MIMO Cognitive Radio: A Game Theoretical Approach

Gesualdo Scutari, *Member, IEEE*, and Daniel P. Palomar, *Senior Member, IEEE*

***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 and analyze a totally decentralized approach, based on game theory, to design cognitive MIMO transceivers, who compete with each other to maximize their information rate. The formulation incorporates constraints on the transmit power as well as null and/or soft shaping constraints on the transmit covariance matrix, so that the interference generated by secondary users be confined within the temperature-interference limit required by the primary users. We provide a unified set of conditions that guarantee the uniqueness and global asymptotic stability of the Nash equilibrium of all the proposed games through totally distributed and asynchronous algorithms. Interestingly, the proposed algorithms overcome the main drawback of classical waterfilling based algorithms—the violation of the temperature-interference limit—and they have the desired features required for CR applications, such as low-complexity, distributed implementation, robustness against missing or outdated updates of the users, and fast convergence behavior.**

***Index Terms*—Cognitive radio, game theory, interference constraints, MIMO Gaussian interference channel, Nash equilibrium, totally asynchronous distributed algorithms.**

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 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 the 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]. Within this context, alternative approaches have been considered to allow concurrent communications (see [5] for a recent tutorial on the topic).

In this article, we focus on *opportunistic* communications in hierarchical cognitive networks (also known in the CR literature as *interweave communications* [5]), as they seem to be the most suitable for the current spectrum management policies and legacy wireless systems [4]. In particular, our interest is in devising the most appropriate form of concurrent communications of cognitive users competing over the physical resources that primary users make available. Looking at the opportunistic communication paradigm from a broad signal processing perspective, the secondary users are allowed to transmit over a multidimensional space, whose coordinates may represent time slots, frequency bins, and/or angles, with the goal of finding out the most appropriate transmission strategy exploring all available degrees of freedom, under the constraint of inducing a limited interference, or no interference at all, on the primary users.

One approach to devise such systems would be using global optimization techniques, under the framework of network utility maximization (NUM) [see, e.g., [6]]. However, recent results in [7] have shown that the nonconvex 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. Roughly speaking, this means that there is no hope to obtain an algorithm, even centralized, that can efficiently compute the exact globally optimal (i.e., Pareto dominant) solution. Consequently, several attempts have been pursued in the literature to deal with the nonconvexity of such a problem. Some works proposed suboptimal or closed-to-optimal algorithms based on duality theory (see, e.g., [7]–[10]). Others works applied the theory of cooperative games (based on the Nash bargaining optimality criterion) to compute, under technical conditions and/or simplifying assumptions on the users' transmission strategies, the largest achievable rate region of the system. Two good tutorials on the topic are [11] and [12]. However, current algorithms based on global optimization or Nash bargaining solution lack any mechanism to control the amount of aggregate interference generated by the transmitters. Moreover, they are centralized and computationally expensive. This raises some practical issues that are insurmountable

Manuscript received January 25, 2009; accepted August 07, 2009. First published September 18, 2009; current version published January 13, 2010. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Roberto Lopez-Valcarce. This work was supported by the NSFC/RGC N_HKUST604/08 research grant.

The authors are with the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Hong Kong (e-mail: eealdo@ust.hk; palomar@ust.hk).

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.2009.2032039

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. For these reasons, in this paper, we follow a different approach and we 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 noncooperative game theory (see, e.g., [13] and [14]).

Because of the inherently competitive nature of multi-user systems, it is not surprising indeed that game theory has been already adopted to solve distributively many resource allocation problems in communications. An early application of noncooperative game theory in a communication system is [15], where the information rates of the users were maximized with respect to the power allocation in a DSL system modeled as a frequency-selective (in practice, multicarrier) Gaussian interference channel. Extension of the basic problem to ad hoc frequency-selective and MIMO networks were given in [16]–[20] and [21]–[24], respectively. However, the power control schemes proposed in the aforementioned papers have been recognized not to be applicable to CR systems because they lack any mechanism to control the amount of interference generated by the secondary users on the primary users [3].

Thus, a distributed design of cognitive MIMO transceivers in a CR network composed of *many* primary and secondary users that preserves the QoS of the primary users is, at the best of our knowledge, up-to-date missing. In this paper we fill this gap and propose a novel decentralized approach, based on noncooperative game theory, suitable to design cognitive MIMO transceivers within the paradigm of opportunistic communications. More specifically, adopting an information theoretical perspective, we formulate the resource allocation problem among secondary users as a strategic noncooperative game, where each player (transmit–receive pair) competes against the others to maximize the information rate over his own MIMO channel, under constraints on the transmit power and on the maximum interference induced against the primary users. The interference constraints are established by the primary users in terms of *null* and/or *soft shaping* constraints on the transmit covariance matrix of secondary users. Null constraints are enforced to prevent secondary users from transmitting over prescribed subspaces, which can identify, for example, portions of the spectrum or spatial angular directions. Soft shaping constraints allow the transmissions of secondary users, provided that the interference that they generate in specific frequency bands and/or geographic locations be confined within the temperature-interference limits [3]. The main question is then to establish whether, and under what conditions, the overall system can eventually converge to an equilibrium—the NE—from which every user is not willing to unilaterally deviate as this would lead to a performance loss.

The presence of interference constraints makes the analysis of the proposed games much harder than that of classical rate maximization game over the MIMO interference channel [22], [24]. Nevertheless, we show that the proposed games always admit a NE, irrespective of the transmission strategies of primary users and the null/soft shaping constraints. Moreover, using a classical signal processing tool—the projection operator onto a signal subspace—we prove that, even in the presence of both null and soft shaping constraints, all the Nash equilibria can be equivalently rewritten as the solutions of a MIMO *waterfilling* nonlinear fixed-point equation, provided that the original channel is properly modified. This result is also instrumental to derive sufficient conditions guaranteeing the uniqueness of the NE of the proposed games. To reach these Nash equilibria we then propose a class of totally distributed and (possibly) asynchronous algorithms (in the sense of [19], [25]) along with their convergence properties. Interestingly, the proposed framework is sufficiently general to incorporate, as special cases, the algorithms proposed in the literature to solve the rate maximization game in MIMO [22] (or SISO frequency-selective [15], [18], [19], [26]) Gaussian interference channels. However, our proposed algorithms overcome the main drawback of classical waterfilling based algorithms [15], [18], [19], [26]—the violation of the interference temperature limits [3]—and they have all the desired features required for cognitive radio applications, such as low-complexity, distributed and (possibly) asynchronous implementation, robustness against missing or outdated updates of the users, and fast convergence behavior.

The paper is organized as follows. Section II gives the system model and introduces the null and soft interference constraints as a distributed mechanism to control the maximum interference induced against the primary users by the secondary users. As a warm-up, in Section III, we start formulating the optimization problem as a strategic noncooperative game under transmit power constraints and null constraints. Section IV considers an alternative game formulation to include null constraint, with improved convergence properties. The game with both null and soft shaping constraint is studied in Section V. Section VI provides some numerical results validating our theoretical findings, and Section VII draws some conclusions.

Notation: The following notation is used in the paper. Uppercase and lowercase boldface denote matrices and vectors respectively. The operators $(\cdot)^*$, $(\cdot)^H$, $(\cdot)^\sharp$, $\mathcal{E}\{\cdot\}$, and $\mathrm{Tr}(\cdot)$ are conjugate, Hermitian, Moore–Penrose pseudoinverse [27], expectation, and trace operators, respectively. The range space and null space are denoted by $\mathcal{R}(\cdot)$ and $\mathcal{N}(\cdot)$, respectively. The maximum and the minimum eigenvalue of a Hermitian matrix $\mathbf{A}$ are denoted by $\lambda_{\max}(\mathbf{A})$ and $\lambda_{\min}(\mathbf{A})$, respectively. The operators $\leq$ and $\geq$ for vectors and matrices are defined component-wise, while $\mathbf{A} \succeq \mathbf{B}$ (or $\mathbf{A} \preceq \mathbf{B}$) means that $\mathbf{A} - \mathbf{B}$ is positive (or negative) semidefinite. The operator $\mathrm{Diag}(\cdot)$ is the diagonal matrix with the same diagonal elements as the matrix (or vector) argument; $\mathrm{bdiag}(\mathbf{A}, \mathbf{B}, \ldots)$ is the diagonal matrix, whose diagonal blocks are the matrices $\mathbf{A}, \mathbf{B}, \ldots$; the operator $\perp$ means that two (complex) vectors $\mathbf{x}$ and $\mathbf{y}$ are orthogonal, i.e., $\mathbf{x} \perp \mathbf{y} \Leftrightarrow \mathbf{x}^H\mathbf{y} = 0$. The operators $(\cdot)^+$ and $[\cdot]_a^b$, with $0 \leq a \leq b$, are defined as $(x)^+ \triangleq \max(0, x)$ and $(x)^+ \triangleq \min(b, \max(x, a))$, respectively; when the argument of the operators is a vector or a matrix, then they are assumed to be applied component-wise. The spectral radius of a matrix $\mathbf{A}$ is denoted by $\rho(\mathbf{A})$, and is defined as $\rho(\mathbf{A}) \triangleq \max\{|\lambda| : \lambda \in \sigma(\mathbf{A})\}$ with $\sigma(\mathbf{A})$ denoting the spectrum (set of eigenvalues) of $\mathbf{A}$ [28]. The orthogonal projection onto the null space (or the range space) of matrix $\mathbf{A}$ is denoted by $\mathbf{P}_{\mathcal{N}(\mathbf{A})} = \mathbf{N}_A\left(\mathbf{N}_A^H\mathbf{N}_A\right)^{-1}\mathbf{N}_A^H$ (or $\mathbf{P}_{\mathcal{R}(\mathbf{A})} = \mathbf{R}_A\left(\mathbf{R}_A^H\mathbf{R}_A\right)^{-1}\mathbf{R}_A^H$), where $\mathbf{N}_A$ (or $\mathbf{R}_A$) is any matrix whose columns are linear independent vectors spanning $\mathcal{N}(\mathbf{A})$ (or $\mathcal{R}(\mathbf{A})$) [28]. The operator $[\mathbf{X}]_{\mathcal{Q}} = \arg\min_{\mathbf{Z}\in\mathcal{Q}} \|\mathbf{Z}-\mathbf{X}\|_F$ denotes the matrix projection with respect to the Frobenius norm of matrix $\mathbf{X}$ onto the (convex) set $\mathcal{Q}$, where $\|\mathbf{X}\|_F$ is defined as $\|\mathbf{X}\|_F \triangleq \left(\mathrm{Tr}(\mathbf{X}^H\mathbf{X})\right)^{1/2}$ [28]. We denote by $\mathbf{I}_n$ the $n\times n$ identity matrix. The sets $\mathbb{C}, \mathbb{R}, \mathbb{R}_+, \mathbb{R}_{++}, \mathbb{N}_+$, and $\mathbb{S}_+^{n\times n}$ (or $\mathbb{S}_{++}^{n\times n}$) stand for the set of complex, real, nonnegative real, positive real, nonnegative integer numbers, and $n\times n$ complex positive semidefinite (or definite) matrices, respectively.

II. System Model and Problem Formulation

We consider a multiuser environment composed of Q secondary users and several primary users, sharing the same physical resources, e.g., time, frequency, and space. The setup may include MIMO peer-to-peer links, multiple access, or broadcast (single or multi-antenna, flat or frequency-selective) channels. 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 Gaussian vector interference channel, where the transmission over the generic qth MIMO channel with n_{T_q} transmit and n_{R_q} receive dimensions is given by the following baseband signal model:

$$\mathbf{y}_q = \mathbf{H}_{qq}\mathbf{x}_q + \sum_{r\neq q}\mathbf{H}_{rq}\mathbf{x}_r + \mathbf{n}_q \qquad (1)$$

where $\mathbf{x}_q \in \mathbb{C}^{n_{T_q}}$ is the vector transmitted by source q, $\mathbf{y}_q \in \mathbb{C}^{n_{R_q}}$ is the vector received by destination q, $\mathbf{H}_{qq} \in \mathbb{C}^{n_{R_q}\times n_{T_q}}$ is the channel matrix between the qth transmitter and the intended receiver, $\mathbf{H}_{rq} \in \mathbb{C}^{n_{R_q}\times n_{T_r}}$ is the cross-channel matrix between source r and destination q, and $\mathbf{n}_q \in \mathbb{C}^{n_{R_q}}$ is a zero-mean circularly symmetric complex Gaussian noise vector with arbitrary (nonsingular) covariance matrix $\mathbf{R}_{n_q}$, collecting the effect of both thermal noise and interference generated by the primary users. The first term in the right-hand side of (1) is the useful signal for link q, the second and third terms represent the multi-user interference (MUI) received by secondary user q and caused from the other secondary users and the primary users, respectively. The power constraint for each transmitter is

$$\mathcal{E}\left\{\|\mathbf{x}_q\|_2^2\right\} = \mathsf{Tr}\left(\mathbf{Q}_q\right) \leq P_q \qquad (2)$$

where $\mathbf{Q}_q$ denotes the covariance matrix of the symbols transmitted by user q and P_q is the transmit power in units of energy per transmission.

The model in (1) represents a fairly general MIMO setup, describing multiuser transmissions over multiple channels, which may represent frequency channels (as in OFDM systems) [17]–[19], time slots (as in TDMA systems) [17], [18], or spatial channels (as in transmit–receive beamforming systems) [22].

Due to distributed nature of the CR system, where 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. Each channel is assumed to change sufficiently slowly to be considered fixed during the whole transmission, so that the information theoretical results are meaningful. Moreover, perfect channel state information at both transmitter and receiver sides of each link is assumed. This includes the direct channel $\mathbf{H}_{qq}$ (but not the cross-channels $\{\mathbf{H}_{rq}\}_{r\neq q}$ from the other secondary users) as well as the covariance matrix of the noise plus MUI:

$$\mathbf{R}_{-q}(\mathbf{Q}_{-q}) \triangleq \mathbf{R}_{n_q} + \sum_{r\neq q}\mathbf{H}_{rq}\mathbf{Q}_r\mathbf{H}_{rq}^H. \qquad (3)$$

In some of the proposed algorithms, each secondary user is also assumed to perfectly estimate the cross-channels between his own transmitter and the primary receivers. Within the assumptions made above, the maximum information rate on link q for a given set of user covariance matrices $\mathbf{Q}_1,\ldots,\mathbf{Q}_Q$, is [29]

$$R_q(\mathbf{Q}_q,\mathbf{Q}_{-q}) = \log\det\left(\mathbf{I} + \mathbf{H}_{qq}^H\mathbf{R}_{-q}^{-1}(\mathbf{Q}_{-q})\mathbf{H}_{qq}\mathbf{Q}_q\right) \qquad (4)$$

where $\mathbf{Q}_{-q} \triangleq (\mathbf{Q}_r)_{r\neq q}$ is the set of all the users covariance matrices, except the qth one.

A. Interference Constraints

Differently from traditional static or centralized spectrum assignment, opportunistic communications in CR systems enable secondary users 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]. How to impose interference constraints on the secondary users is a complex and open regulatory issue. Both deterministic and probabilistic interference constraints have been suggested in the literature [3], [4]. In this paper, we consider deterministic interference constraints, expressed in the following very general form:

—*Null shaping constraints*:

$$\mathbf{U}_q^H\mathbf{Q}_q = \mathbf{0} \qquad (5)$$

where $\mathbf{U}_q \in \mathbb{C}^{n_{T_q}\times r_{U_q}}$ is a tall matrix whose columns represent the spatial and/or the frequency "directions" along with user q is not allowed to transmit. We assume, without loss of generality (w.l.o.g.) that each matrix $\mathbf{U}_q$ is full-column rank and, to avoid the trivial solution $\mathbf{Q}_q = \mathbf{0}$, $r_{U_q} < n_{T_q}$.

—*Soft and peak power shaping constraints*:

$$\mathsf{Tr}\left(\mathbf{G}_q^H\mathbf{Q}_q\mathbf{G}_q\right) \leq P_{\mathrm{SU},q}^{\mathrm{ave}} \text{ and } \lambda_{\max}\left(\mathbf{G}_q^H\mathbf{Q}_q\mathbf{G}_q\right) \leq P_{\mathrm{SU},q}^{\mathrm{peak}} \qquad (6)$$

which represent a relaxed version of the null constraints with a constraint on the total average and peak average power radiated along the range space of matrix $\mathbf{G}_q \in$

$\mathbb{C}^{n_{T_q} \times n_{G_q}}$, where $P_{\text{SU},q}^{\text{ave}}$ and $P_{\text{SU},q}^{\text{peak}}$ are the maximum average and average peak power respectively that can be transmitted along the spatial and/or the frequency directions spanned by $\mathbf{G}_q$.

The null constraints are motivated in practice by the interference-avoiding paradigm in CR communications (also called white-space filling approach) [5], [30]: CR nodes sense the spatial, temporal or spectral voids and adjust their transmission strategy to fill in the sensed white spaces. This white-space filling strategy is often considered to be the key motivation for the introduction and development of CR idea and has already been adopted as a core platform in emerging wireless access standards such as the IEEE 802.22-Wireless Regional Area Networks (WRANs) [31], [32]. On a note of interest, the FCC is in the second phase testing white-space devices from a number of companies and research labs, meaning that the white-space filling paradigm is approaching. Observe that the structure of the null constraints in (5) is a very general form and includes, as particular cases, the imposition of nulls over 1) frequency bands occupied by the primary users (the range space of $\mathbf{U}_q$ coincides with the subspace spanned by a set of IDFT vectors); 2) the time slots used by the primary users (the set of canonical vectors); 3) angular directions identifying the primary receivers as observed from the secondary transmitters (the set of steering vectors representing the directions of the primary receivers as observed from the secondary transmitters).

While the white-space filling paradigm demands that cognitive transmissions be orthogonal (in space, time, or frequency) to primary transmissions, opportunistic communications involve simultaneous transmissions between primary and secondary users, provided that the required QoS of the primary users is preserved (also called interference-temperature controlled transmissions [3], [30], [33]). This can be done using soft shaping constraints expressed in (6) that represent a constraint on the total average and peak average power allowed to be radiated (projected) along the directions spanned by the column space of matrix $\mathbf{G}_q$. For example, in a MIMO setup, the matrix $\mathbf{G}_q$ in (6) may contain, in its columns, the steering vectors identifying the directions of the primary receivers. By using these constraints, we assume that the power thresholds $P_{\text{SU},q}^{\text{ave}}$ and $P_{\text{SU},q}^{\text{peak}}$ at each secondary transmitter have been fixed in advance (imposed, e.g., by the network service provider, or legacy systems, or the spectrum body agency) so that the interference-temperature limit constraints at the primary receivers are met. This assumption is motivated by all the practical CR scenarios where primary terminals are oblivious to the presence of secondary users, thus behaving as if no secondary activity was present (also called *commons model*).

It is worth emphasizing that the use of the spatial domain can greatly improve the capabilities of cognitive users, as it allows them to transmit over the same frequency band and time slot without interfering. This however requires an opportunity identification phase, through a proper sensing mechanism: Secondary users need to reliably detect weak primary signals of possibly different type over a targeted region and wide frequency band in order to identify white-space halls. In particular, the use of spatial null constraints requires the identification of the primary receivers, a task that is much more demanding than the detection of primary transmitters, if the primary receivers are passive devices, as in TV network or downlink cellular systems. Examples of solutions to this problem have recently been proposed in [4], [33], and [34]. A recent overview of the challenges and possible solutions for the design of collaborative wideband sensing in CR networks can be found in [35] and [36]. The study of sensing in CR networks goes beyond the scope of this paper. Hereafter, we thus assume perfect sensing from the secondary users.

B. Game Theoretical Formulation

We formulate the system design as a strategic noncooperative game in which the players are the secondary users that attempt to maximize their rate (4), subject to power and interference constraints. As a warm-up, we start considering power constraints (2) and null constraints (5), since they are suitable to model the white-space filling paradigm. More specifically, the problem formulated leads directly to what we call game $\mathcal{G}_{\text{null}}$. We also consider an alternative game formulation, $\mathcal{G}_{\infty}$, with improved convergence properties; however, it does not correspond to any physical scenario so it is a rather artificial formulation. The missing ingredient is provided by another game formulation, $\mathcal{G}_{\alpha}$, that does have a nice physical interpretation and asymptotically is equivalent to $\mathcal{G}_{\infty}$ (in the sense specified in Section IV); thus inheriting the improved convergence properties as well as the physical interpretation. After that, we consider more general opportunistic communications by allowing also soft shaping interference constraints (6) through the game $\mathcal{G}_{\text{soft}}$.

III. Game With Null Constraints

Given the rate functions in (4), the rate maximization game among the secondary users in the presence of the power constraints (2) and null constraints (5) is formally defined as

$$(\mathcal{G}_{\text{null}}): \begin{array}{ll} \underset{\mathbf{Q}_q \succeq \mathbf{0}}{\text{maximize}} & R_q(\mathbf{Q}_q, \mathbf{Q}_{-q}) \\ \text{subject to} & \mathsf{Tr}(\mathbf{Q}_q) \leq P_q, \quad \mathbf{U}_q^H \mathbf{Q}_q = \mathbf{0} \end{array} \tag{7}$$

for all $q \in \Omega$, where $\Omega \triangleq \{1, 2, \ldots, Q\}$ is the set of the players (the secondary users) and $R_q(\mathbf{Q}_q, \mathbf{Q}_{-q})$ is the payoff function of player q, defined in (4). Without the null constraints, the solution of each optimization problem in (7) would lead to the well-known MIMO waterfilling solution [29]. The presence of the null constraints modifies the problem and the solution for each user is not necessarily a waterfilling anymore. Nevertheless, we show now that introducing a proper projection matrix the solutions of (7) can still be efficiently computed via a waterfilling-like expression. To this end, we rewrite game $\mathcal{G}_{\text{null}}$ in a more convenient form as detailed next.

We need the following intermediate definitions. For any $q \in \Omega$, given $r_{H_{qq}} \triangleq \text{rank}(\mathbf{H}_{qq})$ and $r_{U_q} \triangleq \text{rank}(\mathbf{U}_q)$, with $r_{U_q} < n_{T_q}$ w.l.o.g., let $\mathbf{U}_q^{\perp} \in \mathbb{C}^{n_{T_q} \times r_{U_q^{\perp}}}$ be the semi-unitary matrix orthogonal to $\mathbf{U}_q$, with $r_{U_q^{\perp}} \triangleq \text{rank}\left(\mathbf{U}_{qq}^{\perp}\right) = n_{T_q} - r_{U_q}$ and $\mathbf{P}_{\mathcal{R}(\mathbf{U}_q^{\perp})} = \mathbf{U}_q^{\perp}\mathbf{U}_q^{\perp H}$ be the orthogonal projection onto $\mathcal{R}(\mathbf{U}_q^{\perp})$. We can then rewrite the null constraints in (7) as

$$\mathbf{Q}_q = \mathbf{P}_{\mathcal{R}(\mathbf{U}_q^{\perp})} \mathbf{Q}_q \mathbf{P}_{\mathcal{R}(\mathbf{U}_q^{\perp})}, \tag{8}$$

where we used that fact that each $\mathbf{Q}_q \succeq \mathbf{0}$ and thus also Hermitian [28, Sec. 7.1]. At this point, the problem can be further simplified by noting that the constraint $\mathbf{Q}_q = \mathbf{P}_{\mathcal{R}(\mathbf{U}_q^\perp)}\mathbf{Q}_q\mathbf{P}_{\mathcal{R}(\mathbf{U}_q^\perp)}$ in (7) is redundant, provided that the original channels $\mathbf{H}_{rq}$ are replaced by the modified channels $\mathbf{H}_{rq}\mathbf{P}_{\mathcal{R}(\mathbf{U}_r^\perp)}$. The final equivalent formulation then becomes: denoting $\mathbf{P}_q^\perp \triangleq \mathbf{P}_{\mathcal{R}(\mathbf{U}_q^\perp)}$,

$$\begin{array}{ll} \underset{\mathbf{Q}_q \succeq \mathbf{0}}{\text{maximize}} & \log\det\left(\mathbf{I} + \mathbf{P}_q^\perp \mathbf{H}_{qq}^H \tilde{\mathbf{R}}_{-q}^{-1}(\mathbf{Q}_{-q})\mathbf{H}_{qq}\mathbf{P}_q^\perp\mathbf{Q}_q\right) \\ \text{subject to} & \mathsf{Tr}(\mathbf{Q}_q) \leq P_q \end{array} \tag{9}$$

for all $q \in \Omega$, where

$$\tilde{\mathbf{R}}_{-q}(\mathbf{Q}_{-q}) \triangleq \mathbf{R}_{n_q} + \sum_{r\neq q} \mathbf{H}_{rq}\mathbf{P}_r^\perp\mathbf{Q}_r\mathbf{P}_r^\perp\mathbf{H}_{rq}^H \succ \mathbf{0}. \tag{10}$$

This is due to the fact that, for any user q, any optimal solution $\mathbf{Q}_q^\star$ in (9)—the MIMO waterfilling solution—will be orthogonal to the null space of $\mathbf{H}_{qq}\mathbf{P}_q^\perp$, whatever $\tilde{\mathbf{R}}_{-q}^{-1}(\mathbf{Q}_{-q}^\star)$ is [recall that $\tilde{\mathbf{R}}_{-q}^{-1}(\mathbf{Q}_{-q}) \succ \mathbf{0}$ for all feasible $\mathbf{Q}_{-q}$], implying $\mathcal{R}(\mathbf{Q}_q^\star) \subseteq \mathcal{R}\left(\mathbf{U}_q^\perp\right)$.

Building on the equivalence of (7) and (9), we can focus on the game in (9) and apply the framework proposed in [24] to fully characterize game $\mathcal{G}_{\text{null}}$, by deriving the structure of the Nash equilibria and the conditions guaranteeing both the uniqueness of the equilibrium and the global convergence of the proposed distributed algorithms. We address these issues in Sections III-A and B.

A. Nash Equilibria: Existence and Uniqueness

Before studying the game, we introduce the following notations and definitions. Let $\tilde{\Omega}$ denote the set of user' indexes associated to the rank deficient matrices $\mathbf{H}_{qq}\mathbf{U}_q^\perp$, defined as

$$\tilde{\Omega} \triangleq \left\{ q \in \Omega : r_{H_{qq}U_q^\perp} \triangleq \text{rank}\left(\mathbf{H}_{qq}\mathbf{U}_q^\perp\right) < \min\left(n_{R_q}, r_{U_q^\perp}\right)\right\} \tag{11}$$

and let $\mathbf{V}_{q,1} \in \mathbb{C}^{r_{U_q^\perp} \times r_{H_{qq}U_q^\perp}}$ be the semi-unitary matrices such that $\mathcal{R}(\mathbf{V}_{q,1}) = \mathcal{N}\left(\mathbf{H}_{qq}\mathbf{U}_q^\perp\right)^\perp$. Based on these definitions, we introduce: the lower dimensional (with respect to the original channels) modified channel matrices $\tilde{\mathbf{H}}_{rq} \in \mathbb{C}^{n_{R_q} \times r_{H_{rr}U_r^\perp}}$, defined as

$$\tilde{\mathbf{H}}_{rq} = \begin{cases} \mathbf{H}_{rq}\mathbf{U}_r^\perp\mathbf{V}_{r,1}, & \text{if } r \in \tilde{\Omega}, \\ \mathbf{H}_{rq}\mathbf{U}_r^\perp, & \text{otherwise}, \end{cases} \quad \forall\, r,\ q \in \Omega \tag{12}$$

and the nonnegative matrix $\mathbf{S}_{\text{null}} \in \mathbb{R}_+^{Q\times Q}$, defined as

$$[\mathbf{S}_{\text{null}}]_{qr} \triangleq \begin{cases} \widetilde{\text{innr}}_q \cdot \rho\left(\tilde{\mathbf{H}}_{rq}^H\tilde{\mathbf{H}}_{rq}\right)\rho\left(\tilde{\mathbf{H}}_{qq}^{\sharp H}\tilde{\mathbf{H}}_{qq}^\sharp\right), & \text{if } r \neq q \\ 0, & \text{otherwise} \end{cases} \tag{13}$$

with

$$\widetilde{\text{innr}}_q \triangleq \frac{\rho\left(\mathbf{R}_{n_q} + \sum_{r\neq q} P_r \tilde{\mathbf{H}}_{rq}\tilde{\mathbf{H}}_{rq}^H\right)}{\lambda_{\min}\left(\mathbf{R}_{n_q}\right)} \geq 1. \tag{14}$$

Matrix $\mathbf{S}_{\text{null}}$ in (13) is instrumental to obtain sufficient conditions guaranteeing the uniqueness of the NE as well as the convergence of the proposed algorithms. Finally, to write the Nash equilibria of game $\mathcal{G}_{\text{null}}$ in a convenient form, we introduce for any $q \in \Omega$ and given $n_q \in \{1, 2, \ldots, n_{T_q}\}$, the MIMO waterfilling operator $\mathsf{WF}_q : \mathbb{S}_+^{n_q\times n_q} \ni \mathbf{X} \to \mathbb{S}_+^{n_q\times n_q}$, defined as

$$\mathsf{WF}_q(\mathbf{X}) \triangleq \mathbf{U}_X\left(\mu_{q,X}\mathbf{I}_{r_X} - \mathbf{D}_X^{-1}\right)^+\mathbf{U}_X^H \tag{15}$$

where $\mathbf{U}_X \in \mathbb{C}^{n_q\times r_X}$ and $\mathbf{D}_X \in \mathbb{R}_{++}^{r_X\times r_X}$ are the (semi-)unitary matrix of the eigenvectors and the diagonal matrix of the $r_X \triangleq \text{rank}(\mathbf{X}) \leq n_q$ (positive) eigenvalues of $\mathbf{X}$, respectively, and $\mu_{q,X} > 0$ is the water-level chosen to satisfy $\text{Tr}\left\{(\mu_{q,X}\mathbf{I}_{r_X} - \mathbf{D}_X^{-1})^+\right\} = P_q$.

Using the above definitions, the full characterization of the Nash equilibria of $\mathcal{G}_{\text{null}}$ is stated in the following theorem.

Theorem 1 (Existence and Uniqueness of the NE of $\mathcal{G}_{\text{null}}$): Consider the game $\mathcal{G}_{\text{null}}$ in (7) and suppose w.l.o.g. $r_{U_q} < n_{T_q}$, for all $q \in \Omega$. Then, the following hold:

a) there always exists a NE, for any set of channel matrices, power constraints for the users, and null shaping constraints;
b) all the Nash equilibria are the solutions to the following set of nonlinear matrix-value fixed-point equations:

$$\mathbf{Q}_q^\star = \mathbf{U}_q^\perp\mathsf{WF}_q\left(\mathbf{U}_q^{\perp H}\mathbf{H}_{qq}^H\mathbf{R}_{-q}^{-1}(\mathbf{Q}_{-q}^\star)\mathbf{H}_{qq}\mathbf{U}_q^\perp\right)\mathbf{U}_q^{\perp H}, \quad \forall\, q \in \Omega; \tag{16}$$

c) the NE is unique if[1]

$$\rho(\mathbf{S}_{\text{null}}) < 1. \tag{C1}$$

Proof: See Appendix A. ■

Corollary 2: A sufficient condition for (C1) in Theorem 1 is given by one of the two following set of conditions.

Low received MUI:

$$\frac{1}{w_q}\sum_{r\neq q}\widetilde{\text{innr}}_q \cdot \rho\left(\tilde{\mathbf{H}}_{rq}^H\tilde{\mathbf{H}}_{rq}\right)\rho\left(\tilde{\mathbf{H}}_{qq}^{\sharp H}\tilde{\mathbf{H}}_{qq}^\sharp\right)w_r < 1, \quad \forall\, q \in \Omega. \tag{C2}$$

Low generated MUI:

$$\frac{1}{w_r}\sum_{q\neq r}\widetilde{\text{innr}}_q \cdot \rho\left(\tilde{\mathbf{H}}_{rq}^H\tilde{\mathbf{H}}_{rq}\right)\rho\left(\tilde{\mathbf{H}}_{qq}^{\sharp H}\tilde{\mathbf{H}}_{qq}^\sharp\right)w_q < 1, \quad \forall\, r \in \Omega \tag{C3}$$

where $\mathbf{w} \triangleq [w_1, \ldots, w_Q]^T$ is any positive vector. □

[1]Note that a milder (but less easy to check) uniqueness condition than (C1) can be obtained by applying [24, Theorem 9] or [24, Corollary 10] to game $\mathcal{G}_{\text{null}}$. We omit the details because of the space limitation.

Remark 1—Structure of the Nash Equilibria: The structure of the Nash equilibria as given in (16) shows that the null constraints in the transmissions of secondary users can be handled without affecting the computational complexity: The optimal transmission strategy for each user q can be efficiently computed via a MIMO waterfilling solution, provided that the original channel matrix $\mathbf{H}_{qq}$ is replaced by $\mathbf{H}_{qq}\mathbf{U}_q^{\perp}$. Observe that the optimal structure of the covariance matrix in (16) has an intuitive interpretation: To guarantee that each user q does not transmit over a given subspace (spanned by the columns of $\mathbf{U}_q$), *whatever* the strategies of the other users are, while maximizing his information rate, it is enough to induce in the original channel matrix $\mathbf{H}_{qq}$ a null space that coincides with the subspace where the transmission is not allowed. This is precisely what is done in the payoff functions in (9) by replacing $\mathbf{H}_{qq}$ with $\mathbf{H}_{qq}\mathbf{P}_{\mathcal{R}(\mathbf{U}_q^{\perp})}$.

Remark 2—Physical Interpretation of Uniqueness Conditions: Conditions (C1)–(C3) state that the uniqueness of the NE is ensured if the interference among secondary users is sufficiently small. The same conditions will be shown to be sufficient also for the convergence of the distributed algorithms proposed in Section III-B. As expected, what affects the uniqueness of the equilibrium and the convergence of the algorithms is only the MUI generated by secondary users in the subspaces orthogonal to $\mathcal{R}(\mathbf{U}_q)$'s [see (C1)–(C3)], i.e., the subspaces where secondary users are allowed to transmit [note that all the Nash equilibria $\{\mathbf{Q}_q^{\star}\}_{q\in\Omega}$ satisfy $\mathcal{R}(\mathbf{Q}_q^{\star}) \subseteq \mathcal{R}\left(\mathbf{U}_q^{\perp}\right)$, for all $q \in \Omega$]. The importance of conditions (C1)–(C3) is that they quantify how small the interference must be to guarantee that the equilibrium is indeed unique. For example, condition (C2) can be interpreted as a constraint on the maximum amount of interference that each receiver can tolerate, whereas (C3) introduces an upper bound on the maximum level of interference that each transmitter of the secondary users is allowed to generate. Note that one can also obtain sufficient uniqueness/convergence conditions that are independent of the null constraints $\{\mathbf{U}_q\}_{q\in\Omega}$: it is enough to replace in (C1)–(C3) the modified channel matrices $\tilde{\mathbf{H}}_{rq}$ with the original ones $\mathbf{H}_{rq}$.[2] This means that if the NE is unique in a game without null constraints, then it is also unique with null constraint, which is not a trivial statement.

B. Distributed Algorithms

In this section we focus on distributed algorithms that converge to the NE of game $\mathcal{G}_{\text{null}}$. We consider totally asynchronous distributed algorithms, meaning that in the updating procedure some users are allowed to change their strategy more frequently than the others, and they might even perform these updates using *outdated* information on the interference caused by the others. To provide a formal description of the proposed asynchronous MIMO IWFA, we briefly recall some intermediate definitions, as given in [22].

We assume, without loss of generality, that the set of times at which one or more users update their strategies is the discrete set $\mathcal{T} = \mathbb{N}_+ = \{0,1,2,\ldots\}$. Let $\mathbf{Q}_q^{(n)}$ denote the covariance matrix of the vector signal transmitted by user q at the nth iteration, and let $\mathcal{T}_q \subseteq \mathcal{T}$ denote the set of times n at which $\mathbf{Q}_q^{(n)}$ is updated (thus, at time $n \notin \mathcal{T}_q$, $\mathbf{Q}_q^{(n)}$ is left unchanged). Let $\tau_r^q(n)$ denote the most recent time at which the interference from user r is perceived by user q at the nth iteration (observe that $\tau_r^q(n)$ satisfies $0 \leq \tau_r^q(n) \leq n$). Hence, if user q updates his own covariance matrix at the nth iteration, then he chooses his optimal $\mathbf{Q}_q^{(n)}$, according to his best-response (cf. Theorem 1)

$$\mathsf{T}_q(\mathbf{Q}_{-q}) \triangleq \mathbf{U}_q^{\perp}\mathsf{WF}_q\left(\mathbf{U}_q^{\perp H}\mathbf{H}_{qq}^H\mathbf{R}_{-q}^{-1}(\mathbf{Q}_{-q})\mathbf{H}_{qq}\mathbf{U}_q^{\perp}\right)\mathbf{U}_q^{\perp H}, \tag{17}$$

and using the interference level caused by

$$\mathbf{Q}_{-q}^{(\boldsymbol{\tau}^q(n))} \triangleq \left(\mathbf{Q}_1^{\left(\tau_1^q(n)\right)},\ldots,\mathbf{Q}_{q-1}^{\left(\tau_{q-1}^q(n)\right)},\right. \\ \left.\mathbf{Q}_{q+1}^{\left(\tau_{q+1}^q(n)\right)},\ldots,\mathbf{Q}_Q^{\left(\tau_Q^q(n)\right)}\right) \tag{18}$$

where in (17) the MIMO waterfilling mapping $\mathsf{WF}_q(\cdot)$ is defined in (15). Some standard conditions in asynchronous convergence theory that are fulfilled in any practical implementation need to be satisfied by the schedule $\{\tau_r^q(n)\}$ and $\{\mathcal{T}_q\}$; we refer to [19], [22] for the details. Through the whole paper we assume that these conditions are satisfied and call such an updating schedule as feasible.

Using the above notation, the asynchronous MIMO IWFA is formally described in Algorithm 1, where the best-response $\mathsf{T}_q(\cdot)$ of each user is defined in (17). Sufficient conditions guaranteeing the global convergence of the algorithm to the unique NE of $\mathcal{G}_{\text{nat}}$ are given in Theorem 3, whose proof follows from Theorem 1 and [24, Theorem 12] (we omit the proof for the sake of the paper length).

Algorithm 1: MIMO Asynchronous IWFA With Null Constraints

Data: any feasible $\mathbf{Q}_q^{(0)}$, $\forall\, q \in \Omega$;
1: Set $n = 0$;
2: `repeat`
3:

$$\mathbf{Q}_q^{(n+1)} = \begin{cases} \mathsf{T}_q\left(\mathbf{Q}_{-q}^{(\boldsymbol{\tau}^q(n))}\right), & \text{if } n \in \mathcal{T}_q, \\ \mathbf{Q}_q^{(n)}, & \text{otherwise;} \end{cases} \quad \forall\, q \in \Omega \tag{19}$$

4: `until the prescribed convergence criterion is satisfied`

Theorem 3: Suppose that condition (C1) of Theorem 1 is satisfied. Then, any sequence $\{\mathbf{Q}^{(n)}\}_{n=1}^{\infty}$ generated by the asynchronous MIMO IWFA described in Algorithm 1 converges to the unique solution to (16), for any feasible updating schedule of the users. □

Remark 3—Main Properties of the Algorithm: Algorithm 1 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 (i.e., the parameters $\{\tau_r^q(n)\}$ and $\{\mathcal{T}_q\}$). Two well-known special cases are the *sequential* and the *simultaneous* schemes, where the users update their own strategies *sequentially* and *simultaneously*, respectively. The important result stated in Theorem 3 is that all the algorithms resulting as special cases of the asynchronous MIMO IWFA are guaranteed

[2] The proof is based on the Poincaré separation theorem [28, Cor. 4.3.16] and $\mathbf{0} \leq \mathbf{A} \leq \mathbf{B} \implies \rho(\mathbf{A}) \leq \rho(\mathbf{B})$ [37, Cor. 2.2.22]). We omit the details because of the space limitation.

to reach the unique NE of the game, under the same set of convergence conditions, since condition (C1) does not depend on the particular choice of $\{\mathcal{T}_q\}$ and $\{\tau_r^q(n)\}$. Moreover all the algorithms obtained from Algorithm 1 have the following desired properties:

— *Low complexity and distributed implementation*: Even in the presence of null constraints, the best-response $\mathsf{T}_q(\cdot)$ of each user q can be efficiently and locally computed using a MIMO waterfilling based solution, provided that each channel $\mathbf{H}_{qq}$ is replaced by the channel $\mathbf{H}_{qq}\mathbf{U}_q^{\perp}$. Thus, Algorithm 1 can be implemented in a distributed way, since each user only needs to measure the overall interference-plus-noise covariance matrix $\mathbf{R}_{-q}$ and waterfill over $\mathbf{U}_q^{\perp H}\mathbf{H}_{qq}^H\mathbf{R}_{-q}^{-1}(\mathbf{Q}_{-q})\mathbf{H}_{qq}\mathbf{U}_q^{\perp}$.

— *Robustness*: According to Theorem 3, Algorithm 1 is robust against missing or outdated updates of secondary users. This feature strongly relaxes the constraints on the synchronization of the secondary users' updates with respect to those imposed, for example, by the simultaneous or sequential updating schemes.

— *Fast convergence behavior*: We have experienced that, like the classical simultaneous MIMO IWFA (without null constraints) [22], [24], the simultaneous version of the proposed algorithm converges in a very few iterations, even in networks with many active secondary users. The sequential IWFA is slower than the simultaneous IWFA, especially if the number of active secondary users is large, since each user is forced to wait for all the users scheduled in advance, before updating his own covariance matrix. This intuition is formalized in [18], [19], where the authors provided the expression of the asymptotic convergent factor of both the sequential and simultaneous IWFAs. The same results can also be obtained in the presence of null constraints (we omit the details because of the space limitation).

— *Control of the radiated interference*: Thanks to the game theoretical formulation including null constraints, the proposed asynchronous IWFA does not suffer of the main drawback of the classical sequential IWFA [15], i.e., the violation of the interference-temperature limits [3].

IV. Game With Null Constraints Via Virtual Noise Shaping

We have seen how to deal efficiently with null constraints in the rate maximization game. Under condition (C1) [or (C2)–(C3)], the NE is unique and Algorithm 1 asymptotically converges to this solution. However these constraints depend, among all, on the interference generated by the primary users (through the interference-plus-noise to noise ratios $\widetilde{\text{innr}}_q$'s), and thus it may not be satisfied for some interference profile, which is an undesired result. In such a case, the NE might not be unique and there is no guarantee that the proposed algorithms converge to an equilibrium. To overcome this issue, we propose here an alternative approach to impose null constraints (5) on the transmissions of secondary users based on the introduction of *virtual interferers*. This leads to a new game with more relaxed uniqueness and convergence conditions. The solutions of this new game are in general different to the Nash equilibria of $\mathcal{G}_{\text{null}}$, but the two games are numerically shown to have almost the same performance in terms of sum-rate.

The idea behind this alternative approach can be easily understood if one considers the transmission over SISO frequency-selective channels, where all the channel matrices have the same eigenvectors (the DFT vectors): to avoid the use of a given subchannel, it is sufficient to introduce a "virtual" noise with sufficiently high power over that subchannel. The same idea cannot be directly applied to the MIMO case, as arbitrary MIMO channel matrices have different right/left singular vectors from each other. Nevertheless, we show how to bypass this difficulty to design the covariance matrix of the virtual noise (to be added to the noise covariance matrix of each secondary receiver), so that all the Nash equilibria of the game satisfy the null constraints along the specified directions. For the sake of simplicity, we assume here nonsingular square channel matrices $\mathbf{H}_{qq}$, for all $q \in \Omega$. Similar results can be obtained for the rectangular case.

Let us consider the following strategic noncooperative game:

$$(\mathcal{G}_\alpha): \quad \begin{array}{ll} \underset{\mathbf{Q}_q \succeq \mathbf{0}}{\text{maximize}} & \log\det\left(\mathbf{I} + \mathbf{H}_{qq}^H \mathbf{R}_{-q,\alpha}^{-1}(\mathbf{Q}_{-q})\mathbf{H}_{qq}\mathbf{Q}_q\right) \\ \text{subject to} & \mathsf{Tr}\left(\mathbf{Q}_q\right) \leq P_q \end{array} \tag{20}$$

for all $q \in \Omega$, where

$$\mathbf{R}_{-q,\alpha}(\mathbf{Q}_{-q}) \triangleq \mathbf{R}_{n_q} + \sum_{r \neq q} \mathbf{H}_{rq}\mathbf{Q}_r\mathbf{H}_{rq}^H + \alpha\hat{\mathbf{U}}_q\hat{\mathbf{U}}_q^H \succ \mathbf{0} \tag{21}$$

denotes the MUI-plus-noise covariance matrix observed by secondary user q, plus the covariance matrix $\alpha\hat{\mathbf{U}}_q\hat{\mathbf{U}}_q^H$ of the virtual interference along $\mathcal{R}(\hat{\mathbf{U}}_q)$, where $\hat{\mathbf{U}}_q \in \mathbb{C}^{n_{R_q} \times r_{\hat{U}_q}}$ is a (strictly) tall matrix assumed to be full column-rank with $r_{\hat{U}_q} \triangleq \text{rank}(\hat{\mathbf{U}}_q) < r_{H_{qq}} (= n_{T_q} = n_{R_q})$ w.l.o.g., and α is a positive constant.

Our interest is on deriving the asymptotic properties of the solutions of $\mathcal{G}_\alpha$, as $\alpha \to +\infty$, and the structure of $\hat{\mathbf{U}}_q$'s making the null constraints (5) satisfied. To this end, we introduce the following auxiliary game $\mathcal{G}_\infty$:

$$(\mathcal{G}_\infty): \quad \begin{array}{ll} \underset{\mathbf{Q}_q \succeq \mathbf{0}}{\text{maximize}} & \log\det\left(\mathbf{I} + \hat{\mathbf{H}}_{qq}^H \hat{\mathbf{R}}_{-q}^{-1}(\mathbf{Q}_{-q})\hat{\mathbf{H}}_{qq}\mathbf{Q}_q\right) \\ \text{subject to} & \mathsf{Tr}\left(\mathbf{Q}_q\right) \leq P_q \end{array} \tag{22}$$

for all $q \in \Omega$, where

$$\hat{\mathbf{R}}_{-q}(\mathbf{Q}_{-q}) \triangleq \hat{\mathbf{U}}_q^{\perp H}\mathbf{R}_{n_q}\hat{\mathbf{U}}_q^{\perp} + \sum_{r \neq q} \hat{\mathbf{H}}_{rq}\mathbf{Q}_r\hat{\mathbf{H}}_{rq}^H, \tag{23}$$

the modified (strictly fat) channel matrices $\hat{\mathbf{H}}_{rq} \in \mathbb{C}^{r_{\hat{U}_q^{\perp}} \times n_{T_r}}$ are defined as

$$\hat{\mathbf{H}}_{rq} = \hat{\mathbf{U}}_q^{\perp H}\mathbf{H}_{rq} \quad \forall\, r,\, q \in \Omega, \tag{24}$$

and $\hat{\mathbf{U}}_q^{\perp} \in \mathbb{C}^{n_{R_q} \times r_{\hat{U}_q^{\perp}}}$ is the (strictly) tall full column-rank matrix $\hat{\mathbf{U}}_q^{\perp} \in \mathbb{C}^{n_{R_q} \times r_{\hat{U}_q^{\perp}}}$ orthogonal to $\hat{\mathbf{U}}_q$ [i.e., $\mathcal{R}\left(\hat{\mathbf{U}}_q^{\perp}\right) = \mathcal{R}(\hat{\mathbf{U}}_q)^{\perp}$], with $r_{\hat{U}_q^{\perp}} = n_{R_q} - r_{\hat{U}_q} = \text{rank}(\hat{\mathbf{U}}_q^{\perp})$. The relationship between the Nash equilibria of $\mathcal{G}_\alpha$ and $\mathcal{G}_\infty$ is studied next.

A. Nash Equilibria: Existence and Uniqueness

We first introduce the nonnegative matrices $\mathbf{S}, \mathbf{S}_{\infty} \in \mathbb{R}_{+}^{Q\times Q}$, defined as

$$[\mathbf{S}]_{qr} \triangleq \begin{cases} \rho\left(\mathbf{H}_{rq}^{H}\mathbf{H}_{qq}^{-H}\mathbf{H}_{qq}^{-1}\mathbf{H}_{rq}\right), & \text{if } r \neq q \\ 0, & \text{otherwise} \end{cases} \tag{25}$$

$$[\mathbf{S}_{\infty}]_{qr} \triangleq \begin{cases} \rho\left(\hat{\mathbf{H}}_{rq}^{H}\hat{\mathbf{H}}_{qq}^{-H}\hat{\mathbf{H}}_{qq}^{-1}\hat{\mathbf{H}}_{rq}\right), & \text{if } r \neq q \\ 0, & \text{otherwise.} \end{cases} \tag{26}$$

Game $\mathcal{G}_{\alpha}$: Using the above definitions and the MIMO waterfilling best-response $\mathsf{WF}_q(\cdot)$ defined in (15), the full characterization of game $\mathcal{G}_{\alpha}$ is given in the following theorem, whose proof is based on the contraction properties of the multiuser waterfilling mapping as derived in [24, Theorem 5].

Theorem 4 (Existence and Uniqueness of the NE of $\mathcal{G}_{\alpha}$): Given the game $\mathcal{G}_{\alpha}$ in (20), the following hold:

a) there always exists a NE, for any set of channel matrices, transmit power of the users, virtual interference matrices $\hat{\mathbf{U}}_q\hat{\mathbf{U}}_q^H$'s, and $\alpha \geq 0$;

b) all the Nash equilibria are the solutions to the following set of nonlinear matrix-value fixed-point equations:

$$\mathbf{Q}_{q,\alpha}^{\star} = \mathsf{WF}_q\left(\mathbf{H}_{qq}^{H}\mathbf{R}_{-q,\alpha}^{-1}(\mathbf{Q}_{-q,\alpha}^{\star})\mathbf{H}_{qq}\right), \quad \forall\, q \in \Omega; \tag{27}$$

c) the NE is unique if

$$\rho(\mathbf{S}) < 1. \tag{C4}$$

□

Remark 4—On the Properties of Game $\mathcal{G}_{\alpha}$: Game $\mathcal{G}_{\alpha}$ has some interesting properties, namely: i) The Nash equilibria depend on α and the virtual interference covariance matrices $\hat{\mathbf{U}}_q\hat{\mathbf{U}}_q^H$'s, whereas uniqueness condition (C4) *does not*; and ii) as desired, the uniqueness of the NE (and the convergence of the asynchronous algorithms) is not affected by the presence of the primary users. These features allow to choose α and $\hat{\mathbf{U}}_q\hat{\mathbf{U}}_q^H$'s so that the (unique) NE of the game satisfies the null constraints (5), while keeping the uniqueness property of the equilibrium independent of both $\hat{\mathbf{U}}_q\hat{\mathbf{U}}_q^H$'s and the interference level generated by the primary users. The optimal design of α and $\hat{\mathbf{U}}_q\hat{\mathbf{U}}_q^H$'s in $\mathcal{G}_{\alpha}$ passes through the properties of game $\mathcal{G}_{\infty}$, as detailed next.

Game $\mathcal{G}_{\infty}$: The properties of the game $\mathcal{G}_{\infty}$ are given in the following.

Theorem 5 (Existence and Uniqueness of the NE of $\mathcal{G}_{\infty}$): Consider the game $\mathcal{G}_{\infty}$ in (22) and suppose w.l.o.g. $r_{\hat{U}_q} < r_{H_{qq}}(= n_{R_q} = n_{T_q})$, for all $q \in \Omega$. Then, the following hold:

a) there always exists a NE, for any set of channel matrices, transmit power of the users, and virtual interference matrices $\hat{\mathbf{U}}_q$'s;

b) all the Nash equilibria are the solutions to the following set of nonlinear matrix-value fixed-point equations:

$$\mathbf{Q}_{q,\infty}^{\star} = \mathsf{WF}_q\left(\hat{\mathbf{H}}_{qq}^{H}\hat{\mathbf{R}}_{-q}^{-1}(\mathbf{Q}_{-q,\infty}^{\star})\hat{\mathbf{H}}_{qq}\right), \quad \forall\, q \in \Omega, \tag{28}$$

and satisfy $\mathcal{R}\left(\mathbf{Q}_{q,\infty}^{\star}\right) \perp \mathcal{R}\left(\mathbf{H}_{qq}^{-1}\hat{\mathbf{U}}_q\right)$, for all $q \in \Omega$;

c) the NE is unique if

$$\rho(\mathbf{S}_{\infty}) < 1. \tag{C5}$$

□

Remark 5—Null Constraints and Virtual Noise Directions: The orthogonality property $\mathcal{R}(\mathbf{Q}_{q,\infty}^{\star}) \perp \mathcal{R}\left(\mathbf{H}_{qq}^{-1}\hat{\mathbf{U}}_q\right)$ as stated in Theorem 5 provides, for each user q, the desired relationship between the directions of the virtual noise to be introduced in the noise covariance matrix of the user [see (23)]—the matrix $\hat{\mathbf{U}}_q$—and the real directions along with user q will not allocate any power, i.e., the matrix $\mathbf{U}_q$. It turns out that if user q is not allowed to allocate power along $\mathbf{U}_q$, it is sufficient to choose in (23) $\hat{\mathbf{U}}_q \triangleq \mathbf{H}_{qq}\mathbf{U}_q$. Exploring this choice, the structure of the Nash equilibria of game $\mathcal{G}_{\infty}$ can be further simplified and written as the solution of the following set of nonlinear matrix-value fixed-point equations:

$$\mathbf{Q}_{q,\infty}^{\star} = \mathbf{U}_q^{\perp}\mathsf{WF}_q\left(\left(\mathbf{U}_q^{\perp H}\mathbf{H}_{qq}^{-1}\mathbf{R}_{-q}(\mathbf{Q}_{-q,\infty}^{\star}) \times\mathbf{H}_{qq}^{-H}\mathbf{U}_q^{\perp}\right)^{-1}\right)\mathbf{U}_q^{\perp H} \tag{29}$$

for all $q \in \Omega$. Observe that, as desired, any NE $\mathbf{Q}_{q,\infty}^{\star}$ in (29) will be orthogonal to $\mathbf{U}_q$, since $\mathcal{R}(\mathbf{Q}_{q,\infty}^{\star}) \subseteq \mathcal{R}\left(\mathbf{U}_q^{\perp}\right)$, and thus satisfies the null constraints (5). Furthermore, milder conditions for the uniqueness of the solution to (29) can be obtained in such a case, by exploring the structure of the best-response in (29). We refer the interested reader to [38] for the details.

At this point, however, one may ask: What is the physical meaning of a solution to (29)? Does it still correspond to a waterfilling over a real MIMO channel and thus to the maximization of mutual information? The interpretation of game $\mathcal{G}_{\infty}$ and its solutions passes through game $\mathcal{G}_{\alpha}$: we indeed prove next that the solutions to (29) can be reached as Nash equilibria of game $\mathcal{G}_{\alpha}$ for sufficiently large $\alpha > 0$.

Relationship between game $\mathcal{G}_{\alpha}$ **and** $\mathcal{G}_{\infty}$: The asymptotic behavior of the Nash equilibria of $\mathcal{G}_{\alpha}$ as $\alpha \to \infty$, is given in the following.

Theorem 6 (Asymptotic Behavior of $\mathcal{G}_{\alpha}$): Given the games $\mathcal{G}_{\alpha}$ and $\mathcal{G}_{\infty}$, with $r_{\hat{U}_q} < r_{H_{qq}}(= n_{T_q} = n_{R_q})$ for all $q \in \Omega$, suppose that condition (C4) of Theorem 4 is satisfied. Then, the following hold:

a) $\mathcal{G}_{\alpha}$ and $\mathcal{G}_{\infty}$ admit a unique NE, denoted by $\mathbf{Q}_{\alpha}^{\star}$ and $\mathbf{Q}_{\infty}^{\star}$, respectively;

b) the two games are asymptotically equivalent, in the sense that

$$\lim_{\alpha\to+\infty} \mathbf{Q}_{\alpha}^{\star} = \mathbf{Q}_{\infty}^{\star}. \tag{30}$$

Proof: See Appendix B. ■

Invoking Theorem 6 we obtained the following desired property of game $\mathcal{G}_{\alpha}$: Under condition (C4) of Theorem 4, the (unique) NE of $\mathcal{G}_{\alpha}$ tends to satisfy the null constraints (5) for sufficiently large α, provided that the virtual interference matrices in (21) are chosen according to $\hat{\mathbf{U}}_q \triangleq \mathbf{H}_{qq}\mathbf{U}_q$ [see (29) and (30)], and still uniqueness/convergence condition (C4) does not depend on the interference generated by the primary users and the power budget of the secondary users.

B. Distributed Algorithms

To reach the Nash equilibria of game $\mathcal{G}_\alpha$ while satisfying the null constraints (5) (for sufficiently large α), one can use the asynchronous IWFA as given in Algorithm 1, where the best-response $\mathsf{T}_q(\mathbf{Q}_{-q})$ of each user q in (19) is replaced by the following:

$$\mathsf{T}_{q,\alpha}(\mathbf{Q}_{-q}) \triangleq \mathsf{WF}_q\left(\mathbf{H}_{qq}^H \mathbf{R}_{-q,\alpha}^{-1}(\mathbf{Q}_{-q,})\mathbf{H}_{qq}\right) \quad (31)$$

where the MIMO waterfilling operator WF_q is defined in (15), and $\hat{\mathbf{U}}_q$ in $\hat{\mathbf{R}}_{-q,\alpha}$ is $\hat{\mathbf{U}}_q = \mathbf{H}_{qq}\mathbf{U}_q$. Observe that such an algorithm has the same nice properties of the algorithm proposed to reach the Nash equilibria of game $\mathcal{G}_{\text{null}}$ in (7) (see Remark 3 in Section III-B). The convergence of the algorithm is guaranteed under the following conditions.

Theorem 7: Given the games $\mathcal{G}_\alpha$ and $\mathcal{G}_\infty$, with $r_{\hat{U}_q} < r_{H_{qq}} (= n_{T_q} = n_{R_q})$ for all $q \in \Omega$, suppose that condition (C4) of Theorem 4 is satisfied. Then, the following hold:

a) any sequence $\left\{\mathbf{Q}_\alpha^{(n)}\right\}_{n=0}^\infty$ generated by the asynchronous MIMO IWFA described in Algorithm 1 and based on the mapping in (31) converges *uniformly* with respect to $\alpha \in \mathbb{R}_+$ to the *unique* NE of game $\mathcal{G}_\alpha$, for any feasible updating schedule of the users;

b) the sequence $\left\{\mathbf{Q}_\alpha^{(n)}\right\}_{n=0}^\infty$ satisfies:

$$\lim_{n\to+\infty}\lim_{\alpha\to+\infty}\mathbf{Q}_\alpha^{(n)} = \lim_{\alpha\to+\infty}\lim_{n\to+\infty}\mathbf{Q}_\alpha^{(n)} = \mathbf{Q}_\infty^\star \quad (32)$$

where $\mathbf{Q}_\infty^\star$ is the (unique) NE of game $\mathcal{G}_\infty$.

Proof: See Appendix B. ■

Theorem 7 makes the game $\mathcal{G}_\alpha$ useful in practice: Algorithm 1 based on mapping in (31) and sufficiently large α globally converges to the unique NE of $\mathcal{G}_\alpha$ [see the LHS of (32)] that satisfies the null constraints (5) with an arbitrary small error [see (29)].

In Section VI, we will compare the performance of games $\mathcal{G}_{\text{null}}$ and $\mathcal{G}_\alpha$ in terms of conditions guaranteeing the uniqueness of the NE and convergence of the proposed algorithms and achievable rates. Numerical results show that game $\mathcal{G}_\alpha$ is a viable alternative to $\mathcal{G}_{\text{null}}$.

V. Game With Soft and Null Constraints

The rate maximization in the presence of both null and *soft shaping* constraints can be formulated as follows:

$$(\mathcal{G}_{\text{soft}}): \begin{array}{ll} \underset{\mathbf{Q}_q\succeq\mathbf{0}}{\text{maximize}} & R_q(\mathbf{Q}_q,\mathbf{Q}_{-q}) \\ \text{subject to} & \mathsf{Tr}\left(\mathbf{G}_q^H\mathbf{Q}_q\mathbf{G}_q\right) \le P_q^{\text{ave}} \\ & \lambda_{\max}\left(\mathbf{G}_q^H\mathbf{Q}_q\mathbf{G}_q\right) \le P_q^{\text{peak}} \\ & \mathbf{U}_q^H\mathbf{Q}_q = \mathbf{0} \end{array} \quad (33)$$

for all $q \in \Omega$, where we have included both types of soft shaping constraints (6) as well as null constraints (5). The transmit power constraint (2) has been absorbed into the trace soft constraint for convenience. For this, it is necessary that each $r_{G_q} \triangleq \mathsf{rank}(\mathbf{G}_q) = n_{T_q}$ (implying $n_{G_q} \ge n_{T_q}$); otherwise there would be no power constraint along $\mathcal{N}\left(\mathbf{G}_q^H\right)$ [if user q is allowed to transmit along $\mathcal{N}\left(\mathbf{G}_q^H\right)$, i.e., $\mathcal{N}\left(\mathbf{G}_q^H\right) \cap \mathcal{R}(\mathbf{U}_q)^\perp \neq \{\emptyset\}$]. It is worth pointing out that, in practice, a transmit power constraint in (33) will be dominated by the trace shaping constraint, which motivates the absence in (33) of an explicit power constraint. More specifically, the power constraint (2) becomes redundant whenever $P_q^{\text{ave}} \le P_q\lambda_{\min}\left(\mathbf{G}_q\mathbf{G}_q^H\right)$. On the other hand, if $P_q^{\text{ave}} \ge P_q\lambda_{\max}\left(\mathbf{G}_q\mathbf{G}_q^H\right)$, then the soft shaping constraints (6) can be removed without loss of optimality, and game $\mathcal{G}_{\text{soft}}$ cast in the form of game $\mathcal{G}_{\text{null}}$. In the following, we then focus on the former case only.

A. Nash Equilibria: Existence and Uniqueness

Before studying game $\mathcal{G}_{\text{soft}}$, we introduce the following intermediate definitions. For any $q \in \Omega$, define the tall matrix $\overline{\mathbf{U}}_q \in \mathbb{C}^{n_{G_q}\times r_{U_q}}$ as $\overline{\mathbf{U}}_q \triangleq \mathbf{G}_q^\sharp\mathbf{U}_q$ (recall that $n_{G_q} \ge n_{T_q} > r_{U_q}$), and the semi-unitary matrix $\overline{\mathbf{U}}_q^\perp \in \mathbb{C}^{n_{G_q}\times r_{\overline{U}_q^\perp}}$ orthogonal to $\overline{\mathbf{U}}_q$, with $r_{\overline{U}_q^\perp} = n_{G_q} - r_{U_q} = \text{rank}(\overline{\mathbf{U}}_q^\perp)$. Based on these definitions, we introduce the modified channels $\overline{\mathbf{H}}_{rq} \in \mathbb{C}^{n_{R_q}\times r_{\overline{U}_r^\perp}}$, defined as

$$\overline{\mathbf{H}}_{rq} = \mathbf{H}_{rq}\mathbf{G}_r^{\sharp H}\overline{\mathbf{U}}_r^\perp, \quad \forall r,q \in \Omega \quad (34)$$

and the nonnegative matrix $\mathbf{S}_{\text{soft}} \in \mathbb{R}_+^{Q\times Q}$:

$$[\mathbf{S}_{\text{soft}}]_{qr} \triangleq \begin{cases} \overline{\text{innr}}_q \cdot \rho\left(\overline{\mathbf{H}}_{rq}^H\overline{\mathbf{H}}_{rq}\right)\rho\left(\overline{\mathbf{H}}_{qq}^{\sharp H}\overline{\mathbf{H}}_{qq}^\sharp\right), & \text{if } r \neq q \\ 0, & \text{otherwise} \end{cases} \quad (35)$$

with

$$\overline{\text{innr}}_q \triangleq \frac{\rho\left(\mathbf{R}_{n_q} + \sum_{r\neq q} P_r\overline{\mathbf{H}}_{rq}\overline{\mathbf{H}}_{rq}^H\right)}{\lambda_{\min}\left(\mathbf{R}_{n_q}\right)} \ge 1. \quad (36)$$

Matrix $\mathbf{S}_{\text{soft}}$ is instrumental to obtain sufficient conditions for the uniqueness of the NE of $\mathcal{G}_{\text{soft}}$. Finally, we introduce for any $q \in \Omega$ and given $n_q \in \{1,2,\ldots,n_{T_q}\}$, the *modified* MIMO waterfilling operator $\overline{\mathsf{WF}}_q : \mathbb{S}_+^{n_q\times n_q} \ni \mathbf{X} \to \mathbb{S}_+^{n_q\times n_q}$, defined as

$$\overline{\mathsf{WF}}_q(\mathbf{X}) \triangleq \mathbf{U}_X\left[\mu_{q,X}\mathbf{I}_{r_X} - \mathbf{D}_X^{-1}\right]_0^{P_q^{\text{peak}}}\mathbf{U}_X^H \quad (37)$$

where $\mathbf{U}_X \in \mathbb{C}^{n_q\times r_X}$ and $\mathbf{D}_q \in \mathbb{R}_{++}^{r_X\times r_X}$ are defined as in (15) and $\mu_{q,X} > 0$ is the water-level chosen to satisfy $\text{Tr}\left\{\left[\mu_{q,X}\mathbf{I}_{r_X} - \mathbf{D}_X^{-1}\right]_0^{P_q^{\text{peak}}}\right\} = \min(P_q, r_X P_q^{\text{peak}})$ (see, e.g., [39] for practical algorithms to compute the waterlevel $\mu_{q,X}$ in (37)). Using the above definitions, we can now characterize the Nash equilibria of game $\mathcal{G}_{\text{soft}}$, as given next.

Theorem 8 (Existence and Structure of the NE of $\mathcal{G}_{\text{Soft}}$): Consider the game $\mathcal{G}_{\text{soft}}$ in (33) and suppose w.l.o.g. $r_{G_q} = n_{T_q}$, for all $q \in \Omega$ (all matrices $\mathbf{G}_q$ are full row-rank). Then, the following hold:

a) there always exists a NE, for any set of channel matrices and null/soft shaping constraints;

b) if, in addition, $r_{U_q} < r_{H_{qq}}$ and $\text{rank}(\overline{\mathbf{H}}_{qq}) = r_{\overline{U}_q^\perp}$ for all $q \in \Omega$ (all matrices $\overline{\mathbf{H}}_{qq}$ are full column-rank), all the Nash equilibria are the solutions to the following set of nonlinear matrix-value fixed-point equations:

$$\mathbf{Q}_q^\star = \mathbf{G}_q^{\sharp H}\overline{\mathbf{U}}_q^\perp\overline{\mathsf{WF}}_q\left(\overline{\mathbf{H}}_{qq}^H\mathbf{R}_{-q}^{-1}(\mathbf{Q}_{-q}^\star)\overline{\mathbf{H}}_{qq}\right)\overline{\mathbf{U}}_q^{\perp H}\mathbf{G}_q^\sharp,\ \forall q\in\Omega. \quad (38)$$

Proof: See Appendix C. ■

Remark 6—On the Nash Equilibria: The rank constraints in Theorem 8(b) have an interesting physical interpretation: if $\text{rank}(\overline{\mathbf{H}}_{qq}) = r_{\overline{U}_q^{\perp}}$, then the equivalent channel $\overline{\mathbf{H}}_{qq}$ of user q has an empty null space and the optimal solution of the qth convex optimization problem in (33) is unique (see Proposition 18 in Appendix C); otherwise there exist multiple (infinite) globally optimal solutions of the single-user optimization problem q that provide the same maximum rate but allocating power in the null space of $\overline{\mathbf{H}}_{qq}$. In the latter case, the solution given by the best-response in (38) is the one using the minimum transmit power to reach the maximum achievable rate in the single-user problem (termed as *minimum power best-response*). It turns out that, even in the case in which the aforementioned rank constraints are not satisfied, the solutions of the fixed-point equation in (38) are still Nash equilibria of the game $\mathcal{G}_{\text{soft}}$, the ones corresponding to the minimum power best-response of the players.

The structure of the Nash equilibria in (38) states that the optimal transmission strategy of each user leads to a diagonalizing transmission with a proper power allocation, after pre/post multiplication by matrix $\mathbf{G}_q^{\sharp H}\overline{\mathbf{U}}_q^{\perp}$, implying that, even in the presence of soft constraints, the optimal transmission strategy of each user q can be efficiently computed via a MIMO waterfilling-like solution. Finally, observe that the Nash equilibria in (38) satisfy the null constraints (6) since $\mathcal{R}(\overline{\mathbf{U}}_q^{\perp})^{\perp} = \mathcal{R}\left(\mathbf{G}_q^{\sharp}\mathbf{U}_q\right)$, implying $\mathbf{U}_q^H\mathbf{G}_q^{\sharp H}\overline{\mathbf{U}}_q^{\perp} = \mathbf{0}$ and thus $\mathcal{R}(\mathbf{Q}_q^{\star}) \perp \mathcal{R}(\mathbf{U}_q)$, for all $q \in \Omega$.

We provide now a more convenient expression for the Nash equilibria given in (38), that will be instrumental to derive conditions for the uniqueness of the equilibrium and the convergence of the distributed algorithms we propose in Section V-B. Introducing the convex sets $\overline{\mathcal{Q}}_q$ defined as

$$\overline{\mathcal{Q}}_q \triangleq \left\{\mathbf{X} \in \mathbb{S}_+^{n_{T_q}\times n_{T_q}} : \text{Tr}\{\mathbf{X}\} = \overline{P}_q^{\text{ave}}, \lambda_{\max}(\mathbf{X}) \le P_q^{\text{peak}}\right\} \tag{39}$$

where $\overline{P}_q^{\text{ave}} \triangleq \min\left(P_q^{\text{ave}}, r_{\overline{U}_q^{\perp}}P_q^{\text{peak}}\right)$, we have the following equivalent expression for the MIMO waterfilling solutions in (38).

Lemma 9 (NE as a Matrix Projection): The set of nonlinear matrix-value fixed-point equations in (38) can be equivalently rewritten as

$$\begin{aligned}\mathbf{Q}_q^{\star} &= \mathbf{G}_q^{\sharp H}\overline{\mathbf{U}}_q^{\perp}\\ &\times\left[-\left(\left(\overline{\mathbf{H}}_{qq}^H\mathbf{R}_{-q}^{-1}(\mathbf{Q}_{-q}^{\star})\overline{\mathbf{H}}_{qq}\right)^{\sharp} + c_q\mathbf{P}_{\mathcal{N}(\overline{\mathbf{H}}_{qq})}\right)\right]_{\overline{\mathcal{Q}}_q}\\ &\times\overline{\mathbf{U}}_q^{\perp H}\mathbf{G}_q^{\sharp}\end{aligned} \tag{40}$$

for all $q \in \Omega$ and sufficiently large values of $c_q > 0$. An admissible value of c_q is given in (89) in Appendix D.

Proof: See Appendix D. ■

Using Lemma 9, we can now apply [24, Theorem 4] to game $\mathcal{G}_{\text{soft}}$, and obtain sufficient conditions guaranteeing the uniqueness of the NE, as given next.

Theorem 10 (Uniqueness of the NE): The solution to (40) is unique if

$$\rho(\mathbf{S}_{\text{soft}}) < 1. \tag{C6}$$

□

B. Distributed Algorithms

Similarly to games $\mathcal{G}_{\text{null}}$ and $\mathcal{G}_{\alpha}$, the Nash equilibria of game $\mathcal{G}_{\text{soft}}$ can be reached using the asynchronous IWFA algorithm given in Algorithm 1, based on the following mapping:

$$\overline{\mathsf{T}}_q(\mathbf{Q}_{-q}) \triangleq \mathbf{G}_q^{\sharp H}\overline{\mathbf{U}}_q^{\perp}\overline{\mathsf{WF}}_q\left(\overline{\mathbf{H}}_{qq}^H\mathbf{R}_{-q}^{-1}(\mathbf{Q}_{-q})\overline{\mathbf{H}}_{qq}\right)\overline{\mathbf{U}}_q^{\perp H}\mathbf{G}_q^{\sharp} \tag{41}$$

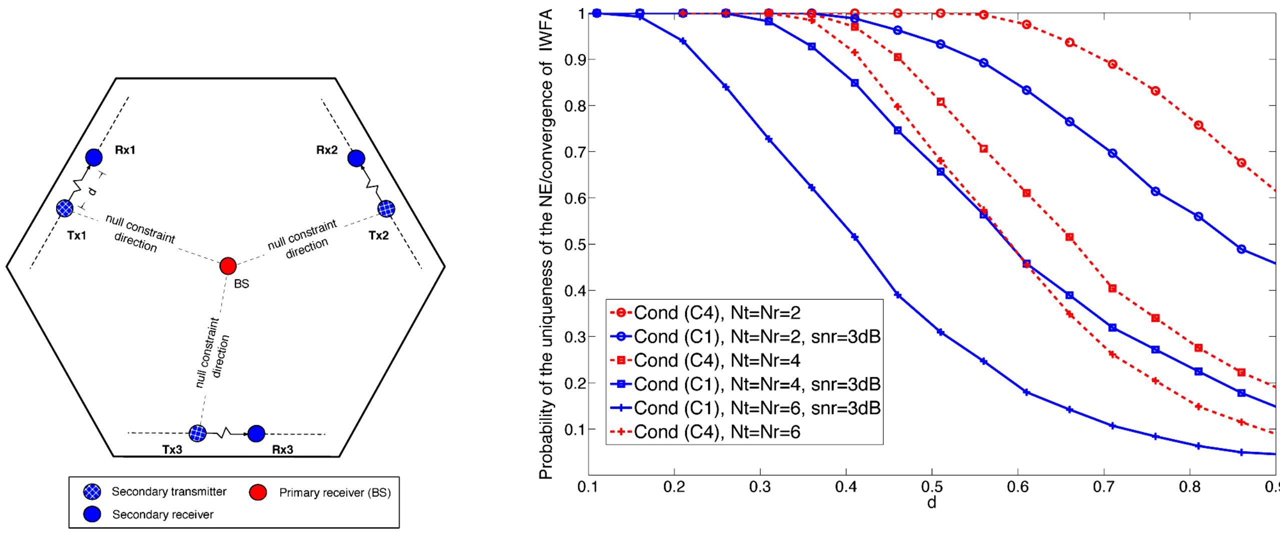
where the MIMO waterfilling operator $\overline{\mathsf{WF}}_q(\cdot)$ is defined in (37) and the modified channels $\overline{\mathbf{H}}_{qq}$'s are defined in (34). Observe that such an algorithm has the same nice properties of the algorithm proposed to reach the Nash equilibria of game $\mathcal{G}_{\text{null}}$ in (7) (see Remark 4 in Section III-B), such as low-complexity, distributed and asynchronous implementation, fast convergence behavior. Moreover, thanks to our game theoretical formulation including null and/or soft shaping constraints, the algorithm does not suffer of the main drawback of the classical sequential IWFA [15], [26], [40], i.e., the violation of the interference-temperature limits [3]. The convergence properties of the algorithm are given in the following.

Theorem 11: Suppose that condition (C6) of Theorem 10 is satisfied. Then, any sequence $\{\mathbf{Q}^{(n)}\}_{n=1}^{\infty}$ generated by the asynchronous MIMO IWFA described in Algorithm 1 and based on the mapping in (41) converges to the unique solution to (40), for any feasible updating schedule of the users. □

VI. Numerical Results

In this section we provide some numerical results validating our theoretical findings. More specifically, we compare the performance of games $\mathcal{G}_{\text{null}}$ and $\mathcal{G}_{\alpha}$ in terms of conditions guaranteeing the uniqueness of the NE and (average and outage) sum-rate. Finally, to show the effectiveness of the proposed game theoretical approach, we compare the performance achievable by our decentralized algorithms with those achievable by a centralized approach. The interesting result is that, in most practical cases of interest, the performance loss resulting from the use of a NE is small and more than acceptable, considering also that the centralized optimization is computationally much more expensive and cannot be implemented in a distributed way.

1) Numerical Example 1—Comparison of Uniqueness/Convergence Conditions: Since the uniqueness/convergence conditions of the proposed games depend on the channel matrices $\{\mathbf{H}_{rq}\}_{r,q\in\Omega}$, there is a nonzero probability that they will not be satisfied for a given channel realization drawn from a given probability space. To quantify the adequacy of our conditions, we tested them over a set of random channel matrices whose elements are generated as circularly symmetric complex Gaussian random variables with variance equal to the inverse of the square distance between the associated transmitter–receiver links (flat-fading channel model). As an example, we consider a hierarchical CR network as depicted in Fig. 1(a), composed of three secondary user MIMO links randomly distributed in the cell and one primary user [the base station (BS)], sharing the same band.



(a) (b)

Fig. 1. (a) CR MIMO system; (b) Probability of the uniqueness of the NE of games $\mathcal{G}_{\rm null}$ and $\mathcal{G}_\alpha$ [or $\mathcal{G}_\infty$ (Theorem 6)] and convergence of the asynchronous IWFA as a function of the normalized intra-pair distance $d \in (0,1)$; $Q=3, n_{T_q}=n_{R_q}=2$ (circle marks), $n_{T_q}=n_{R_q}=4$ (square marks), $n_{T_q}=n_{R_q}=6$ (cross marks), and $\mathrm{snr}_q \triangleq P_q/\sigma^2_{q,\rm tot} = 3$ dB $\forall\, q \in \Omega$. (a) Multicell cellular system. (b) Probability of the uniqueness of the NE and convergence of the asynchronous IWFA.

To preserve the QoS of the primary users, null constraints are imposed on the secondary users in the direction of the receiver of the primary user. In Fig. 1(b), we plot the probability that conditions (C1) (red line dashed curves) or (C4) (blue line solid curves) are satisfied versus the intra-pair distance $d \in (0,1)$ (normalized by the cell's side) between each secondary transmitter and the corresponding receiver (assumed for the simplicity of representation to be equal for all the secondary links), for different values of the transmit–receive antennas. For (C1) we set the value of the SNR at the receivers of the secondary users equal to $\mathrm{snr}_q \triangleq P_q/\sigma^2_{q,\rm tot} = 3$ dB for all $q \in \Omega$, where $\sigma^2_{q,\rm tot}$ is the variance of thermal noise plus the interference generated by the primary user over all the substreams.

As expected, the probability of the uniqueness of the NE of both games $\mathcal{G}_{\rm null}$ and $\mathcal{G}_\alpha$ and convergence of the IWFAs increases as each secondary transmitter approaches his receiver, corresponding to a decrease of the overall MUI. Moreover, condition (C1) is confirmed to be stronger than (C4) whatever the number of transmit–receive antennas, the intra-pair distance d, and the SNR value are, implying that game $\mathcal{G}_\alpha$ admits milder uniqueness/convergence conditions than those of the original game $\mathcal{G}_{\rm null}$.

2) Numerical Example 2—Performance of $\mathcal{G}_{\rm null}$ and $\mathcal{G}_\infty$: Even though the Nash equilibria of $\mathcal{G}_\infty$ (and $\mathcal{G}_\alpha$) may be chosen to satisfy the null constraints, they do not come up from the maximization of the information rate over the MIMO channel of each link. In fact, the best-response of each player in game $\mathcal{G}_\infty$ is a waterfilling solution but over the modified channel $\hat{\mathbf{H}}_{qq}$ [see (24)] rather than $\mathbf{H}_{qq}$, meaning that each player maximizes the information rate over a fictitious channel. In order to make the Nash equilibria of $\mathcal{G}_\infty$ (and $\mathcal{G}_\alpha$) useful in practice, we need to compare the performance of $\mathcal{G}_\infty$ and $\mathcal{G}_{\rm null}$. The numerical example in Fig. 2 shows that, surprisingly, the two games have almost the same performance. More specifically, in the picture we compare games $\mathcal{G}_{\rm null}$ and $\mathcal{G}_\infty$ (or $\mathcal{G}_\alpha$ for large α) in terms of sum-rate. All the Nash equilibria are computed using Algorithm 1 with mapping in (17) for game $\mathcal{G}_{\rm null}$ and (28) for game $\mathcal{G}_\infty$. In Fig. 2(a), we plot the average sum-rate at the (unique) NE of the games $\mathcal{G}_{\rm null}$ and $\mathcal{G}_\infty$ for the CR network depicted in Fig. 1(a) as a function of the intra-pair distance $d \in (0,1)$ among the links, for different numbers of transmit–receive antennas. In Fig. 2(b), we plot the outage sum-rate, for the same systems as in Fig. 2(a) and $d = 0.5$. For each secondary user, a null constraint in the direction of the receiver of the primary user is imposed. From the figures one infers that games $\mathcal{G}_{\rm null}$ and $\mathcal{G}_\infty$ have almost the same performance in terms of sum-rate at the NE; even if in the game $\mathcal{G}_\infty$, given the strategies of the others, each player does not maximize his own rate, as happens in the game $\mathcal{G}_{\rm null}$. This is due to the fact that the Nash equilibria of game $\mathcal{G}_{\rm null}$ are in general not Pareto efficient.

3) Numerical Example 3—How Good Is the NE?: We have seen that a NE might not be Pareto efficient. A formal analysis of the performance loss due to the use of the NE criterion with respect to the Pareto optimal solutions (computing, e.g., the so-called price-of-anarchy) is up-to-date a formidable open problem to solve, even in the absence of interference constraints (recall that solving the system-wide centralized optimization is an NP-hard problem), and it goes beyond the scope of this paper. However, from a practical point of view, it is still interesting to test the validity of the proposed game theoretical approach through numerical results. We thus compare via simulations the performance of the proposed games with those achievable using the state-of-the-art (suboptimal) centralized algorithms. Among several algorithms proposed in the literature to suboptimally solve the nonconvex sum-rate maximization problem for MIMO ad hoc systems [10], [41], [42], the Jacobi gradient projection based algorithm in [10] seems to be the one providing better performance. Hence, as a benchmark for our comparison,

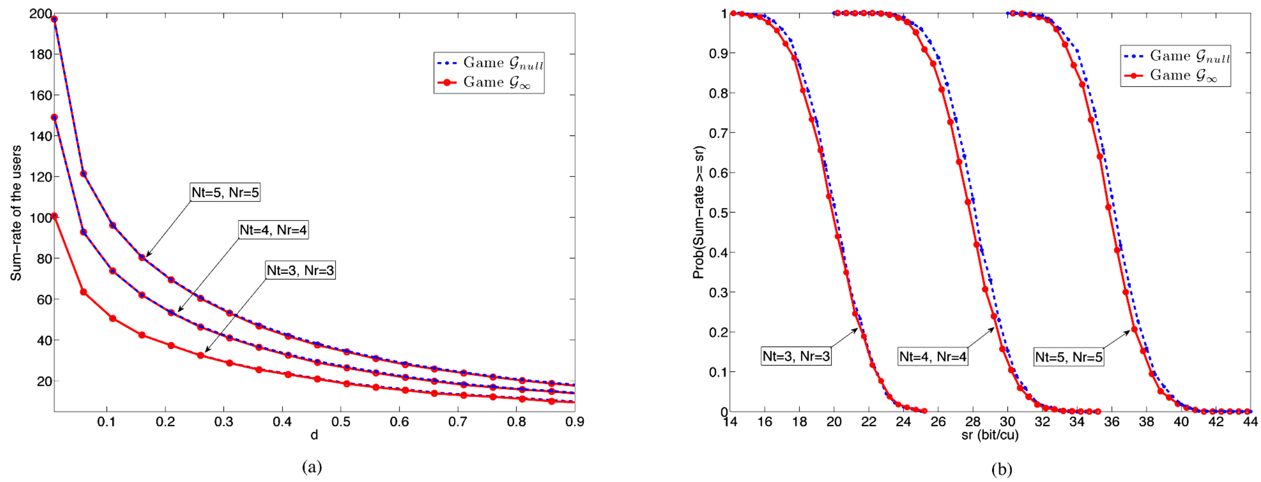


Fig. 2. Performance of games $\mathcal{G}_{\text{null}}$ and $\mathcal{G}_{\infty}$ in terms of Nash equilibria for the cognitive MIMO network given in Fig. 1(a); (a) Average sum-rate at the NE versus $d \in (0, 1)$; (b) Outage sum-rate for $d = 0.5$; $Q = 3$, $P_r = P_q$, and $\text{snr}_q = P_q/\sigma^2_{q,\text{tot}} = 3$ dB, for all $r, q \in \Omega$. (a) Average sum-rate at the NE of $\mathcal{G}_{\text{null}}$ and $\mathcal{G}_{\infty}$. (b) Outage sum-rate at the NE of $\mathcal{G}_{\text{null}}$ and $\mathcal{G}_{\infty}$.

we adopt such an algorithm, properly modified to include the interference constraints in the feasible set of the optimization problem. According to this algorithm, a central node having perfect knowledge of all the channel matrices and interference constraints computes a stationary point of the sum-rate function as limit point of the gradient projection Jacobi scheme (starting from any feasible point)

$$\mathbf{Q}_q^{(n+1)} = \left[\mathbf{Q}_q^{(n)} + \alpha^{(k)} \nabla_{Q_q} \mathsf{sr}\left(\mathbf{Q}^{(n)}\right)\right]_{\mathcal{X}_q} \tag{42}$$

$\forall\, q = 1, \ldots, Q$, where $\nabla_{Q_q} \mathsf{sr}(\mathbf{Q})$ denotes the gradient of the sum-rate function $\mathsf{sr}(\mathbf{Q}) \triangleq \sum_q R_q(\mathbf{Q})$ with respect to $\mathbf{Q}_q$,[3] $[\cdot]_{\mathcal{X}_q}$ is the orthogonal projection onto the feasible set $\mathcal{X}_q$ of each user q [including the interference constraints, see the feasible sets in (7) or (33)], and $\{\alpha^{(k)}\}$ is the sequence of step-sizes following some updating rule (see, e.g., [43, Sec. 2.3]). In our simulations we select the step-size sequence according to the Armijo rule [25, Sec. 2.3.1].

As an example, in Fig. 3, we plot the average sum-rate achievable by the NE of the game $\mathcal{G}_{\text{null}}$ and by the gradient projection algorithm in (42) versus the interference-to-noise ratio $\text{INR} \triangleq P_r/(d^2_{rq}\sigma^2_{q,\text{tot}})$ (assumed, for the sake of simplicity, to be equal for all the links) for the CR 3×3 MIMO system depicted in Fig. 1(a). We considered three different interference regimes among the secondary users (low, medium and high interference), corresponding to different values of the direct signal-to-noise ratio $\text{SNR} \triangleq P_q/(d^2_{qq}\sigma^2_{q,\text{tot}})$ (assumed to be equal for all the links). Since the maximization of the sum-rate is not a convex problem, every limit point of the gradient projection algorithm in (42) is guaranteed only to be a stationary point of the sum-rate and may depend on the initialization of the algorithm. We thus simulated the gradient projection algorithm trying several choices for the initial conditions, namely scaled identity matrix, random feasible covariance matrix, and the covariance matrix corresponding to the NE of $\mathcal{G}_{\text{null}}$. For the latter two cases, we experienced almost the same results (we thus omitted in the plot the curves associated to the NE initialization).

[3]Note that there is no one unique way to consistently define a complex gradient which applies to (necessarily non-complex-analytic) real-valued functions of a complex variable, and authors do not uniformly adhere to the same definition. Here, we use the definition of complex gradient as given in [10].

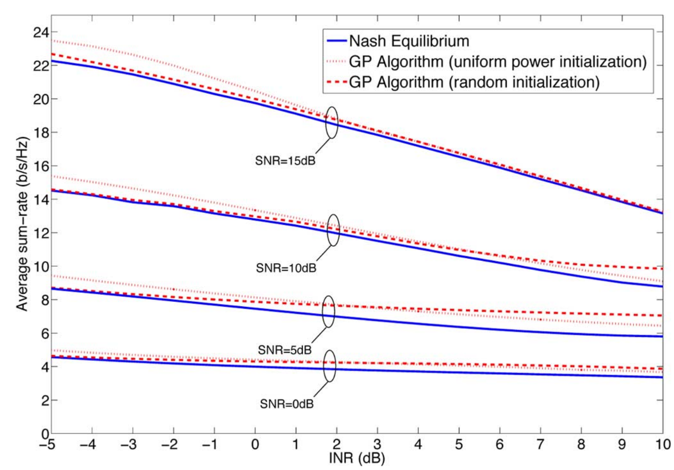


Fig. 3. Average sum-rate versus the interference-to-noise ratio $\text{INR} = P_r/(d^2_{rq}\sigma^2_{q,\text{tot}})$, achievable at the NE of the game $\mathcal{G}_{\text{null}}$ (blue solid line curves) and by the gradient projection algorithm starting from a random feasible point (dashed red line curves) and from the scaled identity matrix (dot red line curves), for the cognitive MIMO network given in Fig. 1(a); $Q = 3$, $N_t = N_r = 3$, $\text{SNR} = P_q/(d^2_{qq}\sigma^2_{q,\text{tot}}) = 0, 5, 10, 15$ dB for all $r, q \in \Omega$.

Fig. 3 suggests the following comments. The centralized approach based on the gradient projection algorithm exhibits better performance than the NE solution. But, as expected, the rate loss of the NE decreases as the interference level among the secondary users decreases, i.e., the INR decreases (keeping the SNR fixed) or the SNR increases (keeping the INR fixed). This is not surprising as, in the case in which the interference is sufficiently low, the interaction (interference) among users becomes more and more negligible. Interestingly, the rate loss experienced in the simulated CR system is more than acceptable for low/medium interference regimes and becomes no more large than 25% of the centralized solution in the worst case interference scenario. Finally, it is important to remark that the better performance of the gradient projection algorithm comes at a high price, namely i) to be implemented, the gradient projection algorithm requires a centralized node with full knowledge of all the channel matrices and interference constraints that computes the optimal covariance matrices of the secondary users and broadcasts them to the corresponding user, which comes at the cost of extra signaling among the secondary users and breaks the noncooperative nature of the CR network; ii) each iteration of the gradient projection algorithm needs the computation of the orthogonal projection onto the feasible set of each secondary user [see (42)], which in the presence of soft and null constraints is computationally much more demanding than the computation of a MIMO waterfilling-like solution [see (17)]; and iii) the final point of the gradient projection algorithm and thus the performance of the algorithm depend on the initialization. The proposed NE solutions instead are suitable to be computed using low-complexity distributed (possibly) asynchronous algorithms that do require no exchange of information among the CR users and provide a good trade-off between performance and implementation issues imposed by the CR systems.

VII. Conclusion

In this paper, we have proposed a novel game theoretical formulation to solve one of the challenging and unsolved resource allocation problems in CR systems: How to allow in a decentralized way concurrent communication over MIMO channels among secondary users, under different constraints imposed to the secondary users on the interference induced to the primary users. The null/soft shaping constraints have been in fact used in a very broad sense, meaning that the projection of the transmitted signal along prescribed subspaces should be null (null constraints) or below a given threshold (soft constraints). In particular, for null shaping constraints we have considered the naturally resulting game $\mathcal{G}_{\text{null}}$ as well as a modified game $\mathcal{G}_{\infty}$ (or $\mathcal{G}_{\alpha}$ for large alpha) with similar performance and more relaxed requirements; for soft shaping constraints plus null constraints, we have studied the game $\mathcal{G}_{\text{soft}}$. For all the games we provided sufficient conditions guaranteeing the uniqueness of the NE and proposed distributed totally asynchronous algorithms that were proved to globally converge under the same conditions guaranteeing the uniqueness of the NE. All the algorithms overcome the main drawback of classical MIMO IWFA—the violation of the temperature-interference limits—and they have the desired features required for CR applications, such as low-complexity, distributed implementation, robustness against missing or outdated updates of the users, and fast convergence behavior.

Appendix A
Proof of Theorem 1

The existence of a NE of game $\mathcal{G}_{\text{null}}$ for any set of channel matrices and transmit powers of the users follows readily from [22, Th. 6], [44] (i.e., quasi-concave payoff functions and convex compact strategy sets).

The structure of the Nash equilibria as given in (16) follows from the equivalence between game $\mathcal{G}_{\text{null}}$ and the relaxed version in (9), as detailed next. Any NE $\left\{\mathbf{Q}_q^{\star}\right\}_{q\in\Omega}$ of game in (9) (whose existence follows from the existence of a solution for $\mathcal{G}_{\text{null}}$) satisfies the following: $\mathcal{R}(\mathbf{Q}_q^{\star}) \subseteq \mathcal{R}\left(\mathbf{P}_{\mathcal{R}(\mathbf{U}_q^{\perp})}\mathbf{H}_{qq}^H\right) \subseteq \mathcal{R}\left(\mathbf{U}_q^{\perp}\right)$, $\forall\, q \in \Omega$. This means that all the solutions to (9) can be written as[4]

$$\mathbf{Q}_q^{\star} = \mathbf{U}_q^{\perp}\overline{\mathbf{Q}}_q^{\star}\mathbf{U}_q^{\perp H}, \quad \forall\, q \in \Omega \tag{43}$$

where $\{\overline{\mathbf{Q}}_q^{\star}\}_{q\in\Omega}$, with each $\overline{\mathbf{Q}}_q^{\star} \in \mathbb{S}_+^{r_{U_q^{\perp}}\times r_{U_q^{\perp}}}$, are the Nash equilibria of the following lower-dimensional game:

$$\begin{array}{ll}\underset{\overline{\mathbf{Q}}_q\succeq\mathbf{0}}{\text{maximize}} & \log\det\left(\mathbf{I}+\mathbf{U}_q^{\perp H}\mathbf{H}_{qq}^H\overline{\mathbf{R}}_{-q}^{-1}(\overline{\mathbf{Q}}_{-q})\mathbf{H}_{qq}\mathbf{U}_q^{\perp}\overline{\mathbf{Q}}_q\right)\\ \text{subject to} & \operatorname{Tr}\left(\overline{\mathbf{Q}}_q\right)\le P_q\end{array} \tag{44}$$

for all $q \in \Omega$, where $\overline{\mathbf{R}}_{-q}(\overline{\mathbf{Q}}_{-q}) \triangleq \mathbf{R}_{n_{R_q}} + \sum_{r\neq q}\mathbf{H}_{rq}\mathbf{U}_r^{\perp}\overline{\mathbf{Q}}_r\mathbf{U}_r^{\perp H}\mathbf{H}_{rq}^H$. The solutions to (44) are the fixed-points of the following nonlinear matrix-value equation [22], [24]:

$$\overline{\mathbf{Q}}_q^{\star} = \mathsf{WF}_q\left(\mathbf{U}_q^{\perp H}\mathbf{H}_{qq}^H\overline{\mathbf{R}}_{-q}^{-1}(\overline{\mathbf{Q}}_{-q}^{\star})\mathbf{H}_{qq}\mathbf{U}_q^{\perp}\right) \quad \forall\, q \in \Omega \tag{45}$$

with $\mathsf{WF}_q(\cdot)$ defined in (15). The structure of the Nash equilibria of $\mathcal{G}_{\text{null}}$ as given in (16) follows readily from (45) and (43).

As far as the uniqueness of the NE of $\mathcal{G}_{\text{null}}$ is concerned, a sufficient condition for the uniqueness of the equilibrium is that the multiuser waterfilling mapping in (45) be a contraction with respect to some norm [25, Prop. 1.1.(a)]. Theorem 5 and Theorem 7 in [24] provide such (sufficient) conditions valid for the cases of full row-rank and full column-rank (direct) channel matrices, respectively. Since the channel matrices $\mathbf{H}_{qq}\mathbf{U}_q^{\perp}$ in (44) may be rank deficient, we cannot apply directly Theorem 5 or Theorem 7 in [24] to game in (44), but we need first to rewrite it in a more useful form, as detailed next.

Following the same approach as before and using the definitions introduced in Section III-A, it is not difficult to show that, for any $q \in \overline{\Omega}$ and $\overline{\mathbf{Q}}_{-q} \succeq \mathbf{0}$, the best-response $\overline{\mathbf{Q}}_q^{\star} = \mathsf{WF}_q\left(\mathbf{U}_q^{\perp H}\mathbf{H}_{qq}^H\overline{\mathbf{R}}_{-q}^{-1}(\overline{\mathbf{Q}}_{-q})\mathbf{H}_{qq}\mathbf{U}_q^{\perp}\right)$—the solution of the rate-maximization problem in (44) for a given $\overline{\mathbf{Q}}_{-q} \succeq \mathbf{0}$—will be orthogonal to the null space of $\mathbf{H}_{qq}\mathbf{U}_q^{\perp}$, whatever $\overline{\mathbf{Q}}_{-q} \succeq \mathbf{0}$ is, implying $\overline{\mathbf{Q}}_q^{\star} = \mathbf{V}_{q,1}\overline{\overline{\mathbf{Q}}}_q^{\star}\mathbf{V}_{q,1}^H$, for some $\overline{\overline{\mathbf{Q}}}_q^{\star} \in \mathbb{S}_+^{r_{H_{qq}U_q^{\perp}}\times r_{H_{qq}U_q^{\perp}}}$ and all $q \in \overline{\Omega}$. Thus, the best-response of each user $q \in \overline{\Omega}$ belongs to the following class of matrices:

$$\overline{\mathbf{Q}}_q = \mathbf{V}_{q,1}\overline{\overline{\mathbf{Q}}}_q\mathbf{V}_{q,1}^H \tag{46}$$

with

$$\overline{\overline{\mathbf{Q}}}_q \in \left\{\mathbf{X} \in \mathbb{S}_+^{r_{H_{qq}U_q^{\perp}}\times r_{H_{qq}U_q^{\perp}}} : \operatorname{Tr}(\mathbf{X}) \le P_q\right\}, \quad q \in \overline{\Omega}. \tag{47}$$

Using (46) and introducing the (possibly) lower-dimensional covariance matrices $\{\widetilde{\mathbf{Q}}_q\}_{q\in\Omega}$, defined as

$$\widetilde{\mathbf{Q}}_q \triangleq \begin{cases}\overline{\overline{\mathbf{Q}}}_q \in \mathbb{S}_+^{r_{H_{qq}U_q^{\perp}}\times r_{H_{qq}U_q^{\perp}}}, & \text{if } q \in \overline{\Omega}\\ \overline{\mathbf{Q}}_q \in \mathbb{S}_+^{r_{U_q^{\perp}}\times r_{U_q^{\perp}}}, & \text{otherwise}\end{cases} \tag{48}$$

[4]Observe that the solution to the rate maximization problem in (9) for any given $\mathbf{Q}_{-q} \succeq \mathbf{0}$ is unique, because of the strict concavity of the rate function in $\mathbf{Q}_q$ on $\mathcal{N}\left(\mathbf{H}_{qq}\mathbf{P}_{\mathcal{R}(\mathbf{U}_q^{\perp})}\right)^{\perp}$.

game in (44) can be recast in the following lower-dimensional game:

$$\begin{aligned}&\underset{\widetilde{\mathbf{Q}}_q\succeq\mathbf{0}}{\text{maximize}} && \log\det\left(\mathbf{I}+\tilde{\mathbf{H}}_{qq}^H\widetilde{\mathbf{R}}_{-q}^{-1}(\widetilde{\mathbf{Q}}_{-q})\tilde{\mathbf{H}}_{qq}\widetilde{\mathbf{Q}}_q\right)\\ &\text{subject to} && \operatorname{Tr}\left(\widetilde{\mathbf{Q}}_q\right)\le P_q\end{aligned} \tag{49}$$

for all $q\in\Omega$, where $\widetilde{\mathbf{R}}_{-q}(\widetilde{\mathbf{Q}}_{-q})\triangleq\mathbf{R}_{n_{R_q}}+\sum_{r\neq q}\tilde{\mathbf{H}}_{rq}\widetilde{\mathbf{Q}}_q\tilde{\mathbf{H}}_{rq}^H$ and $\tilde{\mathbf{H}}_{rq}$ are the modified channel matrices defined in (12). Observe that, in the game given in (49), all channel matrices $\tilde{\mathbf{H}}_{qq}$ are full column-rank matrices by construction. We can thus use Theorem 7 in [24] to obtain the desired sufficient condition for the uniqueness of the NE in (49) (and thus also $\mathcal{G}_{\text{null}}$) as given in (C1). This completes the proof of the theorem. ■

APPENDIX B
PROOF OF THEOREM 6 AND THEOREM 7

We first introduce some intermediate results that are instrumental to prove Theorems 6 and 7.

A. Miscellaneous Results

In this section we provide some interesting properties of the MIMO waterfilling mapping $\mathsf{WF}=(\mathsf{WF}_q)_{q\in\Omega}$ applied to game $\mathcal{G}_\alpha$ in (20) [see (15)] that will be used to study the relationship between $\mathcal{G}_\alpha$ and $\mathcal{G}_\infty$ as $\alpha\to+\infty$. To this end, it is useful to make explicit the dependence of each best-response $\mathsf{WF}_q\left(\mathbf{H}_{qq}^H\mathbf{R}_{-q,\alpha}^{-1}(\mathbf{Q}_{-q})\mathbf{H}_{qq}\right)$ in (27) on the strategies $\mathbf{Q}_{-q}$ of the other players, the channels $\{\mathbf{H}_{rq}\}_{r\in\Omega}$, and the constant $\alpha>0$. We thus introduce the following notation: $\mathsf{WF}_\alpha(\{\mathbf{H}_{rq}\}_{r,q\in\Omega},\mathbf{Q})\triangleq(\mathsf{WF}_{q,\alpha}(\{\mathbf{H}_{rq}\}_{r\in\Omega},\mathbf{Q}_{-q}))_{q\in\Omega}$, where

$$\mathsf{WF}_{q,\alpha}(\{\mathbf{H}_{rq}\}_{r\in\Omega},\mathbf{Q}_{-q})\triangleq\mathsf{WF}_q\left(\mathbf{H}_{qq}^H\mathbf{R}_{-q,\alpha}^{-1}(\mathbf{Q}_{-q})\mathbf{H}_{qq}\right) \tag{50}$$

with $\mathbf{R}_{-q,\alpha}(\mathbf{Q}_{-q})$ defined in (23). Note that in the following we may omit some of these dependencies whenever they are not needed. Similarly, given $\mathcal{G}_\infty$ in (22), the MIMO waterfilling mapping $\mathsf{WF}_\infty(\{\mathbf{H}_{rq}\}_{r,q\in\Omega},\mathbf{Q})\triangleq(\mathsf{WF}_{q,\infty}(\{\mathbf{H}_{rq}\}_{r\in\Omega},\mathbf{Q}_{-q}))_{q\in\Omega}$ of $\mathcal{G}_\infty$ is defined according to (28):

$$\mathsf{WF}_{q,\infty}(\{\mathbf{H}_{rq}\}_{r\in\Omega},\mathbf{Q}_{-q})\triangleq\mathsf{WF}_q\left(\hat{\mathbf{H}}_{qq}^H\hat{\mathbf{R}}_{-q}^{-1}(\mathbf{Q}_{-q,})\hat{\mathbf{H}}_{qq}\right) \tag{51}$$

with $\hat{\mathbf{R}}_{-q}^{-1}(\mathbf{Q}_{-q})$, and $\hat{\mathbf{H}}_{qq}$ given in (23) and (24), respectively.

1) Continuity of the Mapping WF_α*:* We consider here each component $\mathsf{WF}_{q,\alpha}$ and $\mathsf{WF}_{q,\infty}$ in (50) and (51) as a joint function of $(\{\mathbf{H}_{rq}\}_{r\in\Omega},\mathbf{Q}_{-q})$ and $\alpha>0$, and prove the continuity at every point $(\{\mathbf{H}_{rq}\}_{r\in\Omega},\mathbf{Q}_{-q},\alpha)$, with $\mathbf{Q}_{-q}\succeq\mathbf{0}$ and $\alpha>0$. We introduce first the following two lemmas.

Lemma 12: Let $\mathcal{U}$ and $\mathcal{U}^\perp$ be a pair of complementary subspaces of $\mathbb{C}^n$, with dimensions $r\le n$ and $n-r$, respectively, and let $\mathbf{U}\in\mathbb{C}^{n\times r}$ and $\mathbf{U}^\perp\in\mathbb{C}^{n\times n-r}$ be any pair of matrices such that $\mathcal{R}(\mathbf{U})=\mathcal{U}$ and $\mathcal{R}(\mathbf{U}^\perp)=\mathcal{U}^\perp$. Then, for any nonsingular matrix $\mathbf{A}\in\mathbb{C}^{n\times n}$, we have

$$\mathbf{P}_{\mathcal{R}(\mathbf{AU})^\perp}=\mathbf{P}_{\mathcal{R}(\mathbf{A}^{-H}\mathbf{U}^\perp)}. \tag{52}$$

□

Lemma 13: Let $\mathcal{U}$ and $\mathcal{U}^\perp$ be a pair of complementary subspaces of $\mathbb{C}^n$, with dimensions $r\le n$ and $n-r$, respectively, and let $\mathbf{U}\in\mathbb{C}^{n\times r}$ and $\mathbf{U}^\perp\in\mathbb{C}^{n\times n-r}$ be any pair of matrices such that $\mathcal{R}(\mathbf{U})=\mathcal{U}$ and $\mathcal{R}(\mathbf{U}^\perp)=\mathcal{U}^\perp$. Then, for any matrix $\mathbf{R}\in\mathbb{S}_{++}^{n\times n}$, we have

$$\lim_{\alpha\to+\infty}\left(\mathbf{R}+\alpha\mathbf{U}\mathbf{U}^H\right)^{-1}=\mathbf{U}^\perp\left(\mathbf{U}^{\perp H}\mathbf{R}\mathbf{U}^\perp\right)^{-1}\mathbf{U}^{\perp H}. \tag{53}$$

Proof: For any given $\alpha>0$, invoking the matrix inversion lemma,[5] we have

$$\begin{aligned}\left(\mathbf{R}+\alpha\mathbf{U}\mathbf{U}^H\right)^{-1} &=\mathbf{R}^{-1/2}\left(\mathbf{I}-\mathbf{R}^{-H/2}\mathbf{U}\left(\mathbf{U}^H\mathbf{R}^{-1/2}\mathbf{R}^{-H/2}\mathbf{U}+\frac{1}{\alpha}\mathbf{I}\right)^{-1}\right.\\ &\qquad\left.\times\,\mathbf{U}^H\mathbf{R}^{-1/2}\right)\mathbf{R}^{-H/2}\end{aligned} \tag{54}$$

where $\mathbf{R}^{1/2}$ is a square root of the positive definite matrix $\mathbf{R}$, satisfying $\mathbf{R}^{1/2}\mathbf{R}^{H/2}=\mathbf{R}$. Introducing the orthogonal projection onto $\mathcal{R}(\mathbf{R}^{-H/2}\mathbf{U})$, defined as (observe that matrix $\mathbf{R}^{-H/2}\mathbf{U}$ is full column-rank by construction)

$$\mathbf{P}_{\mathcal{R}(\mathbf{R}^{-H/2}\mathbf{U})}=\mathbf{R}^{-H/2}\mathbf{U}\left(\mathbf{U}^H\mathbf{R}^{-1}\mathbf{U}\right)^{-1}\mathbf{U}^H\mathbf{R}^{-1/2}, \tag{55}$$

and using the fact that $\left(\mathbf{U}^H\mathbf{R}^{-1}\mathbf{U}+(1/\alpha)\mathbf{I}\right)^{-1}$ is a continuous function on $\alpha>0$ and thus $\left(\mathbf{U}^H\mathbf{R}^{-1}\mathbf{U}+(1/\alpha)\mathbf{I}\right)^{-1}\underset{\alpha\to+\infty}{\longrightarrow}\left(\mathbf{U}^H\mathbf{R}^{-1}\mathbf{U}\right)^{-1}$—implied from $\operatorname{rank}(\mathbf{U}^H\mathbf{R}^{-1}\mathbf{U}+(1/\alpha)\mathbf{I})=r$, for all $\alpha>0$ [27, Th. 10.7.1]—we obtain

$$\begin{aligned}\lim_{\alpha\to+\infty}\left(\mathbf{R}+\alpha\mathbf{U}\mathbf{U}^H\right)^{-1}&=\mathbf{R}^{-1/2}\mathbf{P}_{\mathcal{R}(\mathbf{R}^{-H/2}\mathbf{U})^\perp}\mathbf{R}^{-H/2} &(56)\\ &=\mathbf{R}^{-1/2}\mathbf{P}_{\mathcal{R}(\mathbf{R}^{1/2}\mathbf{U}^\perp)}\mathbf{R}^{-H/2} &(57)\\ &=\mathbf{U}^\perp\left(\mathbf{U}^{\perp H}\mathbf{R}\mathbf{U}^\perp\right)^{-1}\mathbf{U}^{\perp H} &(58)\end{aligned}$$

where (56) follows from the definition of $\mathbf{P}_{\mathcal{R}(\mathbf{R}^{-H/2}\mathbf{U})}$ as given in (55) and the basic equality $\mathbf{P}_{\mathcal{R}(\mathbf{R}^{-H/2}\mathbf{U})^\perp}=\mathbf{I}-\mathbf{P}_{\mathcal{R}(\mathbf{R}^{-H/2}\mathbf{U})}$ [45]; (57) follows from Lemma 12 [see (52)]; and (58) follows directly from the definition of projection $\mathbf{P}_{\mathcal{R}(\mathbf{R}^{1/2}\mathbf{U}^\perp)}$ (using the fact that $\mathbf{R}^{1/2}\mathbf{U}^\perp$ is a full column-rank matrix). ■

We can now state the main result of this section.

Lemma 14 (Continuity of the WF_α*):* The MIMO waterfilling $\mathsf{WF}_{q,\alpha}(\{\mathbf{H}_{rq}\}_{r\in\Omega},\mathbf{Q}_{-q})$ defined in (50) is continuous at $(\{\mathbf{H}_{rq}\}_{r\in\Omega},\ \mathbf{Q}_{-q},\alpha)$, for any $\{\mathbf{H}_{rq}\}_{r\in\Omega}$, $\mathbf{Q}_{-q}\succeq\mathbf{0}$, and $\alpha>0$. In particular, we have

$$\lim_{\alpha\to+\infty}\mathsf{WF}_{q,\alpha}(\{\mathbf{H}_{rq}\}_{r\in\Omega},\mathbf{Q}_{-q})=\mathsf{WF}_{q,\infty}(\{\mathbf{H}_{rq}\}_{r\in\Omega},\mathbf{Q}_{-q}) \tag{59}$$

where $\mathsf{WF}_{q,\infty}(\{\mathbf{H}_{rq}\}_{r\in\Omega},\mathbf{Q}_{-q})$ is defined in (51).

Proof: Because of the space limitation, we provide only a sketch of the proof. Looking at each $\mathsf{WF}_{q,\alpha}(\{\mathbf{H}_{rq}\}_{r\in\Omega},\mathbf{Q}_{-q})$ as the solution to the qth parametric optimization problem defined in (20), where $\{\mathbf{H}_{rq}\}_{r\in\Omega}$, $\mathbf{Q}_{-q}\succeq\mathbf{0}$, and $\alpha>0$ are the parameters of the problem [46], it is not difficult to check

[5] The expression of the matrix inversion lemma is: $(\mathbf{A}+\mathbf{BCD})^{-1}=\mathbf{A}^{-1}-\mathbf{A}^{-1}\mathbf{B}(\mathbf{DA}^{-1}\mathbf{B}+\mathbf{C}^{-1})^{-1}\mathbf{DA}^{-1}$ [28].

that such a parametric optimization satisfies conditions in [46, Th. 7.11] for any $\{\mathbf{H}_{rq}\}_{r\in\Omega}$, $\mathbf{Q}_{-q}\succeq \mathbf{0}$, and $\alpha>0$; which, invoking [46, Th. 7.9 (ii)], leads to the continuity of the waterfilling $\mathsf{WF}_{q,\alpha}(\{\mathbf{H}_{rq}\}_{r\in\Omega},\mathbf{Q}_{-q})$ at $(\{\mathbf{H}_{rq}\}_{r\in\Omega},\mathbf{Q}_{-q},\alpha)$, for any $\{\mathbf{H}_{rq}\}_{r\in\Omega}$, $\mathbf{Q}_{-q}\succeq \mathbf{0}$, and $\alpha>0$. Building on this result and using Lemma 13 we obtain the equality in (59). ■

2) Contraction Property of the Mapping WF_α*:* Given the multiuser waterfilling mapping $\mathsf{WF}_\alpha=\mathsf{WF}_\alpha(\mathbf{Q})$ defined in (50), we introduce the following block-maximum norm on $\mathbb{C}^{n\times n}$, with $n=n_{T_1}+\cdots+n_{T_Q}$, defined as [25]

$$\|\mathsf{WF}_\alpha(\mathbf{Q})\|_{F,\text{block}}^{\mathbf{w}}\triangleq\max_{q\in\Omega}\frac{\|\mathsf{WF}_{q,\alpha}(\mathbf{Q}_{-q})\|_F}{w_q}\tag{60}$$

where $\|\cdot\|_F$ is the Frobenius norm and $\mathbf{w}\triangleq[w_1,\ldots,w_Q]^T>\mathbf{0}$ is any positive weight vector; and the *matrix* norm

$$\|\mathbf{A}\|_{\infty,\text{mat}}^{\mathbf{w}}\triangleq\max_q\frac{1}{w_q}\sum_{r=1}^{Q}|[\mathbf{A}]_{qr}|\,w_r,\ \text{for }\mathbf{A}\in\mathbb{R}^{Q\times Q}.\tag{61}$$

Introducing the set $\mathcal{Q}=\mathcal{Q}_1\times\cdots\times\mathcal{Q}_Q$, with[6]

$$\mathcal{Q}_q\triangleq\{\mathbf{X}\in\mathbb{S}_+^{n_{T_q}\times n_{T_q}}:\mathrm{Tr}\{\mathbf{X}\}=P_q\}\tag{62}$$

and invoking [24, Th. 5], we have the following contraction properties for $\mathsf{WF}_\alpha(\mathbf{Q})$.

Lemma 15 (Contraction Property of WF_α *Mapping):* Suppose $\mathrm{rank}(\mathbf{H}_{qq})=n_{R_q},\forall\, q\in\Omega$ (i.e., full rank fat/square matrix). Then, for any given $\mathbf{w}\triangleq[w_1,\ldots,w_Q]^T>\mathbf{0}$, the mapping WF_α is Lipschitz continuous on $\mathcal{Q}$:

$$\begin{aligned}\left\|\mathsf{WF}_\alpha(\mathbf{Q}^{(1)})-\mathsf{WF}_\alpha(\mathbf{Q}^{(2)})\right\|_{F,\text{block}}^{\mathbf{w}}\leq\|\mathbf{S}\|_{\infty,\text{mat}}^{\mathbf{w}}\\ \times\left\|\mathbf{Q}^{(1)}-\mathbf{Q}^{(2)}\right\|_{F,\text{block}}^{\mathbf{w}},\end{aligned}\tag{63}$$

$\forall\,\mathbf{Q}^{(1)},\mathbf{Q}^{(2)}\in\mathcal{Q}$, where $\mathbf{S}$ $\|\cdot\|_{F,\text{block}}^{\mathbf{w}}$, and $\|\cdot\|_{\infty,\text{mat}}^{\mathbf{w}}$ are defined in (25), (60) and (61), respectively. □

The important result stated in Lemma 15 is that condition guaranteeing the waterfilling mapping WF_α to be a contraction in the block-maximum norm $\|\cdot\|_{F,\text{block}}^{\mathbf{w}}$, i.e., $-\|\mathbf{S}\|_{\infty,\text{mat}}^{\mathbf{w}}<1$ for some $\mathbf{w}>\mathbf{0}$, is independent of $\alpha>0$. This property together with the continuity of WF_α (Lemma 14) are the key points necessary to prove the asymptotic equivalence between $\mathcal{G}_\alpha$ and $\mathcal{G}_\infty$. We also need the following.

3) Uniform Convergence and Continuity: To study the convergence properties of the asynchronous MIMO IWFA applied to game $\mathcal{G}_\alpha$ as $\alpha\to+\infty$ we need the following result that comes from [47, Theorem 7.11].

Theorem 16: Let $\{\mathbf{f}^{(n)}(\mathbf{x})\}_{n\geq 0}$ denote a sequence of functions $\mathbf{f}^{(n)}:\mathcal{X}\mapsto\mathcal{D}$, where $\mathcal{X}$ and $\mathcal{D}$ are sets in a metric space, and let $\mathbf{x}_0$ be a limit point of $\mathcal{X}$. Suppose that the following conditions are satisfied:

(C.1) The sequence $\{\mathbf{f}^{(n)}(\mathbf{x})\}_n$ converges uniformly on $\mathcal{X}$;
(C.2) $\lim_{\mathbf{x}\to\mathbf{x}_0}\mathbf{f}^{(n)}(\mathbf{x})=\mathbf{a}^{(n)}$, for all $n=0,1,\ldots$.

Then, the sequence $\{\mathbf{a}^{(n)}\}_n$ converges on $\mathcal{D}$, and

$$\lim_{n\to+\infty}\lim_{\mathbf{x}\to\mathbf{x}_0}\mathbf{f}^{(n)}(\mathbf{x})=\lim_{\mathbf{x}\to\mathbf{x}_0}\lim_{n\to+\infty}\mathbf{f}^{(n)}(\mathbf{x}).\tag{64}$$

□

B. Proof of Theorem 6 and Theorem 7

Using the intermediate results introduced in Appendices B–A, we can now prove Theorem 7. Theorem 6 comes as a consequence of Theorem 7(b).

1) Proof of Theorem 7(a) (Uniform Convergence of the Asynchronous IWFA): Assume that condition (C4) of Theorem 4 is satisfied. According to [25, Cor. 6.1], (C4) is equivalent to $\beta=\beta(\mathbf{w},\mathbf{S})\triangleq\|\mathbf{S}\|_{\infty,\text{mat}}^{\mathbf{w}}<1$, for some $\mathbf{w}>\mathbf{0}$, implying that the waterfilling mapping WF_α defined in (50) is a block-contraction in the norm $\|\cdot\|_{F,\text{block}}^{\mathbf{w}}$ (cf. Lemma 15) and thus admits a unique fixed-point [25, Prop. 1.1]. We denote such a fixed-point—the NE of game $\mathcal{G}_\alpha$—by $\mathbf{Q}_\alpha^\star$. The proof of the global convergence of Algorithm 1 based on the mapping WF_α (hereafter called Algorithm 1 only) consists in showing that, under $\beta<1$, the algorithm satisfies the asynchronous convergence theorem in [25] that can be restated for our purpose as follows.

Theorem 17: Given the waterfilling mapping $\mathsf{WF}_\alpha=(\mathsf{WF}_{q,\alpha})_{q\in\Omega}:\mathcal{Q}\mapsto\mathcal{Q}$, with $\mathsf{WF}_{q,\alpha}(\cdot)$ and $\mathcal{Q}_q$ defined in (50) and in (62), respectively, assume that there exists a sequence of nonempty sets $\{\mathcal{X}(k)\}_k$ with $\cdots\subset\mathcal{X}(k+1)\subset\mathcal{X}(k)\subset\cdots\subset\mathcal{Q}$, satisfying the next two conditions:

(C.1) (Synchronous Convergence Condition): $\mathsf{WF}_\alpha(\mathbf{Q})\in\mathcal{X}(k+1)$ for all k and $\mathbf{Q}\in\mathcal{X}(k)$. Furthermore, if $\left\{\mathbf{Q}_\alpha^{(k)}\right\}$ is a sequence such that $\mathbf{Q}_\alpha^{(k)}\in\mathcal{X}(k)$ for every k, then every limit point of $\left\{\mathbf{Q}_\alpha^{(k)}\right\}$ is a fixed point of WF_α.

(C.2) (Box Condition): For every k there exist sets $\mathcal{X}_q(k)\subset\mathcal{Q}_q$ such that: $\mathcal{X}(k)=\mathcal{X}_1(k)\times\cdots\times\mathcal{X}_Q(k)$.

Then, every limit point of $\left\{\mathbf{Q}_\alpha^{(n)}\right\}_{n\in\mathcal{T}}$ generated by the asynchronous MIMO IWFA in Algorithm 1 based on mapping WF_α, starting from $\mathbf{Q}^{(0)}\in\mathcal{X}(0)$, is a fixed point of WF_α. □

Following similar steps as in the proof of [24, Th. 12], one can prove that Algorithm 1 satisfies Theorem 17, using as candidate sets $\mathcal{X}(k)$ the following sets [note that $\mathcal{X}(k)$ in (65) is different from $\mathcal{X}(k)$ used to prove [24, Th.12]]:

$$\mathcal{X}(k)=\left\{\mathbf{Q}\in\mathcal{Q}:\sup_{\alpha>0}\|\mathbf{Q}-\mathbf{Q}_\alpha^\star\|_{F,\text{block}}^{\mathbf{w}}\leq\beta^k e_{\max}^{(0)}\right\}\subset\mathcal{Q}\tag{65}$$

with $k\geq 1$ and $e_{\max}^{(0)}$ defined as

$$e_{\max}^{(0)}\triangleq\sup_{\alpha>0}\max_{\mathbf{Q}^{(0)}\in\mathcal{Q}}\|\mathbf{Q}^{(0)}-\mathbf{Q}_\alpha^\star\|_{F,\text{block}}^{\mathbf{w}}<+\infty.\tag{66}$$

To complete the proof, we need to show that, under $\beta<1$, the convergence of the algorithm is uniform w.r.t. α on $\mathbb{R}_{++}$, i.e.,

$$\begin{aligned}&\forall\delta>0\quad\exists n_\delta\geq 0\text{ such that}\\ &n\geq n_\delta\Rightarrow\left\|\mathbf{Q}_\alpha^{(n)}-\mathbf{Q}_\alpha^\star\right\|_{F,\text{block}}^{\mathbf{w}}\leq\delta,\quad\forall\,\alpha\in\mathbb{R}_{++}.\end{aligned}\tag{67}$$

[6]Note that there is no loss of generality in considering in (62) the trace constraint with equality rather than inequality, since at the optimum to each problem in (20), the power constraint must be satisfied with equality.

Under $\beta < 1$, Algorithm 1 satisfies Theorem 17 with set $\mathcal{X}(k)$ defined in (65). It is not difficult to see that this means that, for any $k \geq 0$, there exists a time index $n_k \in \mathcal{T}$, independent of α, such that

$$\mathbf{Q}_{\alpha}^{(n)} \in \mathcal{X}(k), \quad \forall\, n \geq n_k, \ \forall\, \alpha \in \mathbb{R}_{++}. \tag{68}$$

In words: for any $k \geq 0$, after some time $n_k \in \mathcal{T}$, the set of users' covariance matrices $\mathbf{Q}_{\alpha}^{(n)}$ generated by the algorithm from initial conditions in $\mathcal{Q}$, will eventually enter and stay in the set $\mathcal{X}(k)$. According to our choice of $\mathcal{X}(k)$, (68) implies that, for every $k \geq 0$, there exists a $n_k \in \mathcal{T}$ such that the quantity $\sup_{\alpha>0} \left\| \mathbf{Q}_{\alpha}^{(n)} - \mathbf{Q}_{\alpha}^{\star} \right\|_{F,\text{block}}^{\mathbf{w}}$ eventually becomes smaller than $\beta^k\, e_{\max}^{(0)}$ when $n \geq n_k$. Observe that $\beta^k\, e_{\max}^{(0)}$ converges to 0 as $k \to +\infty$ (recall that $\beta < 1$). Hence, given any $\delta > 0$ in (67), one can always find a (possibly large) time index $n_\delta \in \mathcal{T}$, *independent of* α, such that condition in (67) is satisfied; it is indeed sufficient to choose $k = \bar{k}$ such that $\beta^{\bar{k}}\, e_{\max}^{(0)} \leq \delta$ and n_k in (68) equal to $n_{\bar{k}}$; which guarantees that condition (67) is satisfied with $n_\delta = n_{\bar{k}}$. This completes the proof of the statement (a) of the theorem.

2) Proof of Theorem 7(b) (Relationship Between $\mathbf{Q}_{\infty}^{\star}$ and $\mathbf{Q}_{\alpha}^{\star}$, as $\alpha \to +\infty$): The proof consists in showing that, under (C4) of Theorem 4, Algorithm 1 based on mapping WF_{α} defined in (50) satisfies conditions (C.1) and (C.2) of Theorem 16, implying the equality in (32). We use the following identifications:

$$\begin{aligned} &\mathbf{x} \Leftrightarrow \alpha, \quad \mathbf{x}_0 \Leftrightarrow +\infty, \quad \mathcal{X} \Leftrightarrow \mathbb{R}_{+}, \quad \mathcal{D} \Leftrightarrow \mathcal{Q}, \\ &\mathbf{f}^{(n)}(\mathbf{x}) \Leftrightarrow \mathbf{Q}_{\alpha}^{(n)}, \quad \mathbf{a}^{(n)} \Leftrightarrow \mathbf{Q}_{\infty}^{(n)} \end{aligned} \tag{69}$$

where $\mathbf{Q}_{\alpha}^{(n)} = \left(\mathbf{Q}_{q,\alpha}^{(n)}\right)_{q\in\Omega}$ and $\mathbf{Q}_{\infty}^{(n)} = \left(\mathbf{Q}_{q,\infty}^{(n)}\right)_{q\in\Omega}$ denote the sequence of users covariance matrices generated by Algorithm 1 based on mappings WF_{α} and WF_{∞} defined in (50) and (51), respectively; and $\mathcal{Q} = \mathcal{Q}_1 \times \cdots \times \mathcal{Q}_Q$, with $\mathcal{Q}_q$ denoting the admissible strategy set of user q, defined in (62).

Condition (C.1)—$\mathbf{Q}_{\alpha}^{(n)}$ converges uniformly w.r.t. α on $\mathbb{R}_{++}$, as $n \to +\infty$—follows readily from Theorem 7(a).

Condition (C.2)—$\lim_{\alpha\to+\infty} \mathbf{Q}_{\alpha}^{(n)} = \mathbf{Q}_{\infty}^{(n)}$ for all $n \in \mathcal{T}$—is proved using Lemma 14, as detailed next.

For any given schedule $\{\mathcal{T}_q, \tau_r^q(n)\}_{r,q\in\Omega}$ and $\alpha > 0$, the sequence $\left\{\mathbf{Q}_{\alpha}^{(n)}\right\}_{n\in\mathcal{T}}$ can be expressed as a function of the initial point $\mathbf{Q}^{(0)} \in \mathcal{Q}$ through $\mathbf{Q}_{\alpha}^{(n)} = \mathsf{T}_{\alpha}^{(n)}(\mathbf{Q}^{(0)})$ for all $n \in \mathcal{T}$, where $\mathsf{T}_{\alpha}^{(n)}$ is a proper composition of the best-responses $\mathsf{WF}_{q,\alpha}$'s of the users, which depends on the specific updating schedule $\{\mathcal{T}_q, \tau_r^q(n)\}_{r,q\in\Omega}$ performed by the algorithm. An explicit expression for $\mathsf{T}_{\alpha}^{(n)}$ can be obtained only for special choices of $\{\mathcal{T}_q, \tau_r^q(n)\}_{r,q\in\Omega}$. For example, if $\{\mathcal{T}_q, \tau_r^q(n)\}_{r,q\in\Omega}$ are chosen so that users' updates in Algorithm 1 are *simultaneous*—$\mathcal{T}_q = \mathbb{N}_+$ and $\tau_r^q(n) = n$ for all $r, q, \in \Omega$—then $\mathsf{T}_{\alpha}^{(n)}$ can be written as $\mathsf{T}_{\alpha}^{(n)}(\mathbf{Q}^{(0)}) = \mathsf{T}_{\alpha} \circ \cdots \circ \mathsf{T}_{\alpha}(\mathbf{Q}^{(0)})$, where the composition "$\circ$" is applied n times and $\mathsf{T}_{\alpha}(\mathbf{Q}) = \mathsf{WF}_{\alpha}(\mathbf{Q})$. In a *sequential* update, such as $\mathcal{T}_q = \{kQ + q, k \in \mathbb{N}_+\}$ and $\tau_r^q(n) = n$ for all $r, q, \in \Omega$, we have $\mathsf{T}_{\alpha}(\mathbf{Q}) \triangleq \mathsf{S}_{Q,\alpha} \circ \cdots \circ \mathsf{S}_{1,\alpha}(\mathbf{Q}^{(0)})$ with each $\mathsf{S}_{q,\alpha}(\mathbf{Q}) \triangleq (\mathbf{Q}_1, \ldots, \mathbf{Q}_{q-1}, \mathsf{WF}_{q,\alpha}(\mathbf{Q}_{-q}), \mathbf{Q}_{q+1}, \ldots, \mathbf{Q}_Q)$. Note that we do not need to know explicitly the expression of $\mathsf{T}_{\alpha}^{(n)}$ to complete the proof of the theorem. Similarly, given Algorithm 1 based on WF_{∞} and using the same schedule $\{\mathcal{T}_q, \tau_r^q(n)\}_{r,q\in\Omega}$ and initial point $\mathbf{Q}^{(0)} \in \mathcal{Q}$ as for the mapping WF_{α} above, we define $\mathsf{T}_{\infty}^{(n)}$ as $\mathbf{Q}_{\infty}^{(n)} = \mathsf{T}_{\infty}^{(n)}(\mathbf{Q}^{(0)})$. Then, condition (C.2) of Theorem 16 is equivalent to the following:

$$\lim_{\alpha\to+\infty} \mathsf{T}_{\alpha}^{(n)}(\mathbf{Q}^{(0)}) = \mathsf{T}_{\infty}^{(n)}(\mathbf{Q}^{(0)}), \quad \forall\, \mathbf{Q}^{(0)} \in \mathcal{Q}, \ n \in \mathcal{T}. \tag{70}$$

The proof of (70) follows from (59) (Lemma 14) and the continuity of $\mathsf{T}_{\alpha}(\mathbf{Q}^{(0)})$ on $\alpha > 0$, implied from the following facts: i) each $\mathsf{WF}_{q,\alpha}(\mathbf{Q}_{-q})$ is continuous at $(\alpha, \mathbf{Q}_{-q})$, for all $\alpha > 0$ and $\mathbf{Q}_{-q} \succeq \mathbf{0}$ (cf. Lemma 14); ii) for all $n \in \mathcal{T}$, $\mathsf{T}_{\alpha}^{(n)}(\mathbf{Q}^{(0)})$ is composition of continuous functions $\mathsf{WF}_{q,\alpha}(\mathbf{Q}_{-q})$ at any $(\alpha, \mathbf{Q}_{-q})$; and iii) composition of continuous functions leads to a continuous function on its domain [47, Theorem 4.7].

It follows from Theorem 16 that, under (C4), Algorithm 1 based on WF_{∞} [i.e., $\mathbf{Q}_{\infty}^{(n)} = \mathsf{T}_{\infty}^{(n)}(\mathbf{Q}^{(0)})$] and starting from *any* initial point $\mathbf{Q}^{(0)} \in \mathcal{Q}$ asymptotically converges to the *unique* fixed-point $\mathbf{Q}_{\infty}^{\star}$ of WF_{∞}, and (32) holds true [Note that every limit point of $\left\{\mathbf{Q}_{\infty}^{(n)}\right\}_{n\in\mathcal{T}}$ is a fixed-point $\mathbf{Q}_{\infty}^{\star} = \mathsf{WF}_{\infty}(\mathbf{Q}_{\infty}^{\star})$, since WF_{∞} is continuous on $\mathbf{Q} \in \mathcal{Q}$ (cf. Lemma 14) and the set $\mathcal{Q}$ is closed]. This completes the proof. ■

APPENDIX C
PROOF OF THEOREM 8

Before proving the theorem, we introduce the following intermediate result.

Proposition 18: Given $\mathbb{S}_{+}^{n_T\times n_T} \ni \mathbf{R}_H = \mathbf{V}_H \mathbf{\Lambda}_H \mathbf{V}_H^H$, with $r_{R_H} = \mathsf{rank}(\mathbf{R}_H)$, the solution to the following optimization problem

$$\begin{aligned} \underset{\mathbf{Q}\succeq\mathbf{0}}{\text{maximize}} \quad & \log\det\left(\mathbf{I} + \mathbf{R}_H\mathbf{Q}\right) \\ \text{subject to} \quad & \mathsf{Tr}(\mathbf{Q}) \leq P_T, \quad \lambda_{\max}(\mathbf{Q}) \leq P^{\text{peak}} \end{aligned} \tag{71}$$

with $P_T \leq P^{\text{peak}} r_{R_H}$, is unique and it is given by

$$\mathbf{Q}^{\star} = \mathbf{V}_{H,1} \left[\mu \mathbf{I}_{r_{R_H}} - \mathbf{\Lambda}_{H,1}^{-1}\right]_0^{P^{\text{peak}}} \mathbf{V}_{H,1}^H \tag{72}$$

where $\mathbf{V}_{H,1} \in \mathbb{C}^{n_T\times r_{R_H}}$ is the semi-unitary matrix of the eigenvectors of matrix $\mathbf{R}_H$ corresponding to the r_{R_H} positive eigenvalues in the diagonal matrix $\mathbf{\Lambda}_{H,1}$, and $\mu > 0$ satisfies $\mathrm{Tr}\left(\left[\mu \mathbf{I}_{r_{R_H}} - \mathbf{\Lambda}_{H,1}^{-1}\right]_0^{P^{\text{peak}}}\right) = P_T$.

Proof: Invoking the well-known diagonality result of the capacity-achieving solution to the single user vector Gaussian channel [29] and using the eigendecomposition $\mathbf{R}_H = \mathbf{V}_H \mathbf{\Lambda}_H \mathbf{V}_H^H$, with $\mathbf{V}_H = (\mathbf{V}_{H,1}, \mathbf{V}_{H,2}) \in \mathbb{C}^{n_T\times n_T}$ and $\mathbf{\Lambda}_H = \mathrm{Diag}\left(\{\lambda_{H,k}\}_{k=1}^{n_T}\right)$ (the eigenvalues $\lambda_{H,k}$'s are arranged in decreasing order), we have

$$\log\det\left(\mathbf{I} + \mathbf{R}_H\mathbf{Q}\right) \leq \sum_{k=1}^{r_{R_H}} \log\left(1 + \lambda_{H,k}\left[\tilde{\mathbf{Q}}\right]_{kk}\right) \tag{73}$$

where $\tilde{\mathbf{Q}} \triangleq \mathbf{V}_H^H \mathbf{Q} \mathbf{V}_H$ and the inequality follows from the Hadamard's inequality [29]. Observe that: i) the equality in (73) is reached if and only if $\tilde{\mathbf{Q}}$ is diagonal; ii) the power constraint $\mathsf{Tr}(\tilde{\mathbf{Q}}) = \mathsf{Tr}(\mathbf{Q}) \leq P_T$ depends only on the diagonal elements of $\tilde{\mathbf{Q}}$; and iii) $\lambda_{\max}(\tilde{\mathbf{Q}}) \geq \lambda_{\max}(\mathrm{Diag}(\tilde{\mathbf{Q}}))$. It follows that the optimal solution to (71) must be diagonal, i.e., $\tilde{\mathbf{Q}} = \mathrm{Diag}(\mathbf{p})$, with $p(k) = 0$ for all $k > r_{R_H}$. In fact, for any nondiagonal $\tilde{\mathbf{Q}}$ with possibly $|[\tilde{\mathbf{Q}}]_{ij}| \neq 0$ for $i, j > r_{R_H}$—implying $[\tilde{\mathbf{Q}}]_{ii} > 0$ and $[\tilde{\mathbf{Q}}]_{jj} > 0$ (recall that $\tilde{\mathbf{Q}} \succeq \mathbf{0}$) [28, p. 398]—one can always reach the upper bound in (73) or increase its value (if $\tilde{\mathbf{Q}}$ already satisfies $|[\tilde{\mathbf{Q}}]_{ij}| = 0$ for $i, j \leq r_{R_H}$ and $i \neq j$) by using instead a diagonal matrix $\tilde{\mathbf{Q}}' \succeq \mathbf{0}$ with $|[\tilde{\mathbf{Q}}']_{ij}| = 0$ for $i, j > r_{R_H}$, and $[\tilde{\mathbf{Q}}']_{ii} > [\tilde{\mathbf{Q}}']_{ii}$ for at least one $i \leq r_{R_H}$ (recall that $P_T \leq P^{\text{peak}} r_{R_H}$), without affecting the trace constraint and still satisfying $\lambda_{\max}(\tilde{\mathbf{Q}}') \leq P^{\text{peak}}$. This proves the optimality of $\mathbf{Q}^\star$ as given in (72). Introducing (72) in (71), the optimization problem reduces to the following scalar power allocation problem:

$$\begin{aligned} &\underset{\mathbf{p}}{\text{maximize}} && \sum_{k=1}^{r_{R_H}} \log\left(1 + \lambda_{H,k} p(k)\right) \\ &\text{subject to} && \sum_{k=1}^{r_{R_H}} p(k) \leq P_T, \\ & && 0 \leq p(k) \leq P^{\text{peak}}, k \in \{1, \ldots, r_{R_H}\}. \end{aligned} \tag{74}$$

The optimal power allocation $\mathbf{p}^\star$—the solution to (74)—as given in (72), follows easily from the KKT optimality conditions of the convex problem (see, e.g., [18, Lemma 1]). The uniqueness of $\mathbf{p}^\star$ follows from the strict concavity of the objective function in (74). This also guarantees the uniqueness of $\mathbf{Q}^\star$ in (72). ■

Proof of Theorem 8: Under $r_{G_q} = n_{T_q}$, for all $q \in \Omega$, game $\mathcal{G}_{\text{soft}}$ admits at least one NE, since it satisfies [22, Theorem 6] (or [44]). We prove now (38), under $\mathrm{rank}(\overline{\mathbf{H}}_{qq}) = \mathrm{rank}(\mathbf{H}_{qq}\mathbf{G}_q^{\sharp H}\overline{\mathbf{U}}_q^\perp) = r_{\overline{U}_q^\perp}$, for all $q \in \Omega$. To this end, we rewrite $\mathcal{G}_{\text{soft}}$ in (33) in a more convenient form. For the sake of notation, we introduce $\overline{\mathbf{P}}_q^\perp \triangleq \mathbf{P}_{\mathcal{R}(\overline{\mathbf{U}}_q^\perp)}$ for all q. Defining, for each q, the transformation $\overline{\mathbf{Q}}_q \triangleq \mathbf{G}_q^H \mathbf{Q}_q \mathbf{G}_q$, one can rewrite $\mathcal{G}_{\text{soft}}$ in terms of $\overline{\mathbf{Q}}_q$ as

$$\begin{aligned} &\underset{\overline{\mathbf{Q}}_q \succeq \mathbf{0}}{\text{maximize}} && \log\det\Big(\mathbf{I} + \overline{\mathbf{P}}_q^\perp \mathbf{G}_q^\sharp \mathbf{H}_{qq}^H \overline{\mathbf{R}}_{-q}^{-1}(\overline{\mathbf{Q}}_{-q}) \\ & && \qquad \times \mathbf{H}_{qq}\mathbf{G}_q^{\sharp H}\overline{\mathbf{P}}_q^\perp \overline{\mathbf{Q}}_q\Big) \\ &\text{subject to} && \mathsf{Tr}(\overline{\mathbf{Q}}_q) \leq P_q^{\text{ave}}, \quad \lambda_{\max}(\overline{\mathbf{Q}}_q) \leq P_q^{\text{peak}} \\ & && \overline{\mathbf{Q}}_q = \overline{\mathbf{P}}_q^\perp \overline{\mathbf{Q}}_q \overline{\mathbf{P}}_q^\perp \end{aligned} \tag{75}$$

for all $q \in \Omega$, where $\overline{\mathbf{R}}_{-q}(\overline{\mathbf{Q}}_{-q}) \triangleq \mathbf{R}_{n_q} + \sum_{r\neq q} \mathbf{H}_{rq}\mathbf{G}_r^{\sharp H}\overline{\mathbf{P}}_r^\perp \overline{\mathbf{Q}}_r \mathbf{H}_{rq}^H \mathbf{G}_r^\sharp \overline{\mathbf{P}}_r^\perp$. Observe now that, because of the null constraint, any solution $\overline{\mathbf{Q}}_q^\star$ to (75) will satisfy $\mathrm{rank}(\overline{\mathbf{Q}}_q^\star) \leq r_{\overline{U}_q^\perp}$, *whatever* the strategies $\overline{\mathbf{Q}}_{-q}$ of the others are. This implies that, for each users q, the trace constraint in (75) can be equivalently rewritten as $\mathsf{Tr}(\overline{\mathbf{Q}}_q) = \sum_{k=1}^{r_{\overline{U}_q^\perp}} \lambda_k(\overline{\mathbf{Q}}_q) \leq P_q^{\text{ave}}$ (the eigenvalues $\lambda_k(\overline{\mathbf{Q}}_q)$ are assumed to be arranged in decreasing order); which, together to $\lambda_{\max}(\overline{\mathbf{Q}}_q) \leq P_q^{\text{peak}}$, leads to the following equivalent reformulation of game in (75):

$$\begin{aligned} &\underset{\overline{\mathbf{Q}}_q \succeq \mathbf{0}}{\text{maximize}} && \log\det\Big(\mathbf{I} + \overline{\mathbf{P}}_q^\perp \mathbf{G}_q^\sharp \mathbf{H}_{qq}^H \overline{\mathbf{R}}_{-q}^{-1}(\overline{\mathbf{Q}}_{-q}) \\ & && \qquad \times \mathbf{H}_{qq}\mathbf{G}_q^{\sharp H}\overline{\mathbf{P}}_q^\perp \overline{\mathbf{Q}}_q\Big) \\ &\text{subject to} && \mathsf{Tr}(\overline{\mathbf{Q}}_q) \leq \overline{P}_q^{\text{ave}}, \quad \lambda_{\max}(\overline{\mathbf{Q}}_q) \leq P_q^{\text{peak}} \\ & && \overline{\mathbf{Q}}_q = \overline{\mathbf{P}}_q^\perp \overline{\mathbf{Q}}_q \overline{\mathbf{P}}_q^\perp \end{aligned} \tag{76}$$

for all $q \in \Omega$, where $\overline{P}_q^{\text{ave}} \triangleq \min\left(P_q^{\text{ave}}, r_{\overline{U}_q^\perp} P_q^{\text{peak}}\right)$. Invoking Proposition 18, game in (76) can be further simplified removing the null constraint:

$$\begin{aligned} &\underset{\overline{\mathbf{Q}}_q \succeq \mathbf{0}}{\text{maximize}} && \log\det\Big(\mathbf{I} + \overline{\mathbf{P}}_q^\perp \mathbf{G}_q^\sharp \mathbf{H}_{qq}^H \overline{\mathbf{R}}_{-q}^{-1}(\overline{\mathbf{Q}}_{-q}) \\ & && \qquad \times \mathbf{H}_{qq}\mathbf{G}_q^{\sharp H}\overline{\mathbf{P}}_q^\perp \overline{\mathbf{Q}}_q\Big) \\ &\text{subject to} && \mathsf{Tr}(\overline{\mathbf{Q}}_q) \leq \overline{P}_q^{\text{ave}}, \quad \lambda_{\max}(\overline{\mathbf{Q}}_q) \leq P_q^{\text{peak}} \end{aligned} \tag{77}$$

for all $q \in \Omega$. In fact, according to (72) in Proposition 18, any optimal solution $\overline{\mathbf{Q}}_q^\star$ will satisfy $\mathcal{R}(\overline{\mathbf{Q}}_q^\star) \perp \mathcal{R}(\overline{\mathbf{U}}_q)$ [recall that $\mathrm{rank}(\mathbf{H}_{qq}\mathbf{G}_q^{\sharp H}\overline{\mathbf{U}}_q^\perp) = r_{\overline{U}_q^\perp}$], so that the null constraint $\overline{\mathbf{Q}}_q = \overline{\mathbf{P}}_q^\perp \overline{\mathbf{Q}}_q \overline{\mathbf{P}}_q^\perp$ in (76) is redundant.

Given the game in (77), all the Nash equilibria satisfy the following MIMO waterfilling-like equation:

$$\begin{aligned} \overline{\mathbf{Q}}_q^\star &= \overline{\mathsf{WF}}_q\left(\overline{\mathbf{P}}_q^\perp \mathbf{G}_q^\sharp \mathbf{H}_{qq}^H \overline{\mathbf{R}}_{-q}^{-1}(\overline{\mathbf{Q}}_{-q}^\star)\mathbf{H}_{qq}\mathbf{G}_q^{\sharp H}\overline{\mathbf{P}}_q^\perp\right) && (78) \\ &= \overline{\mathbf{U}}_q^\perp \overline{\mathsf{WF}}_q\left(\overline{\mathbf{U}}_q^{\perp H}\mathbf{G}_q^\sharp \mathbf{H}_{qq}^H \overline{\mathbf{R}}_{-q}^{-1}(\overline{\mathbf{Q}}_{-q}^\star)\mathbf{H}_{qq}\mathbf{G}_q^{\sharp H}\overline{\mathbf{U}}_q^\perp\right)\overline{\mathbf{U}}_q^{\perp H} && (79) \end{aligned}$$

for all $q \in \Omega$, where with a slight abuse of notation we used the same symbol $\overline{\mathsf{WF}}_q(\cdot)$ to denote the MIMO waterfilling operator in (37) applied to channel matrices with different dimensions, namely $\overline{\mathbf{P}}_q^\perp \mathbf{G}_q^\sharp \mathbf{H}_{qq}^H\ \overline{\mathbf{R}}_{-q}^{-1}(\overline{\mathbf{Q}}_{-q}^\star)\mathbf{H}_{qq}\mathbf{G}_q^{\sharp H}\overline{\mathbf{P}}_q^\perp$ and $\overline{\mathbf{U}}_q^{\perp H}\mathbf{G}_q^\sharp \mathbf{H}_{qq}^H \overline{\mathbf{R}}_{-q}^{-1}(\overline{\mathbf{Q}}_{-q}^\star)\mathbf{H}_{qq}\mathbf{G}_q^{\sharp H}\overline{\mathbf{U}}_q^{\perp H}$.

The structure of the Nash equilibria of game $\mathcal{G}_{\text{soft}}$ in (33) as given in (38) follows readily from (78) (recall that $\overline{\mathbf{Q}}_q \triangleq \mathbf{G}_q^H \mathbf{Q}_q \mathbf{G}_q$). ■

APPENDIX D
PROOF OF LEMMA 9

The proof of Lemma 9 is based on the following result.

Proposition 19: Let $\mathbb{S}_+^{n_T \times n_T} \ni \mathbf{X}_0 = \mathbf{V}_{0,1}\mathbf{D}_{0,1}\mathbf{V}_{0,1}^H$, where $\mathbf{V}_{0,1} \in \mathbb{C}^{n_T \times r_{X_0}}$ is semi-unitary and $\mathbf{D}_{0,1} = \mathrm{Diag}(\{d_{0,k}\}_{k=1}^{r_{X_0}}) > \mathbf{0}$, with $r_{X_0} = \mathsf{rank}(\mathbf{X}_0)$. Let $\overline{\mathcal{Q}}$ be the convex set defined as

$$\overline{\mathcal{Q}} \triangleq \left\{\mathbf{X} \in \mathbb{S}_+^{n_T \times n_T} : \mathrm{Tr}\{\mathbf{X}\} = P_T,\ \lambda_{\max}(\mathbf{X}) \leq P^{\text{peak}}\right\} \tag{80}$$

with $0 < P_T \leq P^{\text{peak}} r_{X_0}$. Given $\mathbf{Y}_0 \triangleq \mathbf{X}_0^\sharp + c\,\mathbf{P}_{\mathcal{N}(\mathbf{X}_0)}$, where c is any fixed positive constant, the matrix projection of $-\mathbf{Y}_0$ w.r.t. the Frobenius norm onto the convex set $\overline{\mathcal{Q}}$ is by definition given by:

$$[-\mathbf{Y}_0]_{\mathcal{Q}} = \underset{\mathbf{Q}\in\mathcal{Q}}{\mathrm{argmin}} \quad \|\mathbf{Q} + \mathbf{Y}_0\|_F^2. \tag{81}$$

For any fixed $c \geq \max_{k\in\{1,\ldots,r_{X_0}\}} d_{0,k}^{-1}+\max\{P_T, P^{\text{peak}}\}$, the unique solution to (81) assumes the following form:

$$[-\mathbf{Y}_0]_{\mathcal{Q}} = \mathbf{V}_{0,1}\left[\mu\mathbf{I}_{r_{X_0}} - \mathbf{D}_{0,1}^{-1}\right]_0^{P^{\text{peak}}}\mathbf{V}_{0,1}^H \tag{82}$$

where $\mu > 0$ satisfies the power constraint $\text{Tr}\left([\mu\mathbf{I}_{r_{X_0}} - \mathbf{D}_{0,1}^{-1}]_0^{P^{\text{peak}}}\right) = P_T$.

Proof: Using $\mathbf{X}_0 = \mathbf{V}_{0,1}\mathbf{D}_{0,1}\mathbf{V}_{0,1}^H$ and introducing the unitary $n_T \times n_T$ matrix $\mathbf{V}_0 \triangleq (\mathbf{V}_{0,1}, \mathbf{V}_{0,2})$, with $\mathbf{V}_{0,2} \in \mathbb{C}^{n_T\times r_{V_{0,2}}}$ and $r_{V_{0,2}} = n_T - r_{X_0}$ [note that $\mathcal{R}(\mathbf{V}_{0,2}) = \mathcal{N}(\mathbf{X}_0)$], the objective function in (81) can be rewritten as

$$\|\mathbf{Q}+\mathbf{Y}_0\|_F^2 = \left\|\widetilde{\mathbf{Q}} + \widetilde{\mathbf{D}}_0^{-1}\right\|_F^2 \tag{83}$$

where $\widetilde{\mathbf{Q}} \triangleq \mathbf{V}_0^H\mathbf{Q}\mathbf{V}_0$, $\widetilde{\mathbf{D}}_0^{-1} \triangleq \text{bdiag}\left(\mathbf{D}_{0,1}^{-1}, c\mathbf{I}_{r_{V_{0,2}}}\right)$, and we used the unitary invariance of the Frobenius norm [28]. Observe that: i) the objective function (83) satisfies $\left\|\widetilde{\mathbf{Q}} + \widetilde{\mathbf{D}}_0^{-1}\right\|_F^2 \geq \left\|\text{Diag}(\widetilde{\mathbf{Q}}) + \widetilde{\mathbf{D}}_0^{-1}\right\|_F^2$, with equality if and only if $\widetilde{\mathbf{Q}}$ is diagonal; ii) the power constraint $\text{Tr}\{\mathbf{Q}\} = \text{Tr}\{\widetilde{\mathbf{Q}}\} = P_T$ depends only on the diagonal elements of $\widetilde{\mathbf{Q}}$; iii) $\lambda_{\max}(\widetilde{\mathbf{Q}}) \geq \lambda_{\max}(\text{Diag}(\widetilde{\mathbf{Q}}))$; and iv) the optimal solution to (81) is unique (implied from the strict convexity of the Frobenius norm). It follows that the optimal solution $\widetilde{\mathbf{Q}}^\star$ must be diagonal, i.e., $\widetilde{\mathbf{Q}}^\star = \text{Diag}(\mathbf{p}^\star)$, which proves the optimality of the structure in (82).

We derive now the optimal power allocation $\mathbf{p}^\star = \text{Diag}(\widetilde{\mathbf{Q}}^\star)$. Introducing the optimal structure $\mathbf{Q} = \mathbf{V}_{0,1}\text{Diag}(\mathbf{p})\mathbf{V}_{0,1}^H$ in (81), with $\mathbf{p} = (p_k)_{k=1}^{n_T}$, and denoting each $\tilde{d}_{0,k} = [\widetilde{\mathbf{D}}_0]_{kk}$, the original matrix-value problem reduces to the following vector convex optimization problem:

$$\begin{aligned} &\underset{\mathbf{p}}{\text{minimize}} && \frac{1}{2}\sum_{k=1}^{n_T}\left(p_k + \tilde{d}_{0,k}^{-1}\right)^2 \\ &\text{subject to} && \sum_{k=1}^{n_T} p_k = P_T, \\ &&& 0 \leq p_k \leq P^{\text{peak}}, \quad \forall\, k \in \{1,\ldots,n_T\}. \end{aligned} \tag{84}$$

whose (unique) solution comes from the KKT optimality conditions (note that constraint qualifications are implied by the polyhedral structure of the feasible set):

$$\begin{aligned} &0 \leq\ p_k \perp p_k + \tilde{d}_{0,k}^{-1} + \lambda_k - \mu \geq 0, && \forall\, k \in \{1,\ldots,n_T\}, \\ &0 \leq\ \lambda_k \perp P^{\text{peak}} - p_k \geq 0, && \forall\, k \in \{1,\ldots,n_T\}, \\ &\mu \text{ free} \quad \sum_{k=1}^{n_T} p_k = P_T. \end{aligned} \tag{85}$$

If $P_T = r_{X_0}P^{\text{peak}}$ and $c \geq \max_{k\in\{1,\ldots,r_{X_0}\}} d_{0,k}^{-1} + P^{\text{peak}}$, then it is straightforward to check by inspection that the solution $\mathbf{p}^\star$ to (85) is (recall that $\tilde{d}_{0,k}^{-1} = c$ for $k > r_{X_0}$)

$$p_k^\star = \begin{cases} P^{\text{peak}}, & \text{if } k \in \{1,\ldots,r_{X_0}\} \\ 0, & \text{otherwise} \end{cases} \tag{86}$$

with the multipliers $\mu^\star$ and $\lambda_k^\star$'s satisfying

$$\begin{aligned} \mu^\star &= \left(P^{\text{peak}} + \max_{k\in\{1,\ldots,r_{X_0}\}} d_{0,k}^{-1}\right) \\ \lambda_k^\star &= \begin{cases} \max_{k\in\{1,\ldots,r_{X_0}\}} d_{0,k}^{-1} - d_{0,k}^{-1}, & \text{if } k \in \{1,\ldots,r_{X_0}\} \\ 0, & \text{otherwise.} \end{cases} \end{aligned} \tag{87}$$

Assume now that $P_T < r_{X_0}P^{\text{peak}}$. Then the solution $\mathbf{p}^\star$ can be written as $p_k^\star = [\mu^\star - \tilde{d}_{0,k}^{-1}]_0^{P^{\text{peak}}}$ for all $k \in \{1,\ldots,n_T\}$, with $\mu^\star > 0$ such that the power constraint $\sum_{k=1}^{n_T} p_k^\star = P_T$ holds true and each $\lambda_k^\star$ satisfies the complementary conditions in (85). It is not difficult to see that if $c \geq \max_{k\in\{1,\ldots,r_{X_0}\}} d_{0,k}^{-1} + P_T$, then $p_k^\star = 0$ for all $k > r_{X_0}$. Therefore, choosing $c \geq \max_{k\in\{1,\ldots,r_{X_0}\}} d_{0,k}^{-1} + \max\{P_T, P^{\text{peak}}\}$ leads to the desired structure of the optimal solution, as given in (82) ■

Proof of Theorem 9: It follows from Proposition 19 that, for each player $q \in \Omega$ of $\mathcal{G}_{\text{soft}}$ and $\mathbf{Q}_{-q} \in \overline{\mathcal{Q}}_{-q}$, the waterfilling mapping $\overline{\mathsf{WF}}_q(\mathbf{Q}_{-q}) = \overline{\mathsf{WF}}_q\left(\overline{\mathbf{H}}_{qq}^H\mathbf{R}_{-q}^{-1}(\mathbf{Q}_{-q})\overline{\mathbf{H}}_{qq}\right)$ defined in (37) can be equivalently written as

$$\overline{\mathsf{WF}}_q(\mathbf{Q}_{-q}) = \left[-\left(\left(\overline{\mathbf{H}}_{qq}^H\mathbf{R}_{-q}^{-1}(\mathbf{Q}_{-q}^\star)\overline{\mathbf{H}}_{qq}\right)^\sharp + c_q\mathbf{P}_{\mathcal{N}(\overline{\mathbf{H}}_{qq})}\right)\right]_{\overline{\mathcal{Q}}_q} \tag{88}$$

for any $c_q \geq c_q(\mathbf{Q}_{-q}) = \max_{k\in\{1,\ldots,r_{X_0}\}} \overline{\lambda}_{q,k}^{-1}(\mathbf{Q}_{-q}) + \max\{\overline{P}_q^{\text{ave}}, P_q^{\text{peak}}\}$, where $\overline{\lambda}_{q,k}^{-1}(\mathbf{Q}_{-q})$ are the eigenvalues of matrix $\overline{\mathbf{H}}_{qq}^H\mathbf{R}_{-q}^{-1}(\mathbf{Q}_{-q})\overline{\mathbf{H}}_{qq}$. A (finite) upper bound of $c_q(\mathbf{Q}_{-q})$, independent of $\mathbf{Q}_{-q}$, is, e.g.,

$$c_q(\mathbf{Q}_{-q}) < \max\{\overline{P}_q^{\text{ave}}, P_q^{\text{peak}}\} + \frac{\rho(\mathbf{R}_{n_q}) + Q\overline{P}_{\max}^{\text{ave}}\max_r \rho\left(\mathbf{H}_{rq}^H\mathbf{H}_{rq}\right)}{\lambda_{\min}\left(\overline{\mathbf{H}}_{qq}^H\overline{\mathbf{H}}_{qq}\right)} < \infty \tag{89}$$

where $\overline{P}_{\max}^{\text{ave}} \triangleq \max_q \overline{P}_q^{\text{ave}}$. It follows from (88) that the fixed-points of (38) are then the solution to (82), which completes the proof. ■

REFERENCES

[1] "Report of the Spectrum Efficiency Working Group," FCC Spectrum Policy Task Force, 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 Mobile Multimedia Communcations (MoMuC)*, 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.

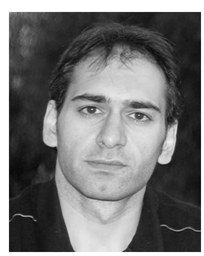
[5] A. Goldsmith, S. A. Jafar, I. Maric, and S. Srinivasa, "Breaking spectrum gridlock with cognitive radios: An information theoretic perspective," *Proc. IEEE*, vol. 97, pp. 894–914, May 2009.

[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] Z.-Q. Luo and S. Zhang, "Spectrum management: Complexity and duality," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 57–72, Feb. 2008.

[8] 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.

[9] 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.
[10] S. Ye and R. S. Blum, "Optimized signaling for MIMO interference systems with feedback," *IEEE Trans. Signal Process.*, vol. 51, no. 11, pp. 2839–2848, Nov. 2003.
[11] E. G. Larsson, E. Jorswieck, J. Lindblom, and R. Mochaourab, "Game theory and the flat fading Gaussian interference channel," *IEEE Signal Process. Mag.*, vol. 26, no. 5, pp. 18–27, Sep. 2009.
[12] A. Leshem and E. Zehavi, "Game theory and the frequency selective interference channel—A tutorial," *IEEE Signal Process. Mag.*, vol. 26, no. 5, pp. 28–40, Sep. 2009.
[13] D. Fudenberg and J. Tirole, *Game Theory*. Cambridge, MA: MIT Press, 1991.
[14] M. J. Osborne and A. Rubinstein, *A Course in Game Theory*. Cambridge, MA: MIT Press, 2004.
[15] 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, Jun. 2002.
[16] 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.
[17] 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.
[18] 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.
[19] 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.
[20] 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.
[21] G. Arslan, M. F. Demirkol, and Y. Song, "Equilibrium efficiency improvement in MIMO interference systems: A decentralized stream control approach," *IEEE Trans. Wireless Commun.*, vol. 6, no. 8, pp. 2984–2993, Aug. 2007.
[22] G. Scutari, D. P. Palomar, and S. Barbarossa, "Competitive design of multiuser MIMO systems based on game theory: A unified view," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 7, pp. 1089–1103, Sep. 2008.
[23] E. Larsson and E. Jorswieck, "Competition versus collaboration on the MISO interference channel," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 7, pp. 1059–1069, Sep. 2008.
[24] G. Scutari, D. P. Palomar, and S. Barbarossa, "The MIMO iterative waterfilling algorithm," *IEEE Trans. Signal Process.*, vol. 57, no. 5, pp. 1917–1935, May 2009.
[25] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*, 2nd ed. Singapore: Athena Scientific Press, 1989.
[26] 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.
[27] S. L. Campbell and C. D. Meyer, *Generalized Inverses of Linear Transformations*. New York: Dover, 1991.
[28] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1985.
[29] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.
[30] N. Devroye, P. Mitran, and V. Tarokh, "Limits on communications in a cognitive radio channel," *IEEE Commun. Mag.*, vol. 44, no. 6, pp. 44–49, Jun. 2006.
[31] Working Group on Wireless Regional Area Networks [Online]. Available: http://www.ieee802.org/22/
[32] C. R. Stevenson, G. Chouinard, Z. Lei, W. Hu, and S. J. Shellhammer, "IEEE 802.22: The first cognitive radio wireless regional area network standard," *IEEE Commun. Mag.*, vol. 47, no. 1, pp. 130–138, Jan. 2009.
[33] N. Devroye, M. Vu, and V. Tarokh, "Cognitive radio networks: Highlights of information theoretic limits, models, and design," *IEEE Signal Process. Mag.*, vol. 25, no. 6, pp. 12–23, Nov. 2008.
[34] B. Wild and K. Ramchandran, "Detecting primary receivers for cognitive radio applications," in *Proc. IEEE 2005 Symp. New Frontiers Dynamic Spectrum Access Networks (DYSPAN 2005)*, Baltimore, MD, pp. 124–130.
[35] Z. Quan, S. Cui, H. V. Poor, and A. H. Sayed, "Collaborative wideband sensing for cognitive radios: An overview of challenges and solutions," *IEEE Signal Process. Mag.*, vol. 25, no. 6, pp. 60–73, Nov. 2008.
[36] T. Yücek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *IEEE Commun. Surveys Tuts.*, vol. 11, no. 1, pp. 116–130, 2009.
[37] R. W. Cottle, J.-S. Pang, and R. E. Stone, *The Linear Complementarity Problem*. Cambridge, U.K.: Cambridge Academic Press, 1992.
[38] G. Scutari, D. P. Palomar, and S. Barbarossa, "Competitive optimization of cognitive radio MIMO systems via game theory," in *Convex Optimization in Signal Processing and Communications*, D. P. Palomar and Y. C. Eldar, Eds. London, U.K.: Cambridge Univ. Press, 2009.
[39] 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.
[40] 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.
[41] Y. Rong and Y. Hua, "Optimal power schedule for distributed MIMO links," *IEEE Trans. Wireless Commun.*, vol. 7, no. 8, pp. 2896–2900, Aug. 2008.
[42] M. Nokleby, A. L. Swindlehurst, Y. Rong, and Y. Hua, "Cooperative power scheduling for wireless MIMO networks," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Washington, DC, Nov. 26–30, 2007, pp. 2982–2986.
[43] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed. Singapore: Athena Scientific, 1999.
[44] J. Rosen, "Existence and uniqueness of equilibrium points for concave n-person games," *Econometrica*, vol. 33, no. 3, pp. 520–534, Jul. 1965.
[45] A. Galantai, *Projectors and Projection Methods*. Norwell, MA: Kluwer Academic, 2003.
[46] S. Zlobec, *Stable Parametric Programming*. Norwell, MA: Kluwer Academic, 2001.
[47] W. Rudin, *Principles of Mathematical Analysis*, 3rd ed. New York: McGraw-Hill, 1976.



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 at the Department of Electrical Engineering and Computer Sciences, University of California at Berkeley; during 2007 and 2008, he was a Postdoctoral Fellow at the Department of Electronic and Computer Engineering at the Hong Kong University of Science and Technology, Hong Kong; during 2009 he was a Research Associate in the Department of Electronic and Computer Engineering at the Hong Kong University of Science and Technology and a Postdoctoral Researcher at the INFOCOM Department, University of Rome, "La Sapienza," Rome, Italy. He is currently a Research Associate in the Department of Industrial and Enterprise Systems Engineering at University of Illinois at Urbana-Champain. 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 in the Department of Electronic and Computer Engineering at the Hong Kong University of Science and Technology (HKUST), Hong Kong. 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, Spain; 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 the Technical University of Catalonia (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.