receivers. A natural way to control the aggregate interference generated by the secondary users while keeping the optimization as decentralized as possible is via pricing through a penalty term. However, the introduction of such coupling constraints among the strategies of the players presents a new challenge for the analysis of the proposed Nash game. None of the current results in the literature provide a satisfactory answer to the study of the proposed game, neither the game theoretical models proposed for the interference channel in [18]–[25] nor the equilibrium models based on pricing [26]–[31] (see Section II for a detailed comparison between our NE problem and current game theoretical models using pricing techniques). For this purpose, we need the framework provided by the more advanced theory of finite-dimensional VIs [32]. Building on this framework, we prove that the proposed NE problem always admits a solution and derive sufficient conditions guaranteeing the uniqueness of this equilibrium. We then focus on distributed algorithms that converge to a solution of the NE problem and on their convergence properties. We propose alternative algorithms that differ in performance, level of protection of the primary users, computational effort and signaling among primary and secondary users, convergence analysis, and convergence speed, which makes them suitable for many different CR systems. Furthermore, at the price of some signaling from the primary to the secondary users, albeit very reduced, all these algorithms outperform (in terms of achievable sum-rate) conservative IWFA using spectral masks [20], [21], [23], [24] and overcome the main drawback of classical IWFA [17], [33], i.e., the violation of the interference temperature-limit.

On top of that, the second major contribution of the paper is to introduce a new line of analysis based on VIs in the literature of distributed power allocation games (possibly based on pricing) that is expected to be broadly applicable to other game models.

The paper is organized as follows. Section II introduces the system model and formulates the NE problem. Section III provides a detailed analysis of the NE problem with exogenous prices, i.e., the game where the prices are fixed and chosen *a priori*. Results obtained in such a case are instrumental to study the original NE problem with exogenous prices, as detailed in Section IV. Section V provides some numerical results comparing the proposed algorithms. Finally, Section VI draws some conclusions. In order to make the paper accessible to readers who are not familiar with VIs, in Appendix A, we review some basic concepts and results of finite-dimensional variational inequalities and complementarity problems that will be used through the paper. A comprehensive treatment of these problems can be found in the two monographs [32], [34].

II. SYSTEM MODEL AND PROBLEM FORMULATION

We consider a scenario composed of P primary users and Q secondary users, each formed by a single transmitter-receiver pair, coexisting in the same area and sharing the same band (see Fig. 1). We focus on block transmissions over SISO frequency-selective channels. It is well known that, in such a case, multicarrier transmission is capacity achieving for large block-

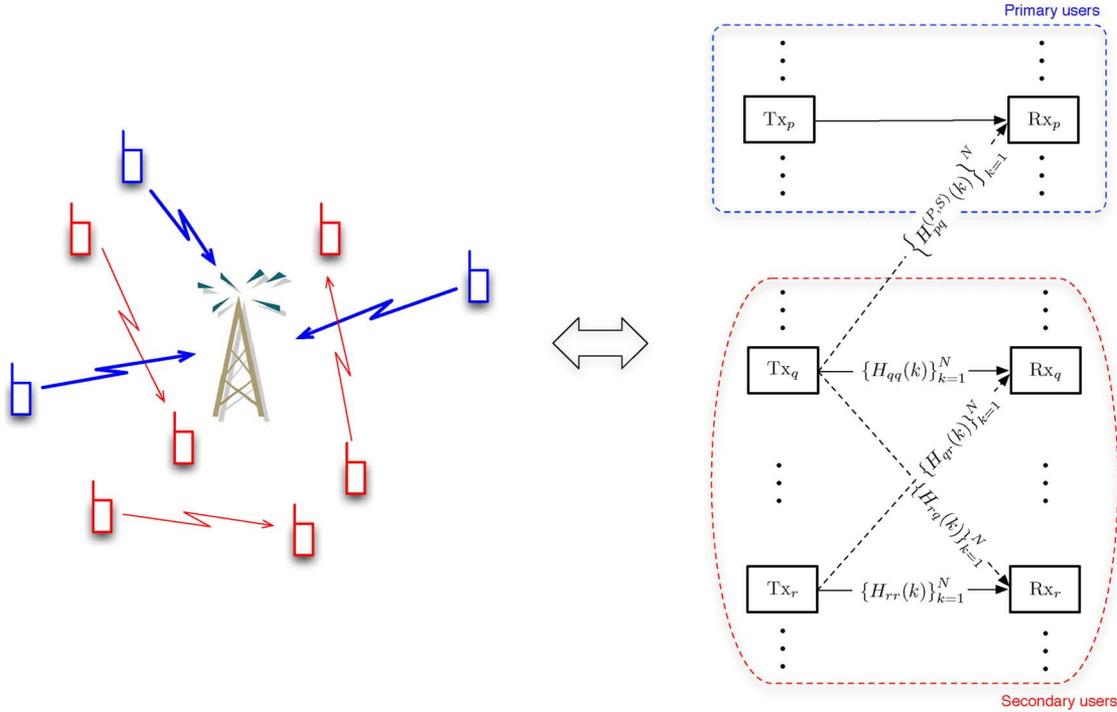


Fig. 1. System model of a hierarchical CR system with primary users (uplink cellular system in blue) and secondary users (red pairs).

length [35]. The systems coexisting in the network do not cooperate with each other, and no centralized authority is assumed to handle the network access for the secondary users. Hence, it is natural to model the set of cognitive secondary users as a frequency-selective N -parallel interference channel, where N is the number of available subcarriers. The transmission strategy of each secondary user q is then the power allocation vector $\mathbf{p}_q = \{p_q(k)\}_{k=1}^N$ over the N subcarriers, subject to the following (local) transmit power constraints:

$$\hat{\mathcal{P}}_q \triangleq \left\{ \mathbf{p} \in \mathbb{R}^N : \sum_{k=1}^N p(k) \leq P_q, \quad \mathbf{0} \leq \mathbf{p} \leq \mathbf{p}_q^{\max} \right\} \quad (1)$$

where we also included local spectral mask constraints $\mathbf{p}_q^{\max} = (p_q^{\max}(k))_{k=1}^N$ [the vector inequality in (1) is component-wise] that may be imposed by radio regulatory bodies to limit the maximum power spectral density (PSD) that each transmitter can use over a specified band. To avoid a trivial solution, we assume throughout the paper that, for all $1 \leq k \leq N$ and $1 \leq q \leq Q$

$$\begin{aligned} p_q^{\max}(k) < P_q \quad \text{and} \quad P_q < \sum_{k=1}^N p_q^{\max}(k) \\ \implies p_q^{\max}(k) < P_q < \sum_{k'=1}^N p_q^{\max}(k') \end{aligned} \quad (2)$$

otherwise, either the upper bound constraints $p_q(k) \leq p_q^{\max}(k)$ or the total power constraint $\sum_{k=1}^N p(k) \leq P_q$ are redundant.

Due to the distributed nature of the CR systems, with neither a centralized control nor coordination among the secondary users, we focus on transmission techniques where no interference cancellation is performed and the MUI is treated as additive colored noise at each receiver. We also assume that the chan-

nels change sufficiently slowly to be considered fixed during the whole transmission, so that the information theoretical results are meaningful, and that perfect channel state information (CSI) is available at both transmitter and receiver sides of each secondary link. This CSI includes, for each secondary pair q , the channel transfer function $\{H_{qq}(k)\}_{k=1}^N$ of the direct link (but not the cross-channels $\{H_{qr}(k)\}_{k,r \neq q}$ from the other secondary users) as well as the overall PSD of the noise plus MUI at each subcarrier, given by $\sigma_q^2(k) + \sum_{r \neq q} |H_{qr}(k)|^2 p_r(k)$, where $\sigma_q^2(k)$ includes both the thermal noise power (zero-mean circularly symmetric complex Gaussian noise) and the MUI PSD due to the active primary users. To implement some of the proposed algorithms, additional CSI is needed from the secondary users; a discussion on this issue is given in Section IV-C, where practical implementation issues of the algorithms are addressed.

Under the setup above, invoking the capacity expression for the single user Gaussian channel—achievable using random Gaussian codes by all the users—the maximum information rate on link q for a specific power allocation profile $\mathbf{p}_1, \dots, \mathbf{p}_Q$ is [35]¹

$$r_q(\mathbf{p}_q, \mathbf{p}_{-q}) = \sum_{k=1}^N \log \left(1 + \frac{|H_{qq}(k)|^2 p_q(k)}{\sigma_q^2(k) + \sum_{r \neq q} |H_{qr}(k)|^2 p_r(k)} \right) \quad (3)$$

where $\mathbf{p}_{-q} \triangleq (\mathbf{p}_r)_{r \neq q}$ is the set of all the users power allocation vectors, except the q th one.

Temperature-Interference Constraints: Differently from traditional static or centralized spectrum assignment, opportunistic communications in CR systems enable secondary users

¹Observe that a NE is obtained if each user transmits using Gaussian signaling, with a proper power allocation. However, Nash equilibria achievable using non-Gaussian codes may exist. In this paper, we focus only on transmissions using Gaussian codebooks.

to transmit with overlapping spectrum and/or coverage with primary users, provided that the degradation induced on the primary users' performance is null or tolerable [3], [4]. In this paper, we consider the following (deterministic) interference constraints that impose an upper bound on the per-carrier and total aggregate interference (the interference temperature-limit [1], [3]) that can be tolerated by each primary user $p = 1, \dots, P$

$$\sum_{q=1}^Q \sum_{k=1}^N \left| H_{pq}^{(P,S)}(k) \right|^2 p_q(k) \leq P_{p,\text{tot}}^{\text{ave}}$$

$$\sum_{q=1}^Q \left| H_{pq}^{(P,S)}(k) \right|^2 p_q(k) \leq P_{p,k}^{\text{peak}}, \quad \forall k=1, \dots, N \quad (4)$$

where $H_{pq}^{(P,S)}(k)$ is the channel transfer function between the transmitter of the q th secondary user and the receiver of the p th primary user, and $P_{p,\text{tot}}^{\text{ave}}$ and $P_{p,k}^{\text{peak}}$ are the interference temperature limit and the maximum interference over subcarrier k tolerable by the p th primary user, respectively. These limits are chosen by each primary user, according to his QoS requirements (see, e.g., [3]).

Problem Formulation: Within the CR context above, we formulate the optimization problem of the transmission strategies of the secondary users as a Nash equilibrium problem (NEP): Each secondary user aims at maximizing his own rate $r_q(\mathbf{p}_q, \mathbf{p}_{-q})$ under the local power constraints in (1) and the interference constraints in (4). The interference constraints however introduce a coupling among the admissible power allocations of all the players, meaning that the secondary users are not allowed to choose their power allocations individually. To keep the optimization as decentralized as possible while imposing global interference constraints, the proposed idea is to introduce a pricing mechanism, properly controlled by the primary users, through a penalty in the payoff function of each player, so that the interference generated by all the secondary will depend on these prices. Stated in mathematical terms, we have the following NEP:

$$\begin{aligned} \underset{\mathbf{p}_q}{\text{maximize}} \quad & r_q(\mathbf{p}_q, \mathbf{p}_{-q}) - \sum_{p=1}^P \sum_{k=1}^N \lambda_{p,k}^{\text{peak}} \left| H_{pq}^{(P,S)}(k) \right|^2 p_q(k) \\ & - \sum_{p=1}^P \lambda_{p,\text{tot}} \sum_{k=1}^N \left| H_{pq}^{(P,S)}(k) \right|^2 p_q(k) \\ \text{subject to} \quad & \mathbf{p}_q \in \hat{\mathcal{P}}_q \end{aligned} \quad (7)$$

for all $q = 1, \dots, Q$, where $\hat{\mathcal{P}}_q$ is defined in (1), $r_q(\mathbf{p}_q, \mathbf{p}_{-q})$ is defined in (3), and the prices $\lambda_{p,\text{tot}}$ and $\lambda_{p,k}^{\text{peak}} = \{\lambda_{p,k}^{\text{peak}}\}_{k=1}^N$ are chosen such that the following complementary conditions are satisfied:

$$0 \leq \lambda_{p,\text{tot}} \perp P_{p,\text{tot}}^{\text{ave}} - \sum_{q=1}^Q \sum_{k=1}^N \left| H_{pq}^{(P,S)}(k) \right|^2 p_q(k) \geq 0$$

$$\forall p = 1, \dots, P$$

$$0 \leq \lambda_{p,k}^{\text{peak}} \perp P_{p,k}^{\text{peak}} - \sum_{q=1}^Q \left| H_{pq}^{(P,S)}(k) \right|^2 p_q(k) \geq 0$$

$$\forall p = 1, \dots, P, \quad \forall k = 1, \dots, N. \quad (8)$$

In (8), the compact notation $0 \leq a \perp b \geq 0$ means $a \cdot b = 0$, $a \geq 0$, and $b \geq 0$. These constraints state that the per-carrier/total interference constraints must be satisfied together with nonnegative pricing; in addition, they imply that if one constraint is trivially satisfied with strict inequality then the corresponding price should be zero (no punishment is needed in that case). With a slight abuse of terminology, we will refer in the following to the NEP (7) with the complementarity constraints (8) as game \mathcal{G} . The challenging goal is then to find the proper decentralized pricing mechanism guaranteeing that the interference constraints [the complementarity conditions (8)] are satisfied while the secondary users reach an equilibrium.

Remark 1 (Pricing Techniques and Related Works): Pricing in game theory is not a new idea. Various auction and pricing approaches have been proposed in the literature to solve spectrum allocation problems in wireline systems (e.g., [20] and [36]), wireless multiple access (see, e.g., [37] and references therein) and *ad hoc* networks (e.g., [26], [27] and [28]–[31]). The current literature can be divided in two large classes, according to the meaning given to the prices: 1) Works where pricing techniques are introduced as heuristics to incentivize the players to reach more socially efficient Nash equilibria [20], [26], [27], [36], [37]; and 2) works dealing with (competitive/cooperative) economy equilibrium models, where prices are used to quantify in monetary terms the value (or worth) of some goods/service involved in the trading between sellers and buyers [28]–[31] (see [31] for a recent overview of the state-of-the-art results in this context). Moreover, most of the algorithms proposed in the cited papers are centralized, and thus they are not applicable to distributed CR networks. In our formulation instead, prices are used to impose coupling constraints (the interference constraints) to the users in a distributed fashion. In the next section, we will see indeed that the prices introduced in the payoff function of the players in (7) are just the multipliers associated to the coupling constraints (4) of the VI reformulation of the game \mathcal{G} .

As far as the analysis of the proposed game \mathcal{G} is concerned, the coupling among the strategies of the players due to the global interference constraints presents a new challenge for the analysis of this class of Nash games that cannot be addressed using game theoretical models for the interference channel in [20]–[24], [33] or market-equilibrium based models [28]–[31]. For this purpose, we need the framework given by the more advanced theory of finite-dimensional VIs [32] that provides a satisfactory resolution to the game \mathcal{G} , as detailed in the forthcoming sections.

III. GAME WITH EXOGENOUS PRICES

Before analyzing the game \mathcal{G} in (7), (8), we start studying a related game, interesting in its own, where the prices are fixed and chosen *a priori*. This assumption may be motivated by all the CR scenarios where a signaling among the primary and secondary users is not allowed (e.g., the so-called *common model* paradigm [4], [5]) and prices are chosen by the network service provider. Some heuristics in the choice of the prices have been proposed, e.g., in [20] and [36] for DSL systems (see also references therein). The results obtained studying the game with exogenous prices provide the building blocks instrumental to analyze the more involved game \mathcal{G} with endogenous prices.

Given the tuple $\boldsymbol{\gamma} = (\boldsymbol{\gamma}_q)_{q=1}^Q \geq \mathbf{0}$, with each $\boldsymbol{\gamma}_q = (\gamma_q(k))_{k=1}^N$, let $f_q(\mathbf{p}_q, \mathbf{p}_{-q}; \boldsymbol{\gamma})$ denote the payoff function of player q , defined as

$$f_q(\mathbf{p}_q, \mathbf{p}_{-q}; \boldsymbol{\gamma}) \triangleq r_q(\mathbf{p}_q, \mathbf{p}_{-q}) - \sum_{k=1}^N \gamma_q(k) p_q(k) \quad (9)$$

with $r_q(\mathbf{p}_q, \mathbf{p}_{-q})$ given in (3). We then consider the following NEP (game in strategic form) $\mathcal{G}_\boldsymbol{\gamma} = \langle \Omega, (\hat{\mathcal{P}}_q)_{q \in \Omega}, (f_q(\mathbf{p}; \boldsymbol{\gamma}))_{q \in \Omega} \rangle$:

$$(\mathcal{G}_\boldsymbol{\gamma}) : \begin{array}{ll} \underset{\mathbf{p}_q \geq \mathbf{0}}{\text{maximize}} & f_q(\mathbf{p}_q, \mathbf{p}_{-q}) \\ \text{subject to} & \mathbf{p}_q \in \hat{\mathcal{P}}_q \end{array} \quad \forall q = 1, \dots, Q \quad (10)$$

where $\Omega \triangleq \{1, \dots, Q\}$ is the set of players; $\hat{\mathcal{P}}_q$ is the strategy set of player q , defined in (1); and $f_q(\mathbf{p}_q, \mathbf{p}_{-q})$ is the payoff function of player q , defined in (9). Note that game \mathcal{G} in (7) with fixed prices $\boldsymbol{\lambda} \triangleq \{(\boldsymbol{\lambda}_p)_{p=1}^P, \boldsymbol{\lambda}_{\text{tot}}\}$ reduces to game $\mathcal{G}_\boldsymbol{\gamma}$ if the vector $\boldsymbol{\gamma} = \boldsymbol{\gamma}(\boldsymbol{\lambda}) = (\boldsymbol{\gamma}_q(\boldsymbol{\lambda}))_{q=1}^Q$ in (10) is chosen as

$$\gamma_q(k; \boldsymbol{\lambda}) = \sum_{p=1}^P \left| H_{pq}^{(P,S)}(k) \right|^2 \left(\lambda_{p,k}^{\text{peak}} + \lambda_{p,\text{tot}} \right) \quad (11)$$

for $k = 1, \dots, N$ and $q = 1, \dots, Q$.

Games with pricing related to $\mathcal{G}_\boldsymbol{\gamma}$ have been studied in [9], [26], [27], [38]. However, only numerical results are available to support the uniqueness of the NE and the convergence of the IWFA with pricing [26]. Furthermore, none of the theoretical results proposed in the literature to solve scalar and vector power control games, such as the approaches based on the theory of standard functions [38]–[40] or the reformulation of the Nash game (with no prices) as a Linearly Complementary Problem [21] (or, equivalently, the interpretation of the waterfilling solution as a projection [18], [23], [24]) are useful for the analysis of game $\mathcal{G}_\boldsymbol{\gamma}$. In this section, we fill this gap and provide a detailed analysis of $\mathcal{G}_\boldsymbol{\gamma}$, valid for arbitrary but fixed tuple $\boldsymbol{\gamma} \geq \mathbf{0}$ [not necessarily in the form (11)]. Our approach provides a new line of analysis of vector power control games with prices that is expected to be broadly applicable to other game models. To study game $\mathcal{G}_\boldsymbol{\gamma}$, we provide first some intermediate definitions and results.

A. Intermediate Definitions and Results

In order to rewrite the game $\mathcal{G}_\boldsymbol{\gamma}$ (and \mathcal{G}) as a VI problem, we introduce the joint admissible strategy set \mathcal{P} , defined as shown in (12) at the bottom of the page, with

$$\hat{\mathcal{P}} = \hat{\mathcal{P}}_1 \times \dots \times \hat{\mathcal{P}}_Q \quad (13)$$

and the vector function $\mathbf{F}(\mathbf{p}) \triangleq (\mathbf{F}_1(\mathbf{p})^T, \dots, \mathbf{F}_Q(\mathbf{p})^T)^T$, where each $\mathbf{F}_q(\mathbf{p})$ is defined as

$$\mathbf{F}_q(\mathbf{p}) \triangleq -\nabla_{\mathbf{p}_q} r_q(\mathbf{p}_q, \mathbf{p}_{-q}) \quad (14)$$

$$= \left(-\frac{1}{\hat{\sigma}_q^2(k) + \sum_{r=1}^Q \left| \hat{H}_{qr}(k) \right|^2 p_r(k)} \right)_{k=1}^N, \quad (15)$$

with $\nabla_{\mathbf{p}_q} r_q(\mathbf{p}_q, \mathbf{p}_{-q})$ denoting the gradient (column) vector of $r_q(\mathbf{p}_q, \mathbf{p}_{-q})$ with respect to \mathbf{p}_q , and

$$\hat{H}_{qr}(k) \triangleq \frac{H_{qr}(k)}{H_{qq}(k)}, \quad \hat{\sigma}_q^2(k) \triangleq \frac{\sigma_q^2(k)}{|H_{qq}(k)|^2}. \quad (16)$$

In what follows, we provide sufficient conditions for the vector function \mathbf{F} to be either a *strongly monotone* function [32, Def. 2.3.1(e)] or a *uniformly P-function* on $\hat{\mathcal{P}}$ (and \mathcal{P}) [32, Def. 3.5.8(b)]. This result will be instrumental to obtain conditions guaranteeing the uniqueness of the NE of game $\mathcal{G}_\boldsymbol{\gamma}$ (and the uniqueness of the NE power allocation of game \mathcal{G}) as well as the global convergence of the proposed iterative algorithms.

Definition 1: The mapping \mathbf{F} in (14) is said to be (a) *strongly monotone* on $\hat{\mathcal{P}}$ (or \mathcal{P}) if there exists a constant $c_{\text{sm}} > 0$ such that for all pairs $\mathbf{p} = (\mathbf{p}_q)_{q=1}^Q$ and $\mathbf{p}' = (\mathbf{p}'_q)_{q=1}^Q$ in $\hat{\mathcal{P}}$ (or \mathcal{P}),

$$(\mathbf{p} - \mathbf{p}')^T (\mathbf{F}(\mathbf{p}) - \mathbf{F}(\mathbf{p}')) \geq c_{\text{sm}} \|\mathbf{p} - \mathbf{p}'\|^2. \quad (17)$$

The constant c_{sm} is called the strong monotonicity constant.

(b) a *uniformly P-function* on $\hat{\mathcal{P}}$ (or \mathcal{P}) if there exists a constant $c_{\text{uP}} > 0$ such that for all pairs $\mathbf{p} = (\mathbf{p}_q)_{q=1}^Q$ and $\mathbf{p}' = (\mathbf{p}'_q)_{q=1}^Q$ in $\hat{\mathcal{P}}$ (or \mathcal{P}),

$$\max_{1 \leq q \leq Q} (\mathbf{p}_q - \mathbf{p}'_q)^T (\mathbf{F}_q(\mathbf{p}) - \mathbf{F}_q(\mathbf{p}')) \geq c_{\text{uP}} \|\mathbf{p} - \mathbf{p}'\|^2. \quad (18)$$

□

We also need the definition of Z-matrix [34, Def. 3.3.1] and P-matrix [34, Def. 3.11.1], as given next.

Definition 2: A matrix $\mathbf{M} \in \mathbb{R}^{n \times n}$ is called Z-matrix if its off-diagonal entries are all nonpositive. A matrix $\mathbf{M} \in \mathbb{R}^{n \times n}$ is called P-matrix if every principal minor of \mathbf{M} is positive. A Z-matrix that is also P-matrix is called a K-matrix. □

Many equivalent characterizations for a P-matrix can be given. The interested reader is referred to [34], [41] for more details. Here we note that any positive definite matrix is a P-matrix, but the reverse does not hold, and we recall a characterizing property of a K-matrix relevant to our forthcoming derivations, as given in the following lemma (proved in [34, Lemma 5.13.14]).

$$\mathcal{P} \triangleq \hat{\mathcal{P}} \cap \left\{ \mathbf{p} : \begin{array}{l} \sum_{q=1}^Q \sum_{k=1}^N \left| H_{pq}^{(P,S)}(k) \right|^2 p_q(k) \leq P_{p,\text{tot}}^{\text{ave}}, \quad \forall p = 1, \dots, P \\ \sum_{q=1}^Q \left| H_{pq}^{(P,S)}(k) \right|^2 p_q(k) \leq P_{p,k}^{\text{peak}}, \quad \forall p = 1, \dots, P, \forall k = 1, \dots, N \end{array} \right\} \quad (12)$$

Lemma 1: Let $\mathbf{M} \in \mathbb{R}^{n \times n}$ be a K-matrix, $\mathbf{N} \in \mathbb{R}^{n \times n}$ a nonnegative matrix, and let $\rho(\mathbf{A})$ denote the spectral radius of \mathbf{A} .² Then $\rho(\mathbf{M}^{-1}\mathbf{N}) < 1$ if and only if $\mathbf{M} - \mathbf{N}$ is a K-matrix. \square

Finally, we introduce the Z-matrices $\mathbf{Y}(k), \hat{\mathbf{Y}}, \tilde{\mathbf{Y}} \in \mathbb{R}^{Q \times Q}$, useful to study the monotonicity properties of the mapping \mathbf{F} and the uniqueness of the NE of the proposed games:

$$[\mathbf{Y}(k)]_{qr} \triangleq \begin{cases} 1, & \text{if } q=r \\ -\frac{|H_{qr}(k)|^2}{|H_{qq}(k)|^2}, & \text{if } q \neq r, \end{cases} \quad k=1, \dots, N \quad (19)$$

$$[\hat{\mathbf{Y}}]_{qr} \triangleq \begin{cases} 1 & \text{if } q=r \\ -\max_{1 \leq k \leq N} \left\{ \frac{|H_{qr}(k)|^2}{|H_{rr}(k)|^2} \cdot \widehat{\text{innr}}_{qr}(k) \right\} & \text{if } q \neq r \end{cases} \quad (20)$$

with

$$\widehat{\text{innr}}_{qr}(k) \triangleq \frac{\sigma_r^2(k) + \sum_{r'} |H_{rr'}(k)|^2 p_r^{\max}(k)}{\sigma_q^2(k)} \quad (21)$$

and

$$[\tilde{\mathbf{Y}}]_{qr} \triangleq \begin{cases} 1, & \text{if } q=r \\ -\max_{1 \leq k \leq N} \left\{ \frac{|H_{qr}(k)|^2}{|H_{rr}(k)|^2} \cdot \widetilde{\text{innr}}_{qr}(k) \right\}, & \text{if } q \neq r \end{cases} \quad (22)$$

with

$$\widetilde{\text{innr}}_{qr}(k) \triangleq \widehat{\text{innr}}_{qr}(k) \cdot \sqrt{1 + \sum_{r'} |H_{qr'}(k)|^2 \frac{p_r^{\max}(k)}{\sigma_q^2(k)}}. \quad (23)$$

Note that $\tilde{\mathbf{Y}} \leq \hat{\mathbf{Y}}$ (componentwise). The next proposition provides sufficient conditions for the map \mathbf{F} to be a strongly monotone or a uniformly P-function on the set $\hat{\mathcal{P}}$ (and \mathcal{P}).

Proposition 2: Given $\mathcal{P}, \hat{\mathcal{P}}, \mathbf{F}$, and $\hat{\mathbf{Y}}$, defined in (12), (13), (14), and (20), respectively, the following hold:

- (a) If $\hat{\mathbf{Y}}$ is a P-matrix, then \mathbf{F} is a uniformly P-function on $\hat{\mathcal{P}}$ (and thus also on \mathcal{P});
- (b) If $\hat{\mathbf{Y}} \succ \mathbf{0}$ (positive definite), then \mathbf{F} is a strongly monotone function on $\hat{\mathcal{P}}$ (and thus also on \mathcal{P}).

Proof: See Appendix B. \blacksquare

It follows readily from the proof of the proposition that a lower bound of the strong monotonicity constant of \mathbf{F} on $\hat{\mathcal{P}}$ is [see (60) in Appendix B)]

$$\hat{c}_{\text{sm}}(\mathbf{F}) = \frac{\lambda_{\min}(\hat{\mathbf{Y}})}{\max_{1 \leq q \leq Q} \max_{1 \leq k \leq N} (\hat{\varphi}_q^{\max}(k))^2} \quad (24)$$

with $\lambda_{\min}(\hat{\mathbf{Y}}) > 0$ denoting the smallest eigenvalue of the symmetric part of $\hat{\mathbf{Y}}$, and $\hat{\varphi}_q^{\max}(k) \triangleq \hat{\sigma}_q^2(k) + \sum_{r=1}^Q |\hat{H}_{qr}(k)|^2 p_r^{\max}(k)$. A lower bound of the uniformly P-constant of \mathbf{F} can be similarly obtained, based on [34, Ex. 5.11.19]. The P-property of matrix $\hat{\mathbf{Y}}$ will be used in Section III-B to prove the uniqueness of the NE of the game \mathcal{G}_γ as well as the convergence of the proposed distributed algorithms; whereas the positive definite property of $\hat{\mathbf{Y}}$ will be exploited in Section IV where the game with endogenous

²The spectral radius of a square matrix \mathbf{A} is defined as $\rho(\mathbf{A}) \triangleq \max\{|\lambda| : \lambda \in \sigma(\mathbf{A})\}$, with $\sigma(\mathbf{A})$ denoting the set of eigenvalues of \mathbf{A} [42].

prices is studied. The constant $\hat{c}_{\text{sm}}(\mathbf{F})$ in (24) has important roles in studying the convergence of the distributed algorithms proposed in the forthcoming sections.

Finally, to write the solutions of \mathcal{G}_γ in a convenient form, we introduce, for each q and $\gamma \geq \mathbf{0}$, the waterfilling-like mapping $\text{wf}_q(\mathbf{p}_{-q}; \gamma) : \hat{\mathcal{P}}_{-q} \ni \mathbf{p}_{-q} \rightarrow \hat{\mathcal{P}}_q$, defined as: for $k = 1, \dots, N$

$$(\text{wf}_q(\mathbf{p}_{-q}; \gamma))_k \triangleq \left[\frac{1}{\mu_q + \gamma_q(k)} - \frac{\sigma_q^2(k) + \sum_{r \neq q} |H_{qr}(k)|^2 p_r(k)}{|H_{qq}(k)|^2} \right]_0^{p_q^{\max}(k)} \quad (25)$$

where $[x]_a^b = \min(b, \max(a, x))$ with $0 \leq a \leq b$ and $\mu_q \geq 0$ is chosen to satisfy the power constraint $\sum_{k=1}^N (\text{wf}_q(\mathbf{p}_{-q}; \gamma))_k \leq P_q$ ($\mu_q = 0$ if the inequality is strictly satisfied). Practical algorithms to compute the water-level μ_q can be found in [43].

B. Nash Equilibria of Game \mathcal{G}_γ

We can now study the game \mathcal{G}_γ in (10) and obtain sufficient conditions guaranteeing the uniqueness of the NE, as given next (basic concepts and results on VIs used extensively through the paper are given in Appendix A).

Theorem 3 (Existence and Uniqueness of the Nash Equilibria): Given $\gamma \geq \mathbf{0}$, consider the game \mathcal{G}_γ in (10) and suppose w.l.o.g. that conditions in (2) are satisfied. Then, the following hold:

- (a) The game is equivalent to the VI($\hat{\mathcal{P}}, \mathbf{F} + \gamma$), which always admits a solution, for any given set of channel matrices, power constraints of the users, and $\gamma \geq \mathbf{0}$. Every NE solution $\mathbf{p}^*(\gamma) = (\mathbf{p}_q^*(\gamma))_{q=1}^Q$ satisfies the following vector waterfilling-like fixed-point equation:

$$\mathbf{p}_q^*(\gamma) = \text{wf}_q(\mathbf{p}_{-q}^*(\gamma); \gamma), \quad q = 1, \dots, Q \quad (26)$$

with wf_q defined in (25).

- (b) If $\hat{\mathbf{Y}}$ is a P-matrix (which is the case if $\hat{\mathbf{Y}}$ is a positive definite matrix), then the NE of \mathcal{G}_γ is unique.

Proof: (a): By the convexity and the first-order (necessary and sufficient) optimality conditions of each of the optimization problems in (10), we know that \mathbf{p}^* is a NE of game \mathcal{G}_γ if and only if for each $q = 1, \dots, Q$

$$(\mathbf{p}_q - \mathbf{p}_q^*)^T (\mathbf{F}_q(\mathbf{p}^*) + \gamma_q) \geq 0 \quad \forall \mathbf{p}_q \in \hat{\mathcal{P}}_q. \quad (27)$$

The set of inequalities above is equivalent to the VI($\hat{\mathcal{P}}, \mathbf{F} + \gamma$) (cf. Appendix A).

The existence of a NE of \mathcal{G}_γ —a solution to VI($\hat{\mathcal{P}}, \mathbf{F} + \gamma$)—follows from [32, Corollary 2.2.5]: the set $\hat{\mathcal{P}}$ is compact and convex and the function \mathbf{F} is continuous.

The waterfilling-like structure of the Nash equilibria as given in (26) follows directly from the fact that, for every q and fixed \mathbf{p}_{-q} , the waterfilling function $\text{wf}_q(\mathbf{p}_{-q}; \gamma)$ in (25) is the unique solution to the q th (strictly convex) optimization problem in (10).

- (b): If $\hat{\mathbf{Y}}$ is a P-matrix, then it follows from Proposition 2(a) that the function \mathbf{F} (and thus also $\mathbf{F} + \gamma$) is a uniformly P-function on $\hat{\mathcal{P}}$, which, together with the Cartesian structure of $\hat{\mathcal{P}}$ (i.e.,

$\widehat{\mathbf{P}} = \widehat{\mathbf{P}}_1 \times \cdots \times \widehat{\mathbf{P}}_Q$), guarantees the uniqueness of the NE [32, Prop. 3.5.10(a)] (cf. Appendix A). ■

To give additional insight into the physical interpretation of the uniqueness conditions of the NE of \mathcal{G}_γ we provide next sufficient conditions for the matrix $\widehat{\mathbf{Y}}$ to be a P-matrix (or positive definite matrix). It follows from Lemma 1 (with $\mathbf{M} = \mathbf{I}$ and $\mathbf{N} = \mathbf{I} - \widehat{\mathbf{Y}}$) that $\widehat{\mathbf{Y}}$ (a Z-matrix) is a P-matrix (and thus a K-matrix) if and only if $\rho(\mathbf{I} - \widehat{\mathbf{Y}}) < 1$, which leads to the following.

Corollary 4: The matrix $\widehat{\mathbf{Y}}$ is a P-matrix (or a positive definite matrix) if one (or both) of the following two sets of conditions are satisfied:

Low received MUI:

$$\frac{1}{w_q} \sum_{r \neq q} w_r \max_k \left\{ \frac{|H_{qr}(k)|^2}{|H_{rr}(k)|^2} \cdot \widehat{\text{innr}}_{qr}(k) \right\} < 1, \quad \forall q = 1, \dots, Q \quad (\text{C1})$$

Low generated MUI:

$$\frac{1}{w_r} \sum_{q \neq r} w_q \max_k \left\{ \frac{|H_{qr}(k)|^2}{|H_{rr}(k)|^2} \cdot \widehat{\text{innr}}_{qr}(k) \right\} < 1, \quad \forall r = 1, \dots, Q \quad (\text{C2})$$

where $\mathbf{w} = (w_q)_{q=1}^Q$ is some positive vector. □

One can also obtain milder conditions for the uniqueness of the NE of game \mathcal{G}_γ following the approach proposed in [22]. The main result is stated next (the proof follows similar steps as in [22, Th. 2] and is omitted, because of the space limitation).

Corollary 5: If each matrix $\mathbf{Y}(k)$ defined in (19) is a P-matrix [or equivalently $\rho(\mathbf{I} - \mathbf{Y}(k)) < 1$, for all $k = 1, \dots, N$], then the NE of \mathcal{G}_γ is unique. □

Remark 2 (Physical Interpretation of the Uniqueness Conditions): Conditions (C1) and (C2) provide a physical interpretation of the uniqueness of the NE: the uniqueness of the NE is ensured for any given γ if the interference among the links is sufficiently small and, interestingly, is not affected by the values of the prices γ . The importance of the above conditions is that they quantify how small the interference must be to guarantee that the equilibrium is indeed unique. Specifically, conditions (C1) can be interpreted as a constraint on the maximum amount of interference that each receiver can tolerate, whereas (C2) introduce an upper bound on the maximum level of interference that each transmitter is allowed to generate. We show next that the uniqueness conditions in Theorem 3 are also sufficient for the convergence of the proposed distributed iterative algorithms, based on the waterfilling best-response (25).

C. Distributed Algorithms

We focus now on distributed algorithms that converge to the NE of game \mathcal{G}_γ . We consider iterative algorithms based on both sequential and simultaneous updates of the power allocation vectors of the users according to the waterfilling-like best-response in (25), and prove their global convergence under the same conditions guaranteeing the uniqueness of the NE of game \mathcal{G}_γ . Interestingly, the proposed algorithms can also be implemented in a totally asynchronous way.

Synchronous Implementation: The sequential (or the simultaneous) iterative algorithm we propose is an instance of the Gauss-Seidel (or Jacobi) scheme: Each player, sequentially (or simultaneously) and according to a fixed updating order, solves

problem (10), performing the single-user water-filling solution in (25). Denoting by $\mathbf{p}_q^{(n)}(\gamma)$ the power allocation vector of user q at the n th iteration, the sequential IWFA with pricing is formally described in Algorithm 1 (the simultaneous IWFA follows similarly). A unified set of convergence conditions for both sequential and simultaneous versions of the algorithm is given in Theorem 6.

Algorithm 1: Sequential IWFA with pricing

Data: Choose any $\mathbf{p}_q^{(0)} \in \widehat{\mathcal{P}}_q$ for all $q = 1, \dots, Q$, and set $n = 0$.

Step 1: If $\mathbf{p}^{(n)}$ satisfies a suitable termination criterion:

STOP

Step 2: Sequentially for $q = 1, \dots, Q$, compute $\mathbf{p}_q^{(n+1)}(\gamma)$ as

$$\mathbf{p}_q^{(n+1)}(\gamma) = \text{wf}_q \left(\mathbf{p}_1^{(n+1)}, \dots, \mathbf{p}_{q-1}^{(n+1)}, \mathbf{p}_{q+1}^{(n)}, \dots, \mathbf{p}_Q^{(n)}; \gamma \right) \quad (28)$$

Step 3: Set $n \leftarrow n + 1$; and go to Step 1.

Theorem 6: Suppose that $\widehat{\mathbf{Y}}$ defined in (20) is a P-matrix. Then, any sequence $\{\mathbf{p}^{(n)}(\gamma)\}_{n=0}^\infty$ generated by the sequential IWFA (or the simultaneous IWFA) described in Algorithm 1 converges to the unique NE of game \mathcal{G}_γ , for any given updating order of the users and $\gamma \geq \mathbf{0}$.

Proof: See Appendix C. ■

The proposed algorithms have some desirable properties that make them appealing in many practical *ad hoc* systems, namely: *low complexity, distributed nature, and fast convergence behavior*. In fact, given the pricing vector γ_q (recall that in the game \mathcal{G}_γ the pricing vector γ is assumed to be fixed), the optimal power allocations of each user q can be efficiently and locally computed using a waterfilling based solution; which requires only the local measure of the overall interference-plus-noise power spectral density. Moreover, the simultaneous version of the algorithm has been experimented to converges in very few iterations, even in networks with many active secondary users (see Section V for some numerical results) and imposes less stringent constraints on the synchronization among the secondary user than the sequential version.

Observe that the convergence of the algorithms is guaranteed under the same conditions obtained for the uniqueness of the solution of the game (cf. Theorem 3). As expected, the convergence is ensured if the level of interference in the network is not too high (see Remark 2).

Asynchronous Implementation: In a real CR network with many secondary users, the synchronization requirement from both sequential and simultaneous IWFAs might not always be acceptable. It is thus natural to ask whether some asynchronous implementation of the proposed algorithm is still guaranteed to globally converge to the NE of \mathcal{G}_γ , under some sufficient conditions. More specifically, we consider *totally asynchronous* schemes (in the sense specified in [44]), where some secondary users may update their power allocations more frequently than others and they may even use an outdated measurement of the interference caused from the others. The only constraints on the updating schedule performed by the secondary users are that

each user updates his power allocation at least once within any sufficiently large, but finite, time interval and that outdated measurements of the interference are eventually replaced by more recent ones. We say that an updating schedule of the users is feasible if the conditions above are satisfied (see [24] for a formal description of the asynchronous algorithm in a different context). Convergence conditions are given in the following theorem (the proof is omitted because of the space limitation; see [24] for a similar approach).

Theorem 7: Suppose that $\tilde{\mathbf{Y}}$ defined in (22) is a P-matrix. Then, any sequence $\{\mathbf{p}^{(n)}(\boldsymbol{\gamma})\}_{n=0}^{\infty}$ generated by the totally asynchronous IWFA [44] and based on the mapping $\text{wf}(\mathbf{p}; \boldsymbol{\gamma}) = (\text{wf}_q(\mathbf{p}; \boldsymbol{\gamma}))_{q=1}^Q$ defined in (25) converges to the unique NE of game $\mathcal{G}_{\boldsymbol{\gamma}}$, for any given feasible updating schedule of the users and $\boldsymbol{\gamma} \geq \mathbf{0}$. \square

Note that the asynchronous IWFA contains as special cases a plethora of algorithms, each one obtained by a possible choice of the scheduling of the users in the updating procedure. The sequential IWFA given in Algorithm 1 and the simultaneous IWFA are indeed two special cases. The important result stated in Theorem 7 is that all the algorithms resulting as special cases of the asynchronous IWFA are guaranteed to reach the unique NE of the game, under the same set of convergence conditions, since the entries of matrix $\tilde{\mathbf{Y}}$ do not depend on the particular choice of the updating schedule. This also means that this class of algorithms is robust against missing or outdated updates of the secondary users; which strongly relaxes the constraints on the network synchronization.

The algorithms proposed so far may suffer of the main drawback of classical IWFA (e.g., [17]), i.e., the violation of the interference temperature limits [3], if the pricing vector $\boldsymbol{\gamma}$ is not properly chosen. The goal of the next section is indeed to show how to design the prices in order to preserve the QoS of the primary users, keeping the proposed algorithms as decentralized as possible.

IV. GAME WITH ENDOGENOUS PRICES

We focus now on game \mathcal{G} defined in (7) and, using results obtained for game $\mathcal{G}_{\boldsymbol{\gamma}}$ in (10), we provide sufficient conditions guaranteeing the uniqueness of the solution of \mathcal{G} and propose a variety of iterative algorithms along with their convergence properties. To this end, we introduce the following preliminary definitions and results.

Under the P-property of matrix $\hat{\mathbf{Y}}$, we have proved in Theorem 3(b) that the NE $\mathbf{p}^*(\boldsymbol{\gamma})$ of game $\mathcal{G}_{\boldsymbol{\gamma}}$ is unique, for any given $\boldsymbol{\gamma} \geq \mathbf{0}$. Under this condition, let us choose $\boldsymbol{\gamma} = \boldsymbol{\gamma}(\boldsymbol{\lambda})$ satisfying (11) and define the map

$$\Phi : \boldsymbol{\lambda} = \begin{pmatrix} (\boldsymbol{\lambda}_p)_{p=1}^P \\ \boldsymbol{\lambda}_{\text{tot}} \end{pmatrix} \mapsto \begin{pmatrix} \left(\left(P_{p,k}^{\text{peak}} - \sum_{q=1}^Q |H_{pq}^{(P,S)}(k)|^2 p_q^*(k; \boldsymbol{\gamma}(\boldsymbol{\lambda})) \right)_{k=1}^N \right)_{p=1}^P \\ \left(P_{p,\text{tot}}^{\text{ave}} - \sum_{q=1}^Q \sum_{k=1}^N |H_{pq}^{(P,S)}(k)|^2 p_q^*(k; \boldsymbol{\gamma}(\boldsymbol{\lambda})) \right)_{p=1}^P \end{pmatrix} \quad (29)$$

which measures the violation of the temperature-interference constraints. We also introduce the nonnegative matrix $\mathbf{H} \in \mathbb{R}^{P(N+1) \times NQ}$, defined as

$$\mathbf{H} \triangleq [\boldsymbol{\Sigma}_1 \dots \boldsymbol{\Sigma}_Q; \mathbf{D}_{1,1} \dots \mathbf{D}_{1,Q}; \dots; \mathbf{D}_{P,1} \dots \mathbf{D}_{P,Q}] \quad (30)$$

with $\mathbf{D}_{p,q} = \text{diag}(|H_{pq}^{(P,S)}(k)|^2)_{k=1}^N$ and $[\boldsymbol{\Sigma}_q]_{ij} \triangleq |H_{iq}^{(P,S)}(j)|^2$, for $1 \leq i \leq P$ and $1 \leq j \leq N$.

To study game \mathcal{G} , we need the following technical property of $\Phi(\boldsymbol{\lambda})$.

Proposition 8: Suppose $\hat{\mathbf{Y}} \succ \mathbf{0}$. Then the map $\Phi(\boldsymbol{\lambda})$ in (29) is a co-coercive function of $\boldsymbol{\lambda} \in \mathbb{R}_+^{P(N+1)}$ with modulus $c_{\text{coc}}(\Phi)$, i.e.

$$\begin{aligned} (\boldsymbol{\lambda}^{(1)} - \boldsymbol{\lambda}^{(2)})^T (\Phi(\boldsymbol{\lambda}^{(1)}) - \Phi(\boldsymbol{\lambda}^{(2)})) &\geq \\ c_{\text{coc}}(\Phi) \|\Phi(\boldsymbol{\lambda}^{(1)}) - \Phi(\boldsymbol{\lambda}^{(2)})\|_2^2, \quad \forall \boldsymbol{\lambda}^{(1)}, \boldsymbol{\lambda}^{(2)} \in \mathbb{R}_+^{P(N+1)} \end{aligned} \quad (31)$$

with

$$c_{\text{coc}}(\Phi) \triangleq \frac{\hat{c}_{\text{sm}}(\mathbf{F})}{\|\mathbf{H}\|_2^2} \quad (32)$$

where $\|\mathbf{H}\|_2$ denotes the spectral norm of \mathbf{H} defined in (30), and $\hat{c}_{\text{sm}}(\mathbf{F})$ is defined in (24).

Proof: See Appendix D. \blacksquare

Note that the co-coercivity of Φ implies that Φ is Lipschitz continuous with modulus $c_{\text{Lip}}(\Phi) \triangleq 1/c_{\text{coc}}(\Phi)$, i.e.

$$\|\Phi(\boldsymbol{\lambda}^{(1)}) - \Phi(\boldsymbol{\lambda}^{(2)})\| \leq c_{\text{Lip}}(\Phi) \|\boldsymbol{\lambda}^{(1)} - \boldsymbol{\lambda}^{(2)}\| \quad (33)$$

for all $\boldsymbol{\lambda}^{(1)}, \boldsymbol{\lambda}^{(2)} \in \mathbb{R}_+^{P(N+1)}$. Using the above results, we focus now on game \mathcal{G} in (7), (8).

A. Equilibrium Solutions

The main properties of the solutions of the game \mathcal{G} are given in the following, where \mathcal{P} , \mathbf{F} , $\hat{\mathbf{Y}}$, and Φ are defined in (12), (14), (20), and (29), respectively, and $\mathbf{p}^*(\boldsymbol{\gamma}(\boldsymbol{\lambda}))$ denotes the unique NE of $\mathcal{G}_{\boldsymbol{\gamma}(\boldsymbol{\lambda})}$ (under the P-property of $\hat{\mathbf{Y}}$), with $\boldsymbol{\gamma}(\boldsymbol{\lambda})$ given in (11).

Theorem 9: Consider the game \mathcal{G} and suppose w.l.o.g. that conditions (2) are satisfied. Then, the following hold:

- The game \mathcal{G} is equivalent to the VI(\mathcal{P} , \mathbf{F}), which always admits a solution, denoted by \mathbf{p}^{VI} . The equivalence is in the following sense: if \mathbf{p}^{VI} is a solution of the VI, then there exists $\boldsymbol{\lambda}^{\text{VI}}$ —the multipliers of the VI associated with the interference constraints (4)—such that $(\mathbf{p}^{\text{VI}}, \boldsymbol{\lambda}^{\text{VI}})$ is an equilibrium pair of \mathcal{G} ; conversely, if $(\mathbf{p}^{\text{NE}}, \boldsymbol{\lambda}^{\text{NE}})$ is an equilibrium pair of \mathcal{G} , then \mathbf{p}^{NE} is a solution of the VI, and $\boldsymbol{\lambda}^{\text{NE}}$ are multipliers of the VI;
- If $\hat{\mathbf{Y}}$ is a P-matrix, then the game \mathcal{G} is equivalent to the non-linear complementarity problem in the price tuple $\boldsymbol{\lambda}$

$$\text{NCP}(\Phi) : \quad \mathbf{0} \leq \boldsymbol{\lambda} \perp \Phi(\boldsymbol{\lambda}) \geq \mathbf{0}. \quad (34)$$

The equivalence is in the following sense: the NCP(Φ) must have a solution and, for any such solution λ^{NCP} , the pair $(\mathbf{p}^*(\gamma(\lambda^{\text{NCP}})), \lambda^{\text{NCP}})$ is an equilibrium of \mathcal{G} ; conversely, if $(\mathbf{p}^{\text{NE}}, \lambda^{\text{NE}})$ is an equilibrium pair of \mathcal{G} , then λ^{NE} is a solution of the NCP(Φ) with $\mathbf{p}^*(\gamma(\lambda^{\text{NE}})) = \mathbf{p}^{\text{NE}}$;

(c) If $\hat{\mathbf{Y}}$ is positive definite, then

(c1) The solution \mathbf{p}^{VI} of the VI(\mathcal{P}, \mathbf{F}) is unique, and it must be $\mathbf{p}^{\text{VI}} = \mathbf{p}^*(\gamma(\lambda^{\text{NCP}}))$, for any solution λ^{NCP} of the NCP(Φ);

(c2) The game \mathcal{G} has a unique least-norm price tuple, denoted by $\lambda^{\text{NE,lm}}$, such that $\|\lambda^{\text{NE,lm}}\|_2 \leq \|\lambda^{\text{NE}}\|_2$, for any price solution λ^{NE} of the game.

Proof: (a) With each $\hat{\mathcal{P}}_q$ being a convex, in particular, a polyhedron, the Nash problem (7) is equivalent to its KKT optimality conditions, which, together with the complementarity conditions in (8), are given in (35) at the bottom of the page (cf. Appendix A), where $\zeta_{q,k}$ denote the multipliers of the mask constraints and μ_q are the multipliers of the transmit power constraints. It follows from [32, Prop. 1.3.4] (see also (49) in Appendix A) that the mixed nonlinear complementary problem above is just the KKT system of the VI(\mathcal{P}, \mathbf{F}), where $\lambda_{p,k}^{\text{peak},s}$ and $\lambda_{p,\text{tot}}$'s are the multipliers of the interference constraints (4). The VI(\mathcal{P}, \mathbf{F}) always admits a solution, since the set \mathcal{P} is convex and compact, and the function $\mathbf{F}(\mathbf{p}) : \mathcal{P} \ni \mathbf{p} \mapsto \mathbb{R}^{N_Q}$ is continuous [32, Corollary 2.2.5].

(b) Let $(\mathbf{p}^{\text{NE}}, \lambda^{\text{NE}})$ be an equilibrium pair of \mathcal{G} [whose existence is guaranteed by statement (a)]. Then, \mathbf{p}^{NE} must be the unique NE of game $\mathcal{G}_{\gamma(\lambda^{\text{NE}})}$, i.e., $\mathbf{p}^*(\gamma(\lambda^{\text{NE}})) = \mathbf{p}^{\text{NE}}$ [the uniqueness follows from Theorem 3(b) and the P-property of $\hat{\mathbf{Y}}$]. Hence, λ^{NE} is a solution of the NCP(Φ). Conversely, if λ^{NCP} is any solution of the NCP(Φ), then $(\mathbf{p}^*(\gamma(\lambda^{\text{NCP}})), \lambda^{\text{NCP}})$ must be an equilibrium of \mathcal{G} .

(c) The first part of (c1) follows from the strong monotonicity of \mathbf{F} on $\hat{\mathcal{P}} \supseteq \mathcal{P}$ (and thus also on \mathcal{P}) [32, Th. 2.3.3(b)], as stated in Proposition 2(b) under $\hat{\mathbf{Y}} \succ \mathbf{0}$. The latter part of (c1) follows readily from statement (b) and the uniqueness of the solution \mathbf{p}^{VI} of the VI(\mathcal{P}, \mathbf{F}). Statement (c2) follows from the fact that, under $\hat{\mathbf{Y}} \succ \mathbf{0}$, the NCP(Φ) is monotone (the function Φ is co-coercive on $\mathbb{R}_+^{P(N+1)}$ and thus monotone, see Proposi-

tion 8) and solvable, which is sufficient for the NCP(Φ) to have a convex solution set [32, Th. 2.3.5(a)]. As a consequence, the NCP(Φ) must have a unique least-norm solution. ■

A graphical representation of the main results in Theorem 9 as well as the relationships among the various games and the associated VI reformulations is given in Fig. 2.

Remark 3 (On the Uniqueness of the Equilibrium Solution): Observe that the uniqueness of the solution of the VI(\mathcal{P}, \mathbf{F}) as stated in Theorem 9(c1) implies only the uniqueness of the power allocation \mathbf{p}^{NE} of the secondary users at the NE of \mathcal{G} , but not of the price tuple λ . The interesting result is that, in such a case, all these prices λ^{NCP} —the solutions of the NCP(Φ)—yield equilibrium pairs $(\mathbf{p}^*(\gamma(\lambda^{\text{NCP}})), \lambda^{\text{NCP}})$ of game \mathcal{G} having the same optimal power allocation \mathbf{p}^{NE} , i.e., $\mathbf{p}^*(\gamma(\lambda^{\text{NCP}})) = \mathbf{p}^{\text{NE}}$. Nevertheless, part (c2) identifies a special price tuple that motivates a distributed algorithm for solving the game \mathcal{G} whose convergence can be established under $\hat{\mathbf{Y}} \succ \mathbf{0}$ (see Section IV-B).

Remark 4 (On the Uniqueness Conditions): Sufficient conditions for matrix $\hat{\mathbf{Y}}$ being a P-matrix (or a positive definite matrix) are given in Corollary 4. We thus refer to this corollary and to Remark 2 for a discussion on the physical interpretation of these conditions. Note that conditions given in Theorem 9 guaranteeing the uniqueness of the solution of the VI(\mathcal{P}, \mathbf{F}) (and thus the uniqueness of the optimal power allocation of game \mathcal{G}) are stronger than those required by Theorem 3 for the uniqueness of the solution of VI($\hat{\mathcal{P}}, \mathbf{F}$) (and thus the NE of game \mathcal{G}_γ , for any given $\gamma \geq \mathbf{0}$). This is due to the fact that the uniform P-property of the mapping \mathbf{F} on \mathcal{P} (Proposition 2) does not directly yield the desired uniqueness result of the solution of VI($\hat{\mathcal{P}}, \mathbf{F}$), because the set \mathcal{P} is not a Cartesian product as instead $\hat{\mathcal{P}}$. For a set with no Cartesian product structure, a stronger condition on \mathbf{F} is required for the associated VI having a unique solution; which is indeed the strong monotonicity property [32, Th. 2.3.3]. However, milder conditions guaranteeing the uniqueness of the solution of the VI(\mathcal{P}, \mathbf{F}) than those given in Theorem 9(c1) can be obtained by imposing the strong monotonicity of \mathbf{F} on \mathcal{P} rather than on $\hat{\mathcal{P}} \supseteq \mathcal{P}$ (we omit the details because of the space limitation).

$$\begin{aligned}
0 \leq p_q(k) \perp & -\frac{1}{\hat{\sigma}_q^2(k) + \sum_r |\hat{H}_{qr}(k)|^2 p_r(k)} + \sum_{p=1}^P \left| H_{pq}^{(P,S)}(k) \right|^2 \left(\lambda_{p,k}^{\text{peak}} + \lambda_{p,\text{tot}} \right) + \zeta_{q,k} + \mu_q \geq 0, \\
& \forall k = 1, \dots, N, \quad \forall q = 1, \dots, Q, \\
0 \leq \zeta_{q,k} \perp & p_q^{\text{max}}(k) - p_q(k) \geq 0, \\
& \forall k = 1, \dots, N, \quad \forall q = 1, \dots, Q, \\
0 \leq \mu_q \perp & P_q - \sum_{k=1}^N p_q(k) \geq 0, \\
& \forall q = 1, \dots, Q, \\
0 \leq \lambda_{p,\text{tot}} \perp & P_{p,\text{tot}}^{\text{ave}} - \sum_{q=1}^Q \sum_{k=1}^N \left| H_{pq}^{(P,S)}(k) \right|^2 p_q(k) \geq 0, \\
& \forall p = 1, \dots, P, \\
0 \leq \lambda_{p,k}^{\text{peak}} \perp & P_{p,k}^{\text{peak}} - \sum_{q=1}^Q \left| H_{pq}^{(P,S)}(k) \right|^2 p_q(k) \geq 0, \\
& \forall k = 1, \dots, N, \quad \forall p = 1, \dots, P.
\end{aligned} \tag{35}$$

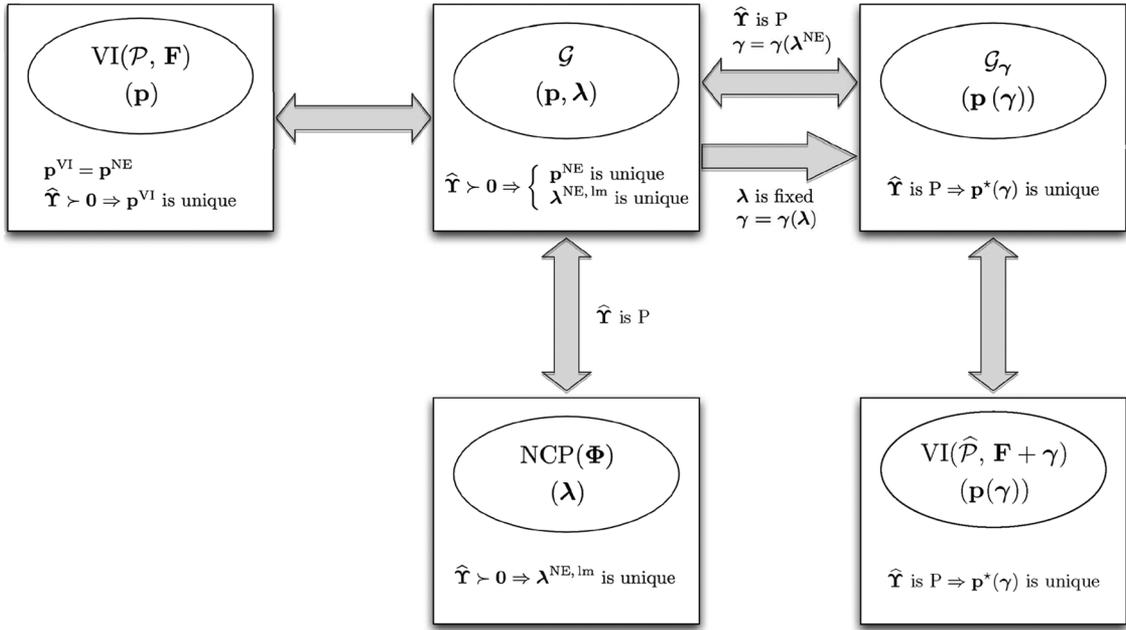


Fig. 2. Graphical representation of the main results stated in Theorem 9.

Remark 5 (On the NCP Reformulation of the Game): Theorem 9 postulates the equivalence of the game \mathcal{G} with the VI(\mathcal{P}, \mathbf{F}) [Theorem 9(a)] as well as with the NCP(Φ) [Theorem 9(b)–(c)]. The reformulation as a VI allows the characterization of the existence and uniqueness of the solution. However, it leads to algorithms that, in principle, require some coordination among the secondary users, since the interference constraints impose a coupling among strategies of the secondary users (the set \mathcal{P} of the VI(\mathcal{P}, \mathbf{F}) is not a Cartesian product). The NCP formulation precisely solves this limitation and offers the possibility of devising iterative algorithms that can be implemented in a distributed fashion among all players and whose convergence can be studied using known results from the theory of VIs (cf. [32, Chapter 12]). In the following, we thus focus on this NCP equivalence to devise distributed algorithms for computing the solutions of game \mathcal{G} .

B. Distributed Algorithms

In this section, we propose several iterative algorithms, along with their convergence properties, for computing an equilibrium solution of game \mathcal{G} . The proposed algorithms differ in: i) the signaling required between primary and secondary users; ii) the computational effort; iii) the convergence speed; and iv) the convergence analysis. A formal description of the proposed algorithms is given next.

Algorithm 2: Based on the NCP(Φ) formulation of the game \mathcal{G} [cf. Theorem 9(b)–(c)], the algorithm is just the *Projection Algorithm with variable steps* [32, Alg. 12.1.4] and is formally described in Algorithm 2, where the waterfilling mapping $\text{wf}_q(\cdot; \boldsymbol{\gamma}(\boldsymbol{\lambda}))$ is defined in (25), with $\boldsymbol{\gamma}(\boldsymbol{\lambda})$ given in (11), and $\Phi(\boldsymbol{\lambda})$ is defined in (29).

Algorithm 2: Projection algorithm with variable steps

Data: Choose any $\boldsymbol{\lambda}^{(0)} \geq \mathbf{0}$; set $n = 0$.

Step 1: If $\boldsymbol{\lambda}^{(n)}$ satisfies a suitable termination criterion: STOP

Step 2: Given $\boldsymbol{\lambda}^{(n)}$, compute $\mathbf{p}^*(\boldsymbol{\gamma}(\boldsymbol{\lambda}^{(n)}))$ as the unique NE of $\mathcal{G}_{\boldsymbol{\gamma}(\boldsymbol{\lambda}^{(n)})}$ (e.g., via Algorithm 1)

$$\mathbf{p}_q^*(\boldsymbol{\gamma}(\boldsymbol{\lambda}^{(n)})) = \text{wf}_q(\mathbf{p}_{-q}^*(\boldsymbol{\gamma}(\boldsymbol{\lambda}^{(n)})); \boldsymbol{\gamma}(\boldsymbol{\lambda}^{(n)})), \quad \forall q \quad (36)$$

Step 3: Choose $\tau_n > 0$ and update the price vectors $\boldsymbol{\lambda}$ according to

$$\boldsymbol{\lambda}^{(n+1)} = [\boldsymbol{\lambda}^{(n)} - \tau_n \Phi(\boldsymbol{\lambda}^{(n)})]^+ \quad (37)$$

Step 4: Set $n \leftarrow n + 1$; go to Step 1.

The algorithm has the following interpretation.³ In the main loop, at the n th iteration, each primary user p measures the received interference generated by the secondary users and, *locally and independently* from the other primary users, adjusts his own set of prices $\boldsymbol{\lambda}_p^{(n)}$ via the simple projection scheme in (37) [i.e., the update in (37) is implemented in a decentralized fashion]. The primary users broadcast their own prices $\boldsymbol{\lambda}_p^{(n)}$ to the secondary users, who then play the game $\mathcal{G}_{\boldsymbol{\gamma}(\boldsymbol{\lambda}^{(n)})}$ defined in (10) corresponding to the price tuple $\boldsymbol{\lambda}^{(n)}$, based on Algorithm 1 (sequential, simultaneous, or asynchronous). The convergence

³Through this section we tacitly assume that some interaction between the primary and the secondary users in the CR system is allowed (as, e.g., in the property-right CR model [5]), so that the updates of the prices can be performed by the primary users. A detailed discussion on the implementation of the proposed algorithms within the context of other CR debate models is given in Section IV-C.

result of Algorithm 2 is given in the following theorem, whose proof follows from Proposition 8 and [32, Th. 12.1.8].⁴

Theorem 10: Suppose $\hat{\mathbf{Y}} \succ \mathbf{0}$. If the scalars τ_n are chosen so that $0 < \inf_n \tau_n \leq \sup_n \tau_n < 2c_{\text{coc}}(\Phi)$, where $c_{\text{coc}}(\Phi)$ is defined in (32), then the sequence $\{\lambda^{(n)}\}_{n=0}^{\infty}$ generated by Algorithm 2 converges to a solution of the NCP(Φ).

Algorithm 3: In this algorithm, we modify the outer loop of Algorithm 2 [i.e., the rule for updating the price tuple in the Step 3] to avoid the choice of the step-size sequence $\{\tau_n\}$ that depends on the co-coercive modulus of $\Phi(\lambda)$ as instead in Algorithm 2 (cf. Theorem 10). The proposed modification is an adaptation of the *Hyperplane Projection Algorithm* presented in [32, Alg. 12.1.12], which adds an *Armijo-type* step-size selection that requires shared information among the primary users. The main loop of Algorithm 3 is formally described in the following, where $\delta \in (0, 1)$ is a given constant.

Algorithm 3: Hyperplane Projection Algorithm

The steps of the algorithm are the same as those of Algorithm 2 except for Step 3, which is modified as follows. Let $\lambda^{(n)} \geq \mathbf{0}$ and $\delta \in (0, 1)$ be given.

Step 3a : Compute

$$\lambda^{(n+\frac{1}{4})} \triangleq [\lambda^{(n)} - \Phi(\lambda^{(n)})]^+ \quad (38)$$

Step 3b : Compute $\lambda^{(n+(1/2))} \triangleq \lambda^{(n)} + 2^{-\ell_n}(\lambda^{(n+(1/4))} - \lambda^{(n)})$, with ℓ_n being the smallest nonnegative integer ℓ (which must be finite) such that, with $\tau \triangleq 2^{-\ell}$,

$$\begin{aligned} [\lambda^{(n)} - \lambda^{(n+\frac{1}{4})}]^T \Phi(\lambda^{(n)} + \tau(\lambda^{(n+\frac{1}{4})} - \lambda^{(n)})) \geq \\ \delta \|\lambda^{(n)} - \lambda^{(n+\frac{1}{4})}\|_2^2. \end{aligned} \quad (39)$$

Step 3c : Update the price tuple

$$\lambda^{(n+1)} \triangleq \left[\lambda^{(n)} - \frac{[\lambda^{(n)} - \lambda^{(n+\frac{1}{2})}]^T \Phi(\lambda^{(n+\frac{1}{2})})}{\|\Phi(\lambda^{(n+\frac{1}{2})})\|_2^2} \Phi(\lambda^{(n+\frac{1}{2})}) \right]^+ \quad (40)$$

The additional computational complexity in the above main loop with respect to the main loop of Algorithm 2 is in the repeated evaluation of the function $\Phi(\lambda^{(n)} + \tau(\lambda^{(n+(1/4))} - \lambda^{(n)}))$ corresponding to a (finite) sequence of decreasing step-sizes τ . When implemented by the primary users, this requires a signaling among them (if there is more than one primary user in the system). Moreover, for each evaluation from the primary users of Φ at a given λ , the secondary users are required to play the game \mathcal{G}_γ with $\gamma = \gamma(\lambda)$. The convergence of Algorithm 3

⁴For the reader convenience, we recall that, according to [32, Th. 12.1.8], the projection algorithm with variable steps [32, Alg. 12.1.4] converges to a solution of the NCP(Φ) [whose existence is guaranteed by Theorem 9(b), under the P-property of $\hat{\mathbf{Y}}$] if the function Φ is co-coercive on $\mathbb{R}_+^{P(N+1)}$ with constant $c_{\text{coc}}(\Phi)$, and $0 < \inf_n \tau_n \leq \sup_n \tau_n < 2c_{\text{coc}}(\Phi)$. Invoking Proposition 8, Theorem 10 follows readily.

is stated in the following theorem, whose proof follows from Proposition 8 and [32, Th. 12.1.16].⁵

Theorem 11: Suppose $\hat{\mathbf{Y}} \succ \mathbf{0}$. Then the sequence $\{\lambda^{(n)}\}_{n=0}^{\infty}$ generated by Algorithm 3 converges to a solution of the NCP(Φ).

Algorithm 4: With respect to Algorithm 2, this algorithm has the added benefits of obtaining the (unique) least-norm solution of the NCP(Φ), as stipulated in Theorem 9(c) (recall that the optimal λ might not be unique; see Remark 13). To this end, the price tuple is updated by successively solving a sequence of non-linear complementarity subproblems defined by the perturbed map $\Phi + \varepsilon_n \mathbf{I}$ (Tikhonov regularization), for a sequence of decreasing scalars $\{\varepsilon_n\} \downarrow 0$, where \mathbf{I} denotes the identity map (i.e., $\mathbf{I} : \lambda \rightarrow \lambda$). For each of these subproblems, we employ a simple *Constant-step Projection Method* [32, Alg. 12.1.1], whose convergence is ensured, under $\hat{\mathbf{Y}} \succ \mathbf{0}$, by the strong monotonicity of the map $\Phi + \varepsilon_n \mathbf{I}$ [32, Th. 12.1.2]. The formal description of the main loop of Algorithm 4 is given in the following and convergence stated in Theorem 12.

Algorithm 4: Tikhonov regularization Algorithm

The steps of the algorithm are the same as those of Algorithm 2 except for Step 3, which is modified as follows. Let $\lambda^{(n)} \geq \mathbf{0}$, $\{\varepsilon_\nu\}_{\nu=1}^{\infty} \downarrow 0$ and $\tau_n > 0$ be given.

Step 3 : Compute the unique solution $\lambda^{(n+1)}$ of the NCP($\Phi + \varepsilon_n \mathbf{I}$) as the limit point of the sequence $\{\hat{\lambda}^{(\ell,n)}\}_{\ell=0}^{\infty}$, generated by the subiteration: given $\hat{\lambda}^{(0,n)}$, let

$$\hat{\lambda}^{(\ell+1,n)} \triangleq \left[\hat{\lambda}^{(\ell,n)} - \tau_n^{-1} \left(\Phi(\lambda^{(n)}) + \varepsilon_n \hat{\lambda}^{(\ell,n)} \right) \right]^+, \quad \ell = 1, 2, \dots, \infty. \quad (41)$$

Theorem 12: Suppose $\hat{\mathbf{Y}} \succ \mathbf{0}$. If each step-size τ_n^{-1} in (41) is such that $\tau_n > ((c_{\text{Lip}}(\Phi) + \varepsilon_n)^2 / 2\varepsilon_n)$, where $c_{\text{Lip}}(\Phi)$ is defined in (33), then the sequence $\{\hat{\lambda}^{(\ell,n)}\}_{\ell=0}^{\infty}$ generated by the subiteration in Algorithm 4 converges to the unique solution $\lambda^{(n+1)} = \lambda^{(n+1)}(\varepsilon_n)$ of the NCP($\Phi + \varepsilon_n \mathbf{I}$). Moreover, the sequence $\{\lambda^{(n)}\}_{n=1}^{\infty}$ converges to the least-norm solution of the NCP(Φ) as $\varepsilon_n \downarrow 0$.

Proof: See Appendix E. ■

A variation (inexact version) of the proposed algorithm can also be considered, in which each $\lambda^{(n+1)}$ is not required to be an exact solution of the NCP($\Phi + \varepsilon_n \mathbf{I}$); i.e., the subiterations (41) can be terminated according to a prescribed criterion that progressively becomes tighter as the iteration in n proceeds. We omit the details and refer the interested reader to [45].

Algorithm 5: As opposed to the previous algorithms based on two nested loops, in this algorithm, there is only a major loop in which the secondary and the primary users update their decisions at the same level either sequentially or simultaneously

⁵For the reader convenience, we recall that, according to [32, Th. 12.1.16], the convergence of the Hyperplane Projection Algorithm [32, Alg. 12.1.12] (Algorithm 3) to a solution of the NCP(Φ) is guaranteed if the function Φ is continuous and monotone on $\mathbb{R}_+^{P(N+1)}$ (weaker conditions on Φ are required in [32, Th. 12.1.16]). Under the assumption $\hat{\mathbf{Y}} \succ \mathbf{0}$, the mapping Φ satisfies these conditions (cf. Proposition 8).

(in the version described in Algorithm 5, the update is simultaneous). Thus, in Algorithm 5, the primary users adjust their prices as soon as the secondary users complete *one iteration* of their non-equilibrium allocation updates, rather than waiting for a full equilibrium response, as in Algorithm 2. The formal description of the algorithm is given in Algorithm 5.

Interestingly, in Algorithm 5, the primary users can be interpreted as an extra player of an augmented game $\widehat{\mathcal{G}}$ (equivalent to the original game \mathcal{G}), who solves a trivial nonnegatively constraint linear program in the variable $\boldsymbol{\lambda}$, parametrized by \mathbf{p} and based on the *proximal regularization* (alternative regularization schemes are possible, such as the well-known Tikhonov regularization [32]): At the beginning of the $(n+1)$ th iteration, given $\zeta > 0$, $\mathbf{p}^{(n+1)}$ and $\boldsymbol{\lambda}^{(n)}$, $\boldsymbol{\lambda}^{(n+1)}$ is the unique minimizer of the convex quadratic program:

$$\begin{aligned} \min_{\boldsymbol{\lambda} \geq \mathbf{0}} & \sum_{p=1}^P \lambda_{p,\text{tot}} \left(P_{p,\text{tot}}^{\text{ave}} - \sum_{q=1}^Q \sum_{k=1}^N \left| H_{pq}^{(P,S)}(k) \right|^2 p_q^{(n+1)}(k) \right) \\ & + \sum_{p=1}^P \sum_{k=1}^N \lambda_{p,k}^{\text{peak}} \left(P_{p,k}^{\text{peak}} - \sum_{q=1}^Q \left| H_{pq}^{(P,S)}(k) \right|^2 p_q^{(n+1)}(k) \right) \\ & + \frac{\zeta}{2} \left\| \boldsymbol{\lambda} - \boldsymbol{\lambda}^{(n)} \right\|^2 \end{aligned} \quad (42)$$

whose solution has the explicit expression given in (44).

Algorithm 5: Proximal regularization algorithm

Data: Choose any $\boldsymbol{\lambda}^{(0)} \geq \mathbf{0}$, $\mathbf{p}_q^{(0)} \in \widehat{\mathcal{P}}_q$ for all $q = 1, \dots, Q$, and $\zeta > 0$; set $n = 0$.

Step 1: If $(\boldsymbol{\lambda}^{(n)}, \mathbf{p}^{(n)})$ satisfies a suitable termination criterion: STOP

Step 2: Given $\boldsymbol{\lambda}^{(n)}$, sequentially for $q = 1, \dots, Q$, compute $\mathbf{p}_q^{(n+1)}(\boldsymbol{\gamma}(\boldsymbol{\lambda}^{(n)}))$ as

$$\begin{aligned} \mathbf{p}_q^{(n+1)}(\boldsymbol{\gamma}(\boldsymbol{\lambda}^{(n)})) = \text{wf}_q & \left(\mathbf{p}_1^{(n+1)}, \dots, \mathbf{p}_{q-1}^{(n+1)}, \mathbf{p}_{q+1}^{(n)}, \dots, \mathbf{p}_Q^{(n)}; \boldsymbol{\gamma}(\boldsymbol{\lambda}^{(n)}) \right) \end{aligned} \quad (43)$$

Step 3: Update the price vectors $\boldsymbol{\lambda}$: for all $p = 1, \dots, P$, and $k = 1, \dots, N$, compute

$$\begin{aligned} \lambda_{p,\text{tot}}^{(n+1)} &= \left[\lambda_{p,\text{tot}}^{(n)} + \right. \\ & \left. - \zeta^{-1} \left(P_{p,\text{tot}}^{\text{ave}} - \sum_{q=1}^Q \sum_{k=1}^N \left| H_{pq}^{(P,S)}(k) \right|^2 p_q^{(n+1)}(k; \boldsymbol{\gamma}(\boldsymbol{\lambda}^{(n)})) \right) \right]^+ \\ \lambda_{p,k}^{\text{peak}^{(n+1)}} &= \left[\lambda_{p,k}^{\text{peak}^{(n)}} + \right. \\ & \left. - \zeta^{-1} \left(P_{p,k}^{\text{peak}} - \sum_{q=1}^Q \left| H_{pq}^{(P,S)}(k) \right|^2 p_q^{(n+1)}(k; \boldsymbol{\gamma}(\boldsymbol{\lambda}^{(n)})) \right) \right]^+ \end{aligned} \quad (44)$$

Step 4: Set $n \leftarrow n + 1$; go to Step 1.

While being closest in spirit to the IWFA algorithms described in Section III-C, the convergence analysis of Algorithm 5 is in jeopardy due to two causes: (i) the coupling of the secondary users' variables due to the interference constraints,

and (ii) the relaxed control of pricing mechanism that imposes too little restriction on the price tuple. Up to date, a proof of convergence of Algorithm 5 is missing. In Section V, we provide some numerical results supporting the convergence of the algorithm in practical CR scenarios.

C. Practical Implementation of the Proposed Algorithms

In the proposed algorithms there are two levels of updates: 1) the computation of the optimal power allocations of the secondary users, given the prices $\gamma_q(k; \boldsymbol{\lambda})$; and 2) the updates of the prices $\lambda_{p,k}^{\text{peak}}$ and $\lambda_{p,\text{tot}}$, given the interference (the power allocation) generated by the secondary users. The former can be performed directly by the secondary users using some of the algorithms (either synchronous or asynchronous) proposed in Section III-C. Note that, once $\gamma_q(k; \boldsymbol{\lambda})$ are given, these algorithms are totally distributed, since the secondary users only need to measure the received MUI over the N subcarriers to perform the waterfilling solution in (25). The update of the prices $\lambda_{p,k}^{\text{peak}}$ and $\lambda_{p,\text{tot}}$, once the interference generated by the secondary users at the receivers of the primary users is given, can be performed in different forms, according to Algorithms 2–5. Depending on the debate position assumed for the CR network—the *property-right model* or the *common model*—this update can be performed by the primary users (as assumed so far) or by the secondary users themselves, which leads to a different computational complexity and amount of signaling between the primary and the secondary users, as described next.

Property-Right CR Model: In a CR network based on the property-right model (also termed as spectrum leasing model), an interaction between the primary and the secondary users is allowed. It is thus natural that the update of the prices is performed by the primary users. In such a case, in all the proposed algorithms, the signaling from the secondary users to the primary users is implicit, since the primary users to update the prices only need to locally measure the global received interference. The signaling from the primary to the secondary users, however, is explicit: the primary users have to broadcast the prices $\lambda_{p,k}^{\text{peak}}$ and $\lambda_{p,\text{tot}}$ and the secondary users receive and estimate their values. Note that, if the transmission of the prices is performed by the primary receivers and the reception by the secondary transmitters, then there is no need from the secondary users to estimate separately the cross-channel transfer functions $H_{pq}^{(P,S)}(k)$ and the prices, since the secondary users receive directly what they really need to compute the waterfilling solution (25), which are the terms $\left| H_{pq}^{(P,S)}(k) \right|^2 (\lambda_{p,k}^{\text{peak}} + \lambda_{p,\text{tot}})$. Regarding the computational complexity of the price update process, Algorithms 2–4 require inner iterations among the secondary users responding to the primary users' announced price directives, whereas Algorithm 5 is composed by only one major loop in which the secondary and primary users update their decisions at the same level, either sequentially or simultaneously. More specifically, to update the prices, Algorithm 2 employs a gradient projection methods with arbitrarily chosen variable steps that are restricted to be less than the co-coercivity modulus of the function Φ (see Theorem 10). By employing a proper variable step-size rule, Algorithm 3 bypasses this restriction, and thus, can be implemented without the knowledge of this

constant. Nevertheless, additional computations are required in the main loop of the algorithm that involve shared information among the primary users. By employing a double iteration within each main loop among the primary users, Algorithm 4 offers a method to compute the least-norm price tuple of the game \mathcal{G} as stipulated by part Theorem 9(c), at the price of some signaling among the primary users. Of course, in Algorithms 3 and 4 there is no extra signaling in the outer loop if there is only one primary users in the system. Finally, being composed by only one major loop, Algorithm 5 has the advantage to be totally distributed and requires minimum interaction among the players.

In some scenarios where the primary users cannot communicate with the secondary users (e.g., when the primary users are legacy systems) and the primary receivers have a fixed geographical location, it may be possible to install some monitoring devices close to each primary receiver having the functionality of interference measurement as well as price computation and broadcasting.

CR Common Model: In a CR network based on the common model, the primary users are oblivious of the presence of the secondary users, thus behaving as if no secondary activity was present. In such a case, the update of $\lambda_{p,k}^{\text{peak}}$ and $\lambda_{p,\text{tot}}$ needs to be performed by the secondary users themselves. Nevertheless, at the price of additional (albeit reduced) signaling among the secondary users and computational complexity, Algorithms 2–5 can be still implemented in a distributed fashion by the secondary users. The only additional assumption we need is that each secondary user q can estimate the cross-channel transfer functions $|H_{pq}^{(P,S)}(k)|^2$ between his transmitter and the receivers of the primary users and thus the interference $\text{MUI}_{pq}(k) \triangleq |H_{pq}^{(P,S)}(k)|^2 p_q(k)$ generated at the receivers of the primary users over each subcarrier. The global interference $\sum_{q=1}^Q \text{MUI}_{pq}(k)$ generated by all the secondary users, which is what each secondary user really needs to update the prices, can be locally computed by each secondary users by running an average consensus algorithm [46], [47] that requires the interaction only between nearby secondary nodes. It is known that if the network of secondary users is connected, under mild assumptions on the consensus algorithm, every secondary users is able to get the average interference $\sum_{q=1}^Q \text{MUI}_{pq}(k)$ in a distributed and (possibly) asynchronous way [47].

Finally, observe that the (almost) distributed nature of the proposed algorithms comes at some price: since the Nash equilibria of the game \mathcal{G} are not in general Pareto efficient (even when there is a unique NE), the solutions achievable by the proposed algorithms might be Pareto dominated. A formal analysis of the performance loss due to the use of the Nash criterion with respect to the Pareto optimal solutions (the so-called price-of-anarchy) is up-to-date a formidable open problem to solve, and it goes beyond the scope of this paper. Recall however that solving the system-wide optimization is an NP-hard problem, even in the absence of interference constraints [8]. Furthermore, there are some practical CR scenarios where a (possibly suboptimal) system-wide optimization cannot be implemented, as secondary users are heterogeneous systems that are not willing to cooperate. Devising possibly distributed algorithms that converge to the globally optimal solutions of the system-wide optimization

problem is up to date a formidable open problem that is worth to be investigated. This, however, goes beyond the scope of the paper.

V. NUMERICAL RESULTS

In this section, we provide some numerical results to illustrate our theoretical findings. More specifically, we compare the performance of the proposed game theoretical formulation via VI with the classical IWFA based on *individual* mask constraints, as studied in [20], [21], [23], and [24]. We also compare some of the proposed algorithms in terms of convergence speed.

Conservative IWFA versus Flexible IWFA: We compare three different approaches, namely the VI-based formulation (Algorithms 2 and 5), the classical IWFA [17], [33], and the IWFA using spectral mask constraints [20], [21], [23], [24], in terms of interference generated at the primary user receivers and the achievable sum-rate from the secondary users; we refer to these algorithms as *flexible IWFA*, *classical IWFA* and *conservative IWFA*, respectively. As an example, we consider a CR system composed of 6 secondary links randomly distributed within an hexagonal cell and one primary user (the BS at the center of the cell). The primary user imposes a constraint on the maximum interference that can tolerate. For simplicity in our description, we assume that the primary user imposes a constant interference threshold over the whole spectrum, namely: $P_{p,k}^{\text{peak}} = 0.01$ (strong interference constraint) and $P_{p,k}^{\text{peak}} = 0.028$ (weak interference constraint) for all $k = 1, \dots, N$ [see (4)]. The individual spectral mask constraints used in the conservative IWFA are chosen so that all the secondary users generate the same interference level at the primary receiver and the aggregate interference satisfies the imposed interference threshold. In Fig. 3 we plot the PSD of the interference generated by the secondary users at the receiver of the primary user, obtained, for a given channel realization, using the flexible IWFA (based on Algorithm 2 with constant step-size) and the conservative IWFA, under two different interference constraints. As benchmark, we also include the PSD of the interference generated by the classical IWFA [17], [33]. We clearly see from the picture that while classical IWFA violates the interference constraints, both conservative and flexible IWFAs satisfy them, but the global interference constraints impose less stringent conditions on the transmit power of the secondary users than those imposed by the individual interference constraints based on the spectral masks. However, this comes at the price of some signaling from the primary to the secondary users. Note also that, in the case of weak interference constraints, the flexible IWFA and the classical IWFA generate almost the same interference profile over the subcarriers where the interference constraint is strictly satisfied [since the complementarity condition (8), the prices $\lambda_{p,k}^{\text{peak}}$ over those subcarriers are zero].

Thanks to less stringent constraints on the transmit powers of the secondary users, the flexible IWFA is expected to exhibit a much better performance than the conservative IWFA also in terms of rates achievable by the secondary user. Fig. 4 confirms this intuition, where we plot the average sum-rate of the secondary users achievable by the conservative IWFA and the flexible IWFA as a function of the maximum tolerable interference

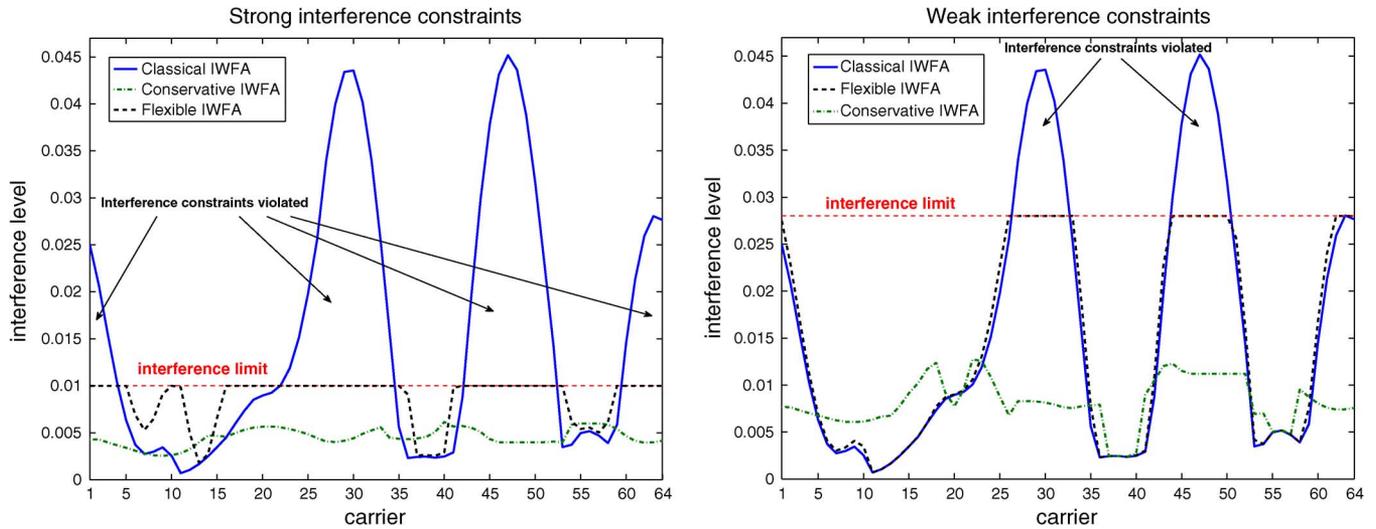


Fig. 3. Comparison of IWFA algorithms: classical IWFA, conservative IWFA, and flexible IWFA, under strong and weak interference constraints; PSD of the interference profile at the primary user's receiver.

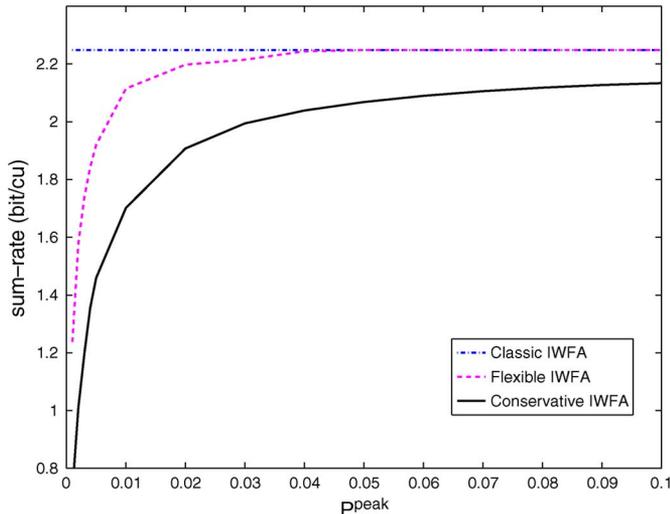


Fig. 4. Comparison of IWFA algorithms: classical IWFA, conservative IWFA, and flexible IWFA: Achievable sum-rate versus the interference constraint.

at the primary receiver, within the same setup of Fig. 3. The curves are averaged over 500 random i.i.d. Gaussian channel realizations.

Convergence Speed: In Fig. 5(a) we plot the worst-case violation of the interference constraint achieved by Algorithms 2 and 5 versus the number of iterations of the outer loop, for a CR system as in Fig. 3, composed now of 15 active secondary links. Interestingly, for the example considered in the figure, both Algorithms 2 and 5 experience the same convergence behavior (provided that the step-size is properly chosen) and converge reasonably fast. Thus, in such a scenario, Algorithm 5 is preferred to Algorithm 2, since it requires less iterations among the secondary users. Finally, in Fig. 5(b), we compare the performance in terms of convergence speed of the sequential and simultaneous IWFA with pricing (see Algorithm 1) used to compute the NE of game $\mathcal{G}_{\gamma(\lambda)}$ in the inner loop of Algorithm 2, for a given price tuple λ and channel realization. In the figure,

we plot the rate evolution of the secondary users' links corresponding to the two cited algorithms as a function of the iteration index. To make the figure not excessively overcrowded, we plot only the curves of 3 out of 15 links. As expected, the sequential IWFA is slower than the simultaneous IWFA, especially if the number of active links Q is large, since each user is forced to wait for all the users scheduled in advance, before updating his own power allocation. The same qualitative behavior has been observed for different channel realizations and value of prices. The fast convergence behavior of the IWFA in the inner loop provides an intuitive explanation of why Algorithms 2 and 5 have been experienced to have almost the same convergence speed [Fig. 5(a)], provided that the step-size is properly chosen: After the first round of the IWFA in the inner loop, the secondary users are expected to be quite close the NE of $\mathcal{G}_{\gamma(\lambda)}$ already.

VI. CONCLUSION

In this paper, we have proposed a novel NE problem based on VI to solve one of the challenging and unsolved resource allocation problems in CR systems: How to allow in a decentralized way concurrent communication over frequency-selective channels among secondary users, under constraints imposed to the secondary users on the maximum per-carrier and global MUI tolerable at the primary receivers. We have seen how VI theory provides the natural framework to solve the proposed NE problem, namely: 1) the establishment of conditions guaranteeing the uniqueness of the equilibrium solution; and 2) the design of fairly decentralized algorithms able to reach the equilibrium points, with minimal coordination among the nodes. The proposed algorithms differ in the tradeoff between performance (in terms of information rate) achievable by the secondary users and the degree of information to be exchanged between the primary and the secondary users, and they have been shown to outperform the classical algorithms based on IWFA with spectral mask constraints (at the price of some signaling, albeit very reduced, from the primary to the secondary users).

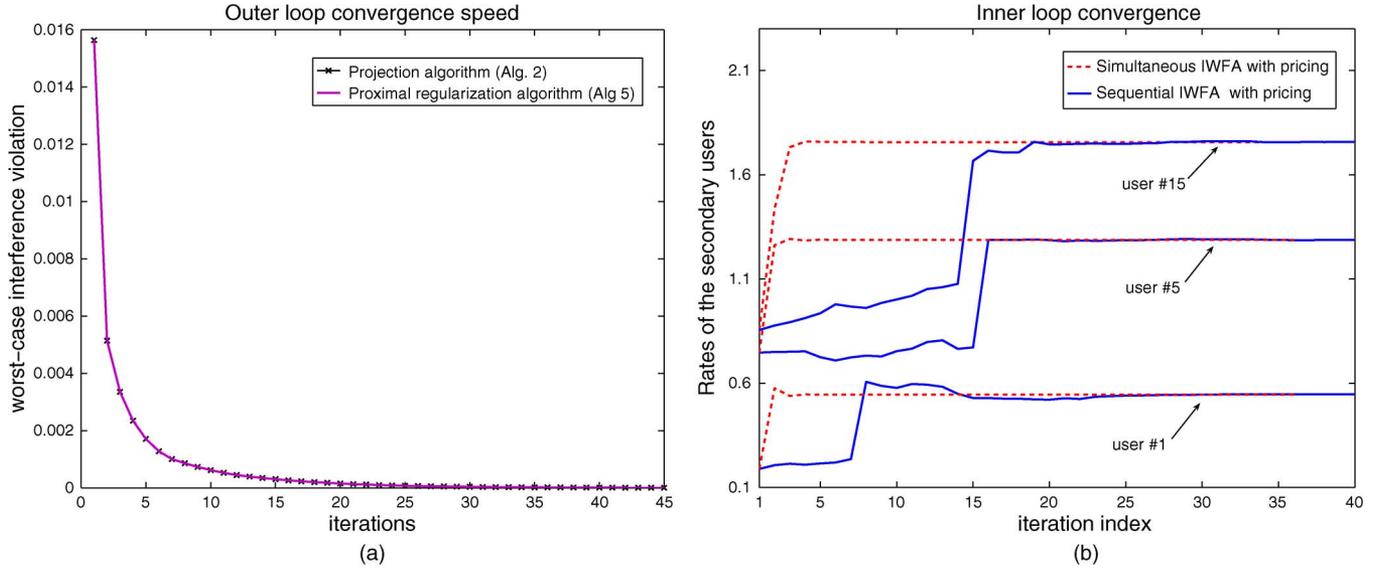


Fig. 5. Example of convergence speed. (a) Worst-case violation of the interference constraints achieved by Algorithms 2 and 5 (flexible IWFA). (b) Secondary users’ rates versus the iteration index achieved in the inner loop of Algorithm 2 by Algorithm 1 (the sequential IWFA with pricing and the simultaneous IWFA with pricing).

APPENDIX A

BACKGROUND ON VARIATIONAL INEQUALITY PROBLEMS

Definition of the VI Problem: Given a subset \mathcal{K} of the Euclidean n -dimensional space \mathbb{R}^n and a mapping $\mathbf{F} : \mathcal{K} \mapsto \mathbb{R}^n$, the VI problem, denoted $\text{VI}(\mathcal{K}, \mathbf{F})$, is to find a vector $\mathbf{x}^* \in \mathcal{K}$ (called a solution of the VI) such that [32, Def. 1.1.1]

$$(\mathbf{x} - \mathbf{x}^*)^T \mathbf{F}(\mathbf{x}^*) \geq 0, \quad \forall \mathbf{x} \in \mathcal{K}. \quad (45)$$

The set of solutions to this problem is denoted $\text{SOL}(\mathcal{K}, \mathbf{F})$.

Several standard problems in nonlinear programming, game theory, and nonlinear analysis can be naturally formulated as a VI problem, and some examples follow (see [32] for more source problems).

– *Solution of systems of equations.* The simplest example of VI is the problem of solving a system of equations. In fact, it is easy to see that if $\mathcal{K} = \mathbb{R}^n$, then $\text{VI}(\mathbb{R}^n, \mathbf{F})$ is equivalent to finding a $\mathbf{x}^* \in \mathbb{R}^n$ such that $\mathbf{F}(\mathbf{x}^*) = \mathbf{0}$. As special case, if the mapping \mathbf{F} is affine, i.e., $\mathbf{F}(\mathbf{x}) = \mathbf{A}\mathbf{x} - \mathbf{b}$, the previous problem is equivalent to the classical system of equation $\mathbf{A}\mathbf{x}^* = \mathbf{b}$.

– *Fixed-point problems.* Given a closed and convex set \mathcal{K} and a mapping $\mathbf{T} : \mathcal{K} \mapsto \mathcal{K}$, the fixed-point problem is to find a vector $\mathbf{x}^* \in \mathcal{K}$ such that $\mathbf{T}(\mathbf{x}^*) = \mathbf{x}^*$. This problem can be converted into a VI format, simply by defining $\mathbf{F}(\mathbf{x}) = \mathbf{x} - \mathbf{T}(\mathbf{x})$.

– *Constrained and unconstrained optimization.* If \mathcal{K} is convex and the mapping \mathbf{F} in $\text{VI}(\mathcal{K}, \mathbf{F})$ is the gradient of a real-valued function $f : \mathcal{K} \mapsto \mathbb{R}$, then $\text{VI}(\mathcal{K}, \mathbf{F})$ represents a necessary conditions of optimality for the following optimization problem: find a point $\mathbf{x}^* \in \mathcal{K}$ such that $f(\mathbf{x}^*) \leq f(\mathbf{x})$, for all $\mathbf{x} \in \mathcal{K}$. Also, if the function f is convex, the reverse assertion is true, meaning that a point $\mathbf{x}^* \in \mathcal{K}$ minimizes f over \mathcal{K} if and only if is a solution to $\text{VI}(\mathcal{K}, \nabla f)$, where ∇f denotes the gradient of f (the VI coincides with the first-order necessary and sufficient

optimality conditions of a convex differentiable function). In particular, if we let $\mathcal{K} = \mathbb{R}^n$, we see that unconstrained convex optimization is also a VI problem.

– *Game theory problems.* Consider a strategic noncooperative game $\mathcal{G} = \langle \Omega, (\mathcal{Q}_q)_{q \in \Omega}, (u_q)_{q \in \Omega} \rangle$, where player q ’s problem is to determine, for each fixed but arbitrary tuple \mathbf{x}_{-q} of the other players’ strategies, an optimal strategy profile \mathbf{x}_q^* that solve the following optimization problem (in the minimization form) in the variable \mathbf{x}_q :

$$\begin{aligned} & \underset{\mathbf{x}_q}{\text{minimize}} && u_q(\mathbf{x}_q, \mathbf{x}_{-q}) \\ & \text{subject to} && \mathbf{x}_q \in \mathcal{Q}_q. \end{aligned} \quad (46)$$

Suppose that each $\mathcal{Q}_q \subset \mathbb{R}^{n_q}$ is convex and closed, and $u_q(\mathbf{x}_q, \mathbf{x}_{-q})$ is convex and continuously differentiable in \mathbf{x}_q . By convexity and the first-order optimality conditions, we infer that a strategy profile \mathbf{x}^* is a NE if and only if $(\mathbf{x}_q - \mathbf{x}_q^*)^T \nabla_{\mathbf{x}_q} u_q(\mathbf{x}^*) \geq 0$, for each $q = 1, \dots, Q$, where $\nabla_{\mathbf{x}_q} u_q(\mathbf{x})$ denotes the gradient of $u_q(\mathbf{x})$ with respect to \mathbf{x}_q . Summing these conditions and taking into account the Cartesian product structure of the strategy set of the game, it is not difficult to see that this set of inequalities is equivalent to the $\text{VI}(\mathcal{K}, \mathbf{F})$, with $\mathcal{K} = \mathcal{Q}_1 \times \dots \times \mathcal{Q}_Q$ and $\mathbf{F}(\mathbf{x}) = (\nabla_{\mathbf{x}_q} u_q(\mathbf{x}))_{q=1}^Q$ [32, Prop. 1.4.2].

Note that the game \mathcal{G} in (7), (8) is not a classical NEP as defined in (46) and thus cannot be readily cast in the VI problem as did above. Nevertheless, we can still gain from the VI framework to analyze the proposed game, as detailed in Section IV.

– *Complementarity problems.* When the set \mathcal{K} is a cone (i.e., $\mathbf{x} \in \mathcal{K} \Rightarrow \tau \mathbf{x} \in \mathcal{K}$ for all scalars $\tau \geq 0$), the VI admits an equivalent form known as a *complementarity problem*, denoted by $\text{CP}(\mathcal{K}, \mathbf{F})$, which is to find a vector \mathbf{x} such that [32, Def. 1.1.2]

$$\mathcal{K} \ni \mathbf{x} \perp \mathbf{F}(\mathbf{x}) \in \mathcal{K}^* \quad (47)$$

where \mathcal{K}^* is the dual cone of \mathcal{K} , defined as $\mathcal{K}^* \triangleq \{\mathbf{d} \in \mathbb{R}^n | \mathbf{v}^T \mathbf{d} \geq 0, \forall \mathbf{v} \in \mathcal{K}\}$. In this paper we are interested in a special class of CPs, when \mathcal{K} is the nonnegative orthant of \mathbb{R}^n . In such a case the CP($\mathbb{R}_+^n, \mathbf{F}$) is known as *nonlinear complementarity problem* (NCP) and denoted NCP(\mathbf{F}). Recognizing that the dual cone of the nonnegative orthant is the nonnegative orthant itself, the NCP(\mathbf{F}) is to find a vector \mathbf{x} such that [32, Def. 1.1.5]

$$\mathbf{0} \leq \mathbf{x} \perp \mathbf{F}(\mathbf{x}) \geq \mathbf{0}. \quad (48)$$

To cast the games proposed in this paper in the VI framework we will build on the Karush-Kuhn-Tucker (KKT) conditions to the VI, which are defined next.

KKT System to the VI: Let \mathcal{K} be represented by finitely many differentiable inequalities and equations, i.e., $\mathcal{K} \triangleq \{\mathbf{x} \in \mathbb{R}^n | \mathbf{h}(\mathbf{x}) = \mathbf{0}, \mathbf{g}(\mathbf{x}) \leq \mathbf{0}\}$, with $\mathbf{h} : \mathbb{R}^n \rightarrow \mathbb{R}^l$ and $\mathbf{g} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ being vector-valued continuously differentiable functions. The following two statements are valid [32, Prop. 1.3.4].

(a) Let $\mathbf{x} \in \text{SOL}(\mathcal{K}, \mathbf{F})$. Under mild conditions on the constraints,⁶ there exist vectors $\boldsymbol{\mu} \in \mathbb{R}^l$ and $\boldsymbol{\lambda} \in \mathbb{R}^m$ such that

$$\begin{aligned} \mathbf{0} &= \mathbf{F}(\mathbf{x}) + \sum_{j=1}^l \mu_j \nabla h_j(\mathbf{x}) + \sum_{i=1}^m \lambda_i \nabla g_i(\mathbf{x}) \\ \mathbf{0} &= \mathbf{h}(\mathbf{x}) \\ \mathbf{0} &\leq \boldsymbol{\lambda} \perp \mathbf{g}(\mathbf{x}) \leq \mathbf{0}. \end{aligned} \quad (49)$$

(b) Conversely, if each function h_j is affine and each function g_i is convex, and if $(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\lambda})$ satisfies (49), then $\mathbf{x} \in \text{SOL}(\mathcal{K}, \mathbf{F})$.

Existence and Uniqueness Results: The theory and solution methods for various kinds of VIs are developed rather well and allow one to choose a suitable way to investigate each particular problem under consideration. Here, we recall some of the basic conditions for the existence and uniqueness of the solution to a VI, as they are used in the paper. A classical existence result reads as follows [32]. The VI(\mathcal{K}, \mathbf{F}) is solvable if: i) \mathcal{K} is a nonempty, convex, and compact subset of a finite-dimensional Euclidean space; and ii) \mathbf{F} is a continuous mapping [32, Corollary 2.2.5]. As far as the uniqueness of the solution, we mention the following condition used in our derivations in the paper: the solution to VI(\mathcal{K}, \mathbf{F}) is unique if \mathbf{F} is continuous and strongly monotone (cf. Definition 1(a) in Section III) on the convex and closed set \mathcal{K} [32, Th. 2.3.3(b)] (the strong monotonicity of \mathbf{F} is sufficient also for the existence of a solution). Uniqueness conditions can be weakened if the set $\mathcal{K} \subseteq \mathbb{R}^n$ has a Cartesian structure, i.e., $\mathcal{K} = \prod_{q=1}^Q \mathcal{K}_q$, $\mathcal{K}_q \subseteq \mathbb{R}^{n_q}$ and $n = \sum_{q=1}^Q n_q$: if each \mathcal{K}_q is closed and convex and \mathbf{F} is a continuous uniformly-P function [cf. Definition 1(b)], then the VI(\mathcal{K}, \mathbf{F}) has a unique solution [32, Prop. 3.5.10].

Several solution methods along with their convergence properties have been proposed for VI in the literature. A treatment on the subject goes beyond the scope of this paper and we refer the interested reader to the technical literature on the subject. A good entry point on parallel and distributed algorithms and their convergence for optimization problems and variational inequalities

⁶One of the most general of constraints qualifications (CQs) is Abadie's CQ (see [32, Sec. 3.2] for more details). However, for the purposes of this paper, where the set \mathcal{K} is defined by only convex inequalities $g_i(\mathbf{x}) \leq 0$, with each g_i being a convex function, the Slater CQ is sufficient.

is the book [44]. A comprehensive and more advanced treatment can be found in the monograph [32]. In Section IV-B, we specialize some of these algorithms to solve the proposed equilibrium problem.

APPENDIX B

PROOF OF PROPOSITION 2

(a) Given $\mathbf{p} = (\mathbf{p}_q)_{q=1}^Q$ and $\mathbf{p}' = (\mathbf{p}'_q)_{q=1}^Q$ in $\hat{\mathcal{P}}$, for the sake of convenience, define for each $k = 1, \dots, N$, and $q = 1, \dots, Q$

$$\hat{\varphi}_q(k) \triangleq \sqrt{\left(\hat{\sigma}_q^2(k) + \sum_{r=1}^Q |\hat{H}_{qr}(k)|^2 p_r^{(k)} \right) \cdot \left(\hat{\sigma}_q^2(k) + \sum_{r=1}^Q |\hat{H}_{qr}(k)|^2 p'_r(k) \right)} \quad (50)$$

and

$$e_q(k) \triangleq \frac{p_q(k) - p'_q(k)}{\hat{\varphi}_q(k)}, \quad k = 1, \dots, N, \quad q = 1, \dots, Q. \quad (51)$$

Note that $\hat{\sigma}_q^2(k) \leq \hat{\varphi}_q(k) \leq \hat{\varphi}_q^{\max}(k) \triangleq \hat{\sigma}_q^2(k) + \sum_{r=1}^Q |\hat{H}_{qr}(k)|^2 p_r^{\max}(k)$ for $\mathbf{p} \in \hat{\mathcal{P}}$. Also recall that $|\hat{H}_{qq}(k)|^2 = 1$, (from the normalization in (16)). Then, for each $q = 1, \dots, Q$

$$(\mathbf{p}_q - \mathbf{p}'_q)^T (\mathbf{F}_q(\mathbf{p}) - \mathbf{F}_q(\mathbf{p}')) \quad (52)$$

$$= \sum_{k=1}^N \frac{(p_q(k) - p'_q(k)) \sum_{r=1}^Q |\hat{H}_{qr}(k)|^2 (p_r(k) - p'_r(k))}{\hat{\varphi}_q(k)} \quad (53)$$

$$\geq \sum_{k=1}^N e_q(k)^2 - \sum_{r \neq q} \left| \sum_{k=1}^N e_q(k) \frac{|\hat{H}_{qr}(k)|^2 \hat{\varphi}_r(k)}{\hat{\varphi}_q(k)} e_r(k) \right| \quad (54)$$

$$\geq \sum_{k=1}^N e_q(k)^2 - \sum_{r \neq q} \left(\sum_{k=1}^N e_q(k)^2 \right)^{1/2} \quad (55)$$

$$\cdot \max_{1 \leq k \leq N} \left\{ \frac{|\hat{H}_{qr}(k)|^2 \hat{\varphi}_r(k)}{\hat{\varphi}_q(k)} \right\} \times \left(\sum_{k=1}^N e_r(k)^2 \right)^{1/2} \quad (56)$$

$$\geq \hat{e}_q \sum_{r=1}^Q \hat{\boldsymbol{\Gamma}}_{qr} \hat{e}_r, \quad \text{with } \hat{e}_r \triangleq \left(\sum_{k=1}^N e_r(k)^2 \right)^{1/2}, \quad (57)$$

where in (56) we used the Cauchy-Schwarz inequality.

Since $\hat{\boldsymbol{\Gamma}}$ is a P-matrix from the assumption, it follows from the characterization [34, Th. 3.3.4(b)] of a P-matrix⁷ that the constant $c(\hat{\boldsymbol{\Gamma}}) \triangleq \min_{\|\mathbf{x}\|_2=1} \{\max_{1 \leq q \leq Q} \{x_q(\hat{\boldsymbol{\Gamma}}\mathbf{x})_q\}\}$ is positive. In other words

$$\max_{1 \leq q \leq Q} x_q(\hat{\boldsymbol{\Gamma}}\mathbf{x})_q \geq c(\hat{\boldsymbol{\Gamma}}) \|\mathbf{x}\|_2^2 \quad (58)$$

⁷An alternative characterization of a P-matrix with respect to that given in Definition 2 is the following: A matrix $\mathbf{M} \in \mathbb{R}^{n \times n}$ is said to be a P-matrix if \mathbf{M} reverses the sign of no nonzero vector, i.e., $[x_q(\mathbf{M}\mathbf{x})_q \leq 0 \text{ for all } q] \Rightarrow [\mathbf{x} = \mathbf{0}]$ [34, Th. 3.3.4(b)].

holds for an arbitrary vector $\mathbf{x} \in \mathbb{R}^Q$. Now, combining (57) and (58), we get

$$\begin{aligned} & \max_{1 \leq q \leq Q} (\mathbf{p}_q - \mathbf{p}'_q)^T (\mathbf{F}_q(\mathbf{p}) - \mathbf{F}_q(\mathbf{p}')) \\ & \geq c(\hat{\mathbf{Y}}) \sum_{r=1}^Q \left(\sum_{k=1}^N e_r(k)^2 \right) \\ & \geq \frac{c(\hat{\mathbf{Y}})}{\max_{1 \leq r \leq Q} \max_{1 \leq k \leq N} (\hat{\varphi}_r^{\max}(k))^2} \left(\sum_{r=1}^Q \sum_{k=1}^N (p_r(k) - p'_r(k))^2 \right), \end{aligned} \quad (59)$$

establishing the uniformly P-property of the function \mathbf{F} on $\hat{\mathcal{P}}$ (and \mathcal{P}).

(b) The proof of the strong monotonicity property of \mathbf{F} under the assumption that $\hat{\mathbf{Y}}$ is positive definite follows similarly: given (57) and summing over q we have

$$\begin{aligned} & (\mathbf{p} - \mathbf{p}')^T (\mathbf{F}(\mathbf{p}) - \mathbf{F}(\mathbf{p}')) \\ & \geq \hat{\mathbf{e}}^T \hat{\mathbf{Y}} \hat{\mathbf{e}} \geq \frac{\lambda_{\min}(\hat{\mathbf{Y}})}{\max_{1 \leq r \leq Q} \max_{1 \leq k \leq N} (\hat{\varphi}_r^{\max}(k))^2} \|\mathbf{p} - \mathbf{p}'\|_2^2. \end{aligned} \quad (60)$$

APPENDIX C

PROOF OF THEOREM 6

We prove next, using contraction arguments, that the sequence $\{\mathbf{p}^{(n)}(\boldsymbol{\gamma})\}_{n=0}^{\infty}$ generated by the sequential IWFA described in Algorithm 1 converges to the unique solution of the VI($\hat{\mathcal{P}}, \mathbf{F} + \boldsymbol{\gamma}$), provided that the matrix $\hat{\mathbf{Y}}$ is a P-matrix. One can similarly show that the P-property of matrix $\hat{\mathbf{Y}}$ is also sufficient for the convergence of the simultaneous implementation of the algorithm (we omit the details because of the space limitation).

Suppose that $\hat{\mathbf{Y}}$ is a P-matrix and let denote by $\mathbf{p}^*(\boldsymbol{\gamma})$ the unique NE of game $\mathcal{G}_\boldsymbol{\gamma}$ (recall from Theorem 3 that the P-property of $\hat{\mathbf{Y}}$ is sufficient for the uniqueness of the NE of $\mathcal{G}_\boldsymbol{\gamma}$). According to the sequential IWFA, given $\mathbf{p}^{(n)}(\boldsymbol{\gamma})$ at iteration $n \geq 1$, at the beginning of iteration $n + 1$, the optimal transmission strategy $\mathbf{p}_q^{(n+1)}(\boldsymbol{\gamma})$ computed sequentially by every user q is the unique solution to the following maximization problem:

$$\begin{aligned} & \underset{\mathbf{p}_q}{\text{maximize}} \quad r_q \left(\mathbf{p}_1^{(n+1)}, \dots, \mathbf{p}_{q-1}^{(n+1)}, \mathbf{p}_q, \mathbf{p}_{q+1}^{(n)}, \dots, \mathbf{p}_Q^{(n)} \right) \\ & \quad - \sum_{k=1}^N \gamma_q(k) p_q(k) \\ & \text{subject to} \quad \mathbf{p}_q \in \hat{\mathcal{P}}_q \end{aligned} \quad (61)$$

with $\hat{\mathcal{P}}_q$ and $r_q(\mathbf{p})$ defined in (1) and (3), respectively. For the sake of notation, we denote by $\widehat{\text{mi}}_q^{(n)}(k)$ the (normalized) multiuser interference-plus-noise PSD over the subcarrier k , measured by user q at the beginning of iteration $n + 1$, defined as (hereafter we omit the dependence of the power vector \mathbf{p} on $\boldsymbol{\gamma}$)

$$\widehat{\text{mi}}_q^{(n)}(k) \triangleq \hat{\sigma}_q^2(k) + \sum_{r=1}^{q-1} \left| \hat{H}_{qr}(k) \right|^2 p_r^{(n+1)}(k) \quad (62)$$

$$+ \sum_{r=q+1}^Q \left| \hat{H}_{qr}(k) \right|^2 p_r^{(n)}(k) > 0. \quad (63)$$

Similarly, at the NE $\mathbf{p}^* = \mathbf{p}^*(\boldsymbol{\gamma})$ of game $\mathcal{G}_\boldsymbol{\gamma}$ we have

$$\begin{aligned} \widehat{\text{mi}}_q^*(k) & \triangleq \hat{\sigma}_q^2(k) + \sum_{r=1}^{q-1} \left| \hat{H}_{qr}(k) \right|^2 p_r^*(k) \\ & \quad + \sum_{r=q+1}^Q \left| \hat{H}_{qr}(k) \right|^2 p_r^*(k) > 0. \end{aligned} \quad (64)$$

We focus now on the definition of a proper error iterates generated by the algorithm that, under the P-property of matrix $\hat{\mathbf{Y}}$, converges to zero, implying the global convergence of the algorithm to the unique NE of $\mathcal{G}_\boldsymbol{\gamma}$. It follows from the first-order optimality conditions of (61) that, $\forall \mathbf{p}_q \in \hat{\mathcal{P}}_q$,

$$\sum_{k=1}^N \left(p_q(k) - p_q^{(n+1)}(k) \right) \left(-\frac{1}{\widehat{\text{mi}}_q^{(n)} + p_q^{(n+1)}(k)} + \gamma_q(k) \right) \geq 0. \quad (65)$$

Furthermore, it follows from the definition of NE that, $\forall \mathbf{p}_q \in \hat{\mathcal{P}}_q$

$$\sum_{k=1}^N \left(p_q(k) - p_q^*(k) \right) \left(-\frac{1}{\widehat{\text{mi}}_q^*(k) + p_q^*(k)} + \gamma_q(k) \right) \geq 0. \quad (66)$$

Adding (65) evaluated at $\mathbf{p}_q = \mathbf{p}_q^*$ and (66) evaluated at $\mathbf{p}_q = \mathbf{p}_q^{(n+1)}$, we obtain

$$\begin{aligned} & \sum_{k=1}^N \left(p_q^*(k) - p_q^{(n+1)}(k) \right) \times \\ & \left(\frac{1}{\widehat{\text{mi}}_q^*(k) + p_q^*(k)} - \frac{1}{\widehat{\text{mi}}_q^{(n)}(k) + p_q^{(n+1)}(k)} \right) \geq 0. \end{aligned} \quad (67)$$

Since

$$\begin{aligned} & \frac{1}{\widehat{\text{mi}}_q^*(k) + p_q^*(k)} - \frac{1}{\widehat{\text{mi}}_q^{(n)}(k) + p_q^{(n+1)}(k)} \\ & = \frac{\left(p_q^{(n+1)}(k) - p_q^*(k) \right)}{\left(\widehat{\text{mi}}_q^*(k) + p_q^*(k) \right) \left(\widehat{\text{mi}}_q^{(n)}(k) + p_q^{(n+1)}(k) \right)} \\ & \quad + \sum_{r=1}^{q-1} \frac{\left| \hat{H}_{qr}(k) \right|^2 \left(p_r^{(n+1)}(k) - p_r^*(k) \right)}{\left(\widehat{\text{mi}}_q^*(k) + p_q^*(k) \right) \left(\widehat{\text{mi}}_q^{(n)}(k) + p_q^{(n+1)}(k) \right)} \\ & \quad + \sum_{r=q+1}^Q \frac{\left| \hat{H}_{qr}(k) \right|^2 \left(p_r^{(n)}(k) - p_r^*(k) \right)}{\left(\widehat{\text{mi}}_q^*(k) + p_q^*(k) \right) \left(\widehat{\text{mi}}_q^{(n)}(k) + p_q^{(n+1)}(k) \right)} \end{aligned}$$

(67) can be rewritten as shown in (68) at the bottom of the next page. Introducing

$$\hat{\varphi}_q^{(n+1)}(k) \triangleq \sqrt{\left(\widehat{\text{mi}}_q^*(k) + p_q^*(k) \right) \left(\widehat{\text{mi}}_q^{(n)}(k) + p_q^{(n+1)}(k) \right)} \quad (69)$$

let define, as in the proof of Proposition 2 (see Appendix B), the error quantity

$$e_q^{(n+1)}(k) \triangleq \frac{p_q^{(n+1)}(k) - p_q^*(k)}{\hat{\varphi}_q^{(n+1)}(k)} \quad (70)$$

for each $k = 1, \dots, N$ and $q = 1, \dots, Q$. Recalling the scalar $\hat{\varphi}_q^{\max}(k) \triangleq \hat{\sigma}_q^2(k) + \sum_{r=1}^Q |\hat{H}_{qr}(k)|^2 p_r^{\max}(k)$, we have the following lower and upper bounds for $\hat{\varphi}_q^{(n+1)}(k)$:

$$\hat{\sigma}_q^2(k) \leq \hat{\varphi}_q^{(n+1)}(k) \leq \hat{\varphi}_q^{\max}(k). \quad (71)$$

Using (70) and the Cauchy-Schwarz inequality, from (68) we deduce (72), shown at the bottom of the page. Hence, letting

$$\hat{e}_q^{(n+1)} \triangleq \sqrt{\sum_{k=1}^N \left(e_q^{(n+1)}(k) \right)^2}, \quad q = 1, \dots, Q \quad (73)$$

using (71), and dividing by $\hat{e}_q^{(n+1)}$ (assuming it is positive), we obtain from (72)

$$\begin{aligned} \hat{e}_q^{(n+1)} - \sum_{r=1}^{q-1} \left[\max_{1 \leq k \leq N} \frac{|\hat{H}_{qr}(k)|^2 \hat{\varphi}_r^{\max}(k)}{\hat{\sigma}_q^2(k)} \right] \hat{e}_r^{(n+1)} \\ \leq \sum_{r=q+1}^Q \left[\max_{1 \leq k \leq N} \frac{|\hat{H}_{qr}(k)|^2 \hat{\varphi}_r^{\max}(k)}{\hat{\sigma}_q^2(k)} \right] \hat{e}_r^{(n)}. \end{aligned} \quad (74)$$

[The last inequality is clearly valid if $\hat{e}_q^{(n+1)} = 0$.]

Introducing the error vector $\hat{\mathbf{e}}^{(n)} = (\hat{e}_q^{(n)})_{q=1}^Q$ and using the definition of Z-matrix $\hat{\mathbf{\Upsilon}}$ as given in (20), which we can write as

$$\hat{\mathbf{\Upsilon}} = \mathbf{D}(\hat{\mathbf{\Upsilon}}) - \mathbf{L}(\hat{\mathbf{\Upsilon}}) - \mathbf{U}(\hat{\mathbf{\Upsilon}}) \quad (75)$$

where $\mathbf{D}(\hat{\mathbf{\Upsilon}})$, $\mathbf{L}(\hat{\mathbf{\Upsilon}})$, and $\mathbf{U}(\hat{\mathbf{\Upsilon}})$ are, respectively, the diagonal, strictly lower triangular, and strictly upper triangular parts of matrix $|\hat{\mathbf{\Upsilon}}|$, concatenating the inequality in (74) for all $q = 1, \dots, Q$, we obtain the following vector inequality (componentwise inequality):

$$\left[\mathbf{D}(\hat{\mathbf{\Upsilon}}) - \mathbf{L}(\hat{\mathbf{\Upsilon}}) \right] \hat{\mathbf{e}}^{(n+1)} \leq \mathbf{U}(\hat{\mathbf{\Upsilon}}) \hat{\mathbf{e}}^{(n)}, \quad n \geq 1. \quad (76)$$

Since the Z-matrix $\mathbf{D}(\hat{\mathbf{\Upsilon}}) - \mathbf{L}(\hat{\mathbf{\Upsilon}})$ is also a P-matrix (due to the fact that all its principal minors are equal to one), the inverse is well-defined and nonnegative entry-wise [34, Th. 3.11.10]. Hence, (76) is equivalent to

$$\hat{\mathbf{e}}^{(n+1)} \leq \left[\mathbf{D}(\hat{\mathbf{\Upsilon}}) - \mathbf{L}(\hat{\mathbf{\Upsilon}}) \right]^{-1} \mathbf{U}(\hat{\mathbf{\Upsilon}}) \hat{\mathbf{e}}^{(n)}. \quad (77)$$

Moreover, using Lemma 1, the P-property of $\hat{\mathbf{\Upsilon}}$ is equivalent to the spectral condition

$$\rho \left(\left[\mathbf{D}(\hat{\mathbf{\Upsilon}}) - \mathbf{L}(\hat{\mathbf{\Upsilon}}) \right]^{-1} \mathbf{U}(\hat{\mathbf{\Upsilon}}) \right) < 1$$

implying that the sequence $\{\hat{\mathbf{e}}^{(n)}\}_{n=1}^{\infty}$ satisfying (77), and thus the sequence $\{\mathbf{p}^{(n)} - \mathbf{p}^*(\boldsymbol{\gamma})\}_{n=1}^{\infty}$, converges to zero. This proves the global convergence of the sequential IWFA, under the P-property of matrix $\hat{\mathbf{\Upsilon}}$. ■

$$\begin{aligned} \sum_{k=1}^N \frac{\left(p_q^{(n+1)}(k) - p_q^*(k) \right)^2 + \left(p_q^{(n+1)}(k) - p_q^*(k) \right) \sum_{r=1}^{q-1} |\hat{H}_{qr}(k)|^2 \left(p_r^{(n+1)}(k) - p_r^*(k) \right)}{\left(\widehat{\text{mi}}_q^*(k) + p_q^*(k) \right) \left(\widehat{\text{mi}}_q^{(n)}(k) + p_q^{(n+1)}(k) \right)} \\ \leq \sum_{k=1}^N \frac{\left(p_q^*(k) - p_q^{(n+1)}(k) \right) \sum_{r=q+1}^Q |\hat{H}_{qr}(k)|^2 \left(p_r^{(n)}(k) - p_r^*(k) \right)}{\left(\widehat{\text{mi}}_q^*(k) + p_q^*(k) \right) \left(\widehat{\text{mi}}_q^{(n)}(k) + p_q^{(n+1)}(k) \right)}. \end{aligned} \quad (68)$$

$$\begin{aligned} \sum_{k=1}^N \left(e_q^{(n+1)}(k) \right)^2 - \left(\sum_{k=1}^N \left(e_q^{(n+1)}(k) \right)^2 \right)^{1/2} \sum_{r=1}^{q-1} \max_{1 \leq k \leq N} \left\{ \frac{|\hat{H}_{qr}(k)|^2 \hat{\varphi}_r^{(n+1)}(k)}{\hat{\varphi}_q^{(n+1)}(k)} \right\} \left(\sum_{k=1}^N \left(e_r^{(n+1)}(k) \right)^2 \right)^{1/2} \\ \leq \sum_{r=q+1}^N \left(\sum_{k=1}^N \left(e_q^{(n+1)}(k) \right)^2 \right)^{1/2} \max_{1 \leq k \leq N} \left\{ \frac{|\hat{H}_{qr}(k)|^2 \hat{\varphi}_r^{(n+1)}(k)}{\hat{\varphi}_q^{(n+1)}(k)} \right\} \left(\sum_{k=1}^N \left(e_r^{(n)}(k) \right)^2 \right)^{1/2}. \end{aligned} \quad (72)$$

APPENDIX D
PROOF OF PROPOSITION 8

Suppose $\hat{\mathbf{Y}} \succ \mathbf{0}$ [recall that this implies that \mathbf{F} is a strongly monotone function on $\hat{\mathcal{P}}$ with strong monotonicity modulus $\hat{c}_{\text{sm}}(\mathbf{F})$, see Proposition 2]. For any given $\boldsymbol{\gamma} \in \mathbb{R}_+^{P(N+1)}$, let $\mathbf{z}^* = \mathbf{z}^*(\boldsymbol{\gamma})$ be the unique solution of the VI($\hat{\mathcal{P}}, \mathbf{F} - \boldsymbol{\gamma}$). Note that $\mathbf{z}^*(\boldsymbol{\gamma})$ is related to the unique NE $\mathbf{p}^*(\boldsymbol{\gamma})$ of game \mathcal{G}_γ in (10)—the solution of the VI($\hat{\mathcal{P}}, \mathbf{F} + \boldsymbol{\gamma}$)—by $\mathbf{p}^*(\boldsymbol{\gamma}) = \mathbf{z}^*(-\boldsymbol{\gamma})$. Furthermore, $\mathbf{z}^*(\boldsymbol{\gamma})$ is a co-coercive function of $\boldsymbol{\gamma}$ with modulus $\hat{c}_{\text{sm}}(\mathbf{F})$ [32, Prop. 2.3.1], i.e.

$$\begin{aligned} & (\boldsymbol{\gamma}^{(1)} - \boldsymbol{\gamma}^{(2)})^T (\mathbf{z}^*(\boldsymbol{\gamma}^{(1)}) - \mathbf{z}^*(\boldsymbol{\gamma}^{(2)})) \\ & \geq \hat{c}_{\text{sm}}(\mathbf{F}) \left\| \mathbf{z}^*(\boldsymbol{\gamma}^{(1)}) - \mathbf{z}^*(\boldsymbol{\gamma}^{(2)}) \right\|_2^2, \quad \forall \boldsymbol{\gamma}^{(1)}, \boldsymbol{\gamma}^{(2)}. \end{aligned} \quad (78)$$

To establish the latter inequality, note that we have

$$\left(\mathbf{z}^*(\boldsymbol{\gamma}^{(1)}) - \mathbf{z}^*(\boldsymbol{\gamma}^{(2)}) \right)^T \left[\mathbf{F}(\mathbf{z}^*(\boldsymbol{\gamma}^{(2)})) - \boldsymbol{\gamma}^{(2)} \right] \geq 0 \quad (79)$$

and

$$\left(\mathbf{z}^*(\boldsymbol{\gamma}^{(2)}) - \mathbf{z}^*(\boldsymbol{\gamma}^{(1)}) \right)^T \left[\mathbf{F}(\mathbf{z}^*(\boldsymbol{\gamma}^{(1)})) - \boldsymbol{\gamma}^{(1)} \right] \geq 0. \quad (80)$$

Adding these two inequalities, rearranging terms, and using the strong monotonicity of \mathbf{F} , the desired co-coercivity property of $\mathbf{z}^*(\boldsymbol{\gamma})$ follows readily.

Choosing $\boldsymbol{\gamma} = \boldsymbol{\gamma}(\boldsymbol{\lambda})$ as in (11) and using $\mathbf{p}^*(\boldsymbol{\gamma}) = \mathbf{z}^*(-\boldsymbol{\gamma})$, it follows from (78) that, for any two tuples $\boldsymbol{\lambda}^{(1)}, \boldsymbol{\lambda}^{(2)} \in \mathbb{R}_+^{P(N+1)}$

$$\begin{aligned} & - \sum_{q=1}^Q \sum_{k=1}^N \sum_{p=1}^P \left(\left| H_{pq}^{(P,S)}(k) \right|^2 \left(\lambda_p^{(1)}(k) - \lambda_p^{(2)}(k) + \lambda_{p,\text{tot}}^{(1)} - \lambda_{p,\text{tot}}^{(2)} \right) \right. \\ & \quad \left. \cdot \left[p_q^* \left(k; \boldsymbol{\gamma}(\boldsymbol{\lambda}^{(1)}) \right) - p_q^* \left(k; \boldsymbol{\gamma}(\boldsymbol{\lambda}^{(2)}) \right) \right] \right) \\ & \geq \hat{c}_{\text{sm}}(\mathbf{F}) \left\| \mathbf{p}^*(\boldsymbol{\gamma}(\boldsymbol{\lambda}^{(1)})) - \mathbf{p}^*(\boldsymbol{\gamma}(\boldsymbol{\lambda}^{(2)})) \right\|_2^2. \end{aligned} \quad (81)$$

Using the definition of mapping $\Phi(\boldsymbol{\lambda})$ in (29), we deduce from (81) that the following inequalities hold:

$$\begin{aligned} & (\boldsymbol{\lambda}^{(1)} - \boldsymbol{\lambda}^{(2)})^T \left[\Phi(\boldsymbol{\lambda}^{(1)}) - \Phi(\boldsymbol{\lambda}^{(2)}) \right] \\ & \geq \hat{c}_{\text{sm}}(\mathbf{F}) \left\| \mathbf{p}^*(\boldsymbol{\gamma}(\boldsymbol{\lambda}^{(1)})) - \mathbf{p}^*(\boldsymbol{\gamma}(\boldsymbol{\lambda}^{(2)})) \right\|_2^2 \\ & \geq \frac{\hat{c}_{\text{sm}}(\mathbf{F})}{\|\mathbf{H}\|_2^2} \left\| \Phi(\hat{\boldsymbol{\lambda}}^1) - \Phi(\hat{\boldsymbol{\lambda}}^2) \right\|_2^2 \end{aligned} \quad (82)$$

where \mathbf{H} is defined in (30), and $\|\mathbf{H}\|_2$ is a positive constant, calculated from the channels $|H_{pq}^{(P,S)}(k)|^2$, such that

$$\begin{aligned} & \left\| \Phi(\boldsymbol{\lambda}^{(1)}) - \Phi(\boldsymbol{\lambda}^{(2)}) \right\|_2 \\ & \leq \|\mathbf{H}\|_2 \left\| \mathbf{p}^*(\boldsymbol{\gamma}(\boldsymbol{\lambda}^{(1)})) - \mathbf{p}^*(\boldsymbol{\gamma}(\boldsymbol{\lambda}^{(2)})) \right\|_2 \end{aligned}$$

for all $\boldsymbol{\lambda}^{(1)}, \boldsymbol{\lambda}^{(2)} \in \mathbb{R}_+^{P(N+1)}$. The inequality in (82) proves the desired co-coercivity property of $\Phi(\boldsymbol{\lambda})$. ■

APPENDIX E
PROOF OF THEOREM 12

The first statement of the theorem follows from [32, Th. 12.1.2]: The subiteration (41) is just the *Basic Projection Algorithm* [32, Alg. 12.1.1] applied to the NCP($\Phi + \varepsilon_n \mathbf{I}$), whose global convergence to the unique solution of NCP($\Phi + \varepsilon_n \mathbf{I}$) is ensured if the mapping $\Phi + \varepsilon_n \mathbf{I}$ is strongly monotone (with modulus μ) and Lipschitz continuous (with Lipschitz constant L) and the step-size τ is such that $\tau > L^2/(2\mu)$ [32, Th. 12.1.2]. It is not difficult to show that the co-coercivity of the mapping Φ with modulus $c_{\text{Lip}}(\Phi)$ (Proposition 8) implies the strong monotonicity of $\Phi + \varepsilon_n \mathbf{I}$ with modulus $\mu = \varepsilon_n$ and the Lipschitz continuity with constant $L = c_{\text{Lip}}(\Phi) + \varepsilon_n$; which proves the claimed convergence of the subiteration (41).

The proof of the second statement of the theorem is based on [32, Th. 12.2.3]: The unique least-norm solution of the monotone and solvable NCP(Φ) (whose existence is guaranteed by the convexity of the nonempty solution set, implied by the monotonicity of Φ [32, Th. 2.3.5]) is the limit point of the Tikhonov trajectory as the perturbation ε_n tends to zero. ■

REFERENCES

- [1] FCC Spectrum Policy Task Force, Report of the Spectrum Efficiency Working Group Nov. 2002 [Online]. Available: <http://www.fcc.gov/sptf/files/SEWGFinalReport1.pdf>
- [2] J. Mitola, "Cognitive radio for flexible mobile multimedia communication," in *Proc. IEEE Int. Workshop on Mobile Multimedia Commun. (MoMuC 1999)*, San Diego, CA, Nov. 15–17, 1999, pp. 3–10.
- [3] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201–220, Feb. 2005.
- [4] Q. Zhao and B. Sadler, "A survey of dynamic spectrum access," *IEEE Commun. Mag.*, vol. 24, no. 3, pp. 79–89, May 2009.
- [5] 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, May 2009.
- [6] D. P. Palomar and M. Chiang, "Alternative distributed algorithms for network utility maximization: Framework and applications," *IEEE Trans. Autom. Control*, vol. 52, no. 12, pp. 2254–2269, Dec. 2007.
- [7] M. Chiang, C. W. Tan, D. P. Palomar, D. O. Neill, and D. Julian, "Power control by geometric programming," *IEEE Trans. Wireless Commun.*, vol. 6, no. 7, pp. 2640–2651, Jul. 2007.
- [8] Z.-Q. Luo and S. Zhang, "Spectrum management: Complexity and duality," *J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 57–72, Feb. 2008.
- [9] R. Cendrillon, W. Yu, M. Moonen, J. Verliden, and T. Bostoen, "Optimal multi-user spectrum management for digital subscriber lines," *IEEE Trans. Commun.*, vol. 54, no. 5, pp. 922–933, May 2006.
- [10] W. Yu and R. Lui, "Dual methods for nonconvex spectrum optimization of multicarrier systems," *IEEE Trans. Commun.*, vol. 54, no. 7, pp. 1310–1322, Jul. 2006.
- [11] S. Hayashi and Z.-Q. Luo, "Spectrum management for interference-limited multiuser communication systems," *IEEE Trans. Inf. Theory*, vol. 55, no. 3, pp. 1153–1175, Mar. 2009.
- [12] A. Leshem and E. Zehavi, "Game theory and the frequency selective interference channel—A tutorial," *IEEE Signal Process. Mag.* [Online]. Available: <http://arxiv.org/abs/0903.2174>, submitted for publication
- [13] Y. Xing, C. N. Mathur, M. Haleem, R. Chandramouli, and K. Subbalakshmi, "Dynamic spectrum access with QoS and interference temperature constraints," *IEEE Trans. Mobile Comput.*, vol. 6, no. 4, pp. 423–433, Apr. 2007.
- [14] Q. Lu, W. Wang, W. Wang, and T. Peng, "Asynchronous distributed power control under interference temperature constraints," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, New Orleans, LA, Nov. 30–Dec. 4 2008.
- [15] W. Wang, T. Peng, and W. Wang, "Optimal power control under interference temperature constraints in cognitive radio network," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC 2007)*, Hong Kong, HK, Mar. 11–15, 2008.

- [16] M. J. Osborne and A. Rubinstein, *A Course in Game Theory*. Cambridge, MA: MIT Press, Jul. 2004.
- [17] W. Yu, G. Ginis, and J. M. Cioffi, "Distributed multiuser power control for digital subscriber lines," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 5, pp. 1105–1115, June 2002.
- [18] G. Scutari, "Competition and cooperation in wireless communication networks," Ph.D. dissertation, INFOCOM Dep., Univ. Rome "La Sapienza", Rome, Italy, Nov. 2004.
- [19] N. Yamashita and Z.-Q. Luo, "A nonlinear complementarity approach to multiuser power control for digital subscriber lines," *Optim. Methods and Software*, vol. 19, pp. 633–652, 2004.
- [20] R. Cendrillon, J. Huang, M. Chiang, and M. Moonen, "Autonomous spectrum balancing for digital subscriber lines," *IEEE Trans. Signal Process.*, vol. 55, no. 8, pp. 4241–4257, Aug. 2007.
- [21] Z.-Q. Luo and J.-S. Pang, "Analysis of iterative waterfilling algorithm for multiuser power control in digital subscriber lines," *EURASIP J. Appl. Signal Process.*, vol. 2006, pp. 1–10, May 2006.
- [22] G. Scutari, D. P. Palomar, and S. Barbarossa, "Optimal linear precoding strategies for wideband noncooperative systems based on game theory—Part I: Nash equilibria," *IEEE Trans. Signal Process.*, vol. 56, no. 3, pp. 1230–1249, Mar. 2008.
- [23] G. Scutari, D. P. Palomar, and S. Barbarossa, "Optimal linear precoding strategies for wideband noncooperative systems based on game theory—Part II: Algorithms," *IEEE Trans. Signal Process.*, vol. 56, no. 3, pp. 1250–1267, Mar. 2008.
- [24] G. Scutari, D. P. Palomar, and S. Barbarossa, "Asynchronous iterative water-filling for Gaussian frequency-selective interference channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 7, pp. 2868–2878, Jul. 2008.
- [25] J.-S. Pang, G. Scutari, F. Facchinei, and C. Wang, "Distributed power allocation with rate constraints in Gaussian parallel interference channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3471–3489, Aug. 2008.
- [26] F. Wang, M. Krunz, and S. Cui, "Price-based spectrum management in cognitive radio networks," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 74–87, Feb. 2008.
- [27] J. Huang, R. Berry, and M. L. Honig, "Distributed interference compensation for wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 5, pp. 1074–1084, May 2006.
- [28] O. Simeone, I. Stanojev, S. Savazzi, Y. Bar-Ness, U. Spagnolini, and R. Pickholtz, "Spectrum leasing to cooperating *ad hoc* secondary networks," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 1, pp. 203–213, Jan. 2008.
- [29] M.-H. Lin, J.-F. Tsai, and Y. Ye, "Budget allocation in a competitive communication spectrum economy," *EURASIP J. Appl. Signal Process.* 2009 [Online]. Available: http://www.stanford.edu/~yye/spectrumbudget7_urasip1.pdf, to be published
- [30] Z. Ji and K. Liu, "Dynamic spectrum sharing: A game theoretical overview," *IEEE Commun. Mag.*, vol. 45, no. 5, pp. 88–94, May 2007.
- [31] D. Niyato and E. Hossain, "Spectrum trading in cognitive radio networks: A market—Equilibrium—Based approach," *IEEE Wireless Commun.*, vol. 15, no. 6, pp. 71–80, Dec. 2008.
- [32] F. Facchinei and J.-S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problem*. New York: Springer-Verlag, 2003.
- [33] R. Etkin, A. Parekh, and D. Tse, "Spectrum sharing for unlicensed bands," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 3, pp. 517–528, Apr. 2007.
- [34] R. W. Cottle, J.-S. Pang, and R. E. Stone, *The Linear Complementarity Problem*. Cambridge, U.K.: Cambridge Academic, 1992.
- [35] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.
- [36] Y. Noam and A. Leshem, "Iterative power pricing for distributed spectrum coordination in DSL," *IEEE Trans. Commun.*, vol. 57, no. 4, Apr. 2009.
- [37] F. Meshkati, V. Poor, and S. Schwartz, "Energy-efficient resource allocation in wireless networks," *IEEE Signal Process. Mag.*, vol. 24, no. 3, pp. 58–68, May 2007.
- [38] C. U. Saraydar, N. B. Mandayam, and D. J. Goodman, "Efficient power control via pricing in wireless data networks," *IEEE Trans. Commun.*, vol. 50, no. 2, pp. 291–303, Feb. 2002.
- [39] R. D. Yates, "A framework for uplink power control in cellular radio systems," *IEEE J. Sel. Areas Commun.*, vol. 52, no. 7, pp. 1341–1347, Sep. 1995.
- [40] C. W. Sung and K. K. Leung, "A generalized framework for distributed power control in wireless networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2625–2635, Jul. 2005.
- [41] A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*. Philadelphia, PA: Society for Industrial Mathematics (SIAM), 1987.
- [42] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1985.
- [43] D. P. Palomar and J. Fonollosa, "Practical algorithms for a family of waterfilling solutions," *IEEE Trans. Signal Process.*, vol. 53, no. 2, pp. 686–695, Feb. 2005.
- [44] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*, 2nd ed. New York: Athena Scientific, 1989.
- [45] R. T. Rockafellar, "Monotone operators and the proximal point algorithm," *SIAM J. Control Optim. (SICON)*, vol. 14, no. 5, pp. 877–898, Aug. 1976.
- [46] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proc. IEEE*, vol. 95, no. 1, pp. 215–223, Jan. 2007.
- [47] G. Scutari, S. Barbarossa, and L. Pescosolido, "Distributed decision through self-synchronizing sensor networks in the presence of propagation delays and asymmetric channels," *IEEE Trans. Signal Process.*, vol. 56, no. 4, pp. 1667–1684, Apr. 2008.

Jong-Shi Pang received the Ph.D. degree in operations research from Stanford University, Stanford, CA.

He is the Caterpillar Professor and Head of the Department of Industrial and Enterprise Systems Engineering at the University of Illinois at Urbana-Champaign. Prior to this position, he was the Margaret A. Darrin Distinguished Professor in Applied Mathematics at Rensselaer Polytechnic Institute, Troy, NY. He had taught at The Johns Hopkins University, Baltimore, MD, The University of Texas at Dallas, and Carnegie-Mellon University, Pittsburgh, PA. His research interests are in continuous optimization and equilibrium programming and their applications in engineering, economics, and finance. Among the current projects, he is studying various extensions of the basic Nash equilibrium problem, including the Stackelberg game and its multileader generalization, and the dynamic version of the Nash problem. The mathematical tool for the latter problem is a new class of dynamical systems known as differential variational inequalities, which provides a powerful framework for dealing with applications that involve dynamics, unilateral constraints, and disjunction.

Dr. Pang has received several awards and honors, most notably the George B. Dantzig Prize in 2003 jointly awarded by the Mathematical Programming Society and the Society for Industrial and Applied Mathematics and the 1994 Lanchester Prize by the Institute for Operations Research and Management Science. He was recently selected as a member in the inaugural 2009 class of the SIAM Fellows. He is an ISI highly cited author in the mathematics category.



Gesualdo Scutari (S'05–M'06) received the Electrical Engineering and Ph.D. degrees (both with honors) from the University of Rome "La Sapienza," Rome, Italy, in 2001 and 2004, respectively.

During 2003, he held a visiting research appointment with the Department of Electrical Engineering and Computer Sciences, University of California at Berkeley; during 2007 and 2008, he was a Postdoctoral Fellow with the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology (HKUST); during 2009, he

was a Research Associate with the Department of Electronic and Computer Engineering, HKUST and a Postdoctoral Researcher with the INFOCOM Department, University of Rome, "La Sapienza." He is currently a Research Associate with the Department of Industrial and Enterprise Systems Engineering, University of Illinois at Urbana-Champaign. He has participated in two European projects on multiantenna systems and multihop systems (IST SATURN and IST-ROMANTIK). He is currently involved in the European project WINSOC, on wireless sensor networks, and in the European project SURFACE, on reconfigurable air interfaces for wideband multiantenna communication systems. His primary research interests include applications of convex optimization theory, game theory, and variational inequality theory to signal processing and communications; sensor networks; and distributed decisions.

Dr. Scutari received the 2006 Best Student Paper Award at the International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2006.



Daniel P. Palomar (S'99–M'03–SM'08) received the Electrical Engineering and Ph.D. degrees (both with honors) from the Technical University of Catalonia (UPC), Barcelona, Spain, in 1998 and 2003, respectively.

Since 2006, he has been an Assistant Professor with the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology (HKUST). He has held several research appointments, namely, at King's College London (KCL), London, U.K.; the Technical University of Catalonia (UPC), Barcelona, Spain; Stanford University, Stanford, CA; the Telecommunications Technological Center of Catalonia (CTTC), Barcelona; the Royal Institute of Technology (KTH), Stockholm, Sweden; the University of Rome "La Sapienza," Rome, Italy; and Princeton University, Princeton, NJ. His current research interests include applications of convex optimization theory, game theory, and variational inequality theory to signal processing and communications.

Dr. Palomar is an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING, a Guest Editor of the IEEE *Signal Processing Magazine* 2010 Special Issue on Convex Optimization for Signal Processing, was a Guest Editor of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS 2008 Special

Issue on Game Theory in Communication Systems, as well as the Lead Guest Editor of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS 2007 Special Issue on Optimization of MIMO Transceivers for Realistic Communication Networks. He serves on the IEEE Signal Processing Society Technical Committee on Signal Processing for Communications (SPCOM). He is a recipient of a 2004/2006 Fulbright Research Fellowship; the 2004 Young Author Best Paper Award by the IEEE Signal Processing Society; the 2002–2003 Best Ph.D. Prize in Information Technologies and Communications by UPC; the 2002–2003 Rosina Ribalta First Prize for the Best Doctoral Thesis in Information Technologies and Communications by the Epson Foundation; and the 2004 prize for the Best Doctoral Thesis in Advanced Mobile Communications by the Vodafone Foundation and COIT.

Francisco Facchinei received the Ph.D. degree in system engineering from the University of Rome, "La Sapienza," Rome, Italy.

He is full professor of Operations Research, Engineering Faculty, University of Rome, "La Sapienza." His research interests focus on theoretical and algorithmic issues related to nonlinear optimization, variational inequalities, complementarity problems, equilibrium programming, and computational game theory.