Optimal Linear Precoding Strategies for Wideband Non-Cooperative Systems Based on Game Theory—Part II: Algorithms

Gesualdo Scutari, *Member, IEEE*, Daniel P. Palomar, *Member, IEEE*, and Sergio Barbarossa, *Member, IEEE*

***Abstract*—In this two-part paper, we address the problem of finding the optimal precoding/multiplexing scheme for a set of noncooperative links sharing the same physical resources, e.g., time and bandwidth. We consider two alternative optimization problems: P.1) the maximization of mutual information on each link, given constraints on the transmit power and spectral mask; and P.2) the maximization of the transmission rate on each link, using finite-order constellations, under the same constraints as in P.1, plus a constraint on the maximum average error probability on each link. Aiming at finding decentralized strategies, we adopted as optimality criterion the achievement of a Nash equilibrium and thus we formulated both problems P.1 and P.2 as strategic noncooperative (matrix-valued) games. In Part I of this two-part paper, after deriving the optimal structure of the linear transceivers for both games, we provided a unified set of sufficient conditions that guarantee the uniqueness of the Nash equilibrium. In this Part II of the paper, we focus on the achievement of the equilibrium and propose alternative distributed iterative algorithms that solve both games. Specifically, the new proposed algorithms are the following: 1) the *sequential* and *simultaneous* iterative waterfilling-based algorithms, incorporating spectral mask constraints and 2) the *sequential* and *simultaneous* gradient-projection-based algorithms, establishing an interesting link with variational inequality problems. Our main contribution is to provide sufficient conditions for the *global* convergence of all the proposed algorithms which, although derived under stronger constraints, incorporating for example spectral mask constraints, have a broader validity than the convergence conditions known in the current literature for the sequential iterative waterfilling algorithm.**

***Index Terms*—Competitive optimality, distributed algorithms, game theory, iterative waterfilling.**

I. Introduction and Motivation

THE goal of this two-part paper is to find the optimal precoding/multiplexing schemes for a set of noncooperative links sharing the same physical resources, e.g., time and bandwidth. Two alternative optimization problems are considered [1]: P.1) the maximization of mutual information on each link, given constraints on the transmit power and spectral emission mask, imposed by radio spectrum regulatory bodies and P.2) the maximization of the transmission rate on each link, using finite-order constellations, under the same constraints as in P.1, plus a constraint on the maximum average error probability on each link. We focus on decentralized strategies to avoid coordination among the separated links and the heavy signaling required by a global controller that would need to collect all relevant information from all the users. The search for decentralized solutions motivated our formulation within the convenient framework of game theory. We thus adopt as optimality criterion the achievement of a Nash equilibrium (NE) [2], and we cast both optimization problems P.1 and P.2 as strategic noncooperative (matrix-valued) games [1], where the goal of each user is to optimize its own precoding/multiplexing matrix. In Part I of this two-part paper [1, Theorem 1], we proved that there is no performance loss in reducing both original matrix-valued games into a unified vector-valued power control game, where the optimal strategy of each user corresponds to finding the power allocation that maximizes its own (information) rate, treating the multiuser interference due to the other users as additive colored noise. We will refer to this power control game as *rate-maximization* game. In Part I of this paper, we proved that the solution set of the rate-maximization game is always nonempty and derived (sufficient) conditions that guarantee the uniqueness of the NE [1, Theorem 2]. In this Part II, we propose alternative algorithms that reach the Nash equilibria of the unified vector game, in a totally distributed manner.

All the distributed algorithms used to compute the Nash equilibria of a (rational [2]) strategic noncooperative game are based on a simple idea: Each player optimizes iteratively its own payoff function following a prescribed updating schedule, for example, simultaneously with the other users (i.e., according to a Jacobi scheme [3]), or sequentially (i.e., according to a Gauss–Seidel scheme [3]). Differently from the optimization of a single-user system, where the optimal transceiver structure can be obtained in a single shot (depending on the interference scenario [4]–[6]), in a competitive multiuser context like a game, it is necessary to adopt an iterative algorithm, as each user's choice affects the interference perceived by the other users. However, the competitive nature of the multiuser system does not guarantee in general the convergence of such an iterative scheme, even if the payoff function of each player is strictly concave (or strictly convex) in its own strategies and the NE is unique. This issue motivated several works in the literature

Manuscript received June 1, 2006; revised June 5, 2007. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Geert Leus. This work was supported by the SURFACE project funded by the European Community under Contract IST-4-027187-STP-SURFACE, and by the National Science Foundation of China under Grant 60702081.

G. Scutari and S. Barbarossa are with the INFOCOM Department, University of Rome "La Sapienza," 00184 Rome, Italy (e-mail: scutari@infocom.uniroma1.it; sergio@infocom.uniroma1.it).

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

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.2007.907808

1053-587X/$25.00 © 2008 IEEE

[7]–[11], [21]–[29], where alternative approaches have been proposed to study the convergence of iterative algorithms in strategic noncooperative games.

A traditional approach comes from classical *scalar* power control problems in *flat-fading* CDMA (or TDMA/FDMA) wireless networks (either cellular or ad-hoc) [7]–[16],[1] where each user has only one variable to optimize: its transmit power. This kind of problems can be elegantly recast as convex optimization problems (see, e.g., [18]–[20]) or as the so-called "standard" problems (in the sense of [9]–[11]), for which distributed (either synchronous or asynchronous) algorithms along with their convergence properties are available [7]–[16], [20]. The rate-maximization game proposed in this paper is more involved, as it falls in the class of *vector* power control problems, where each player has a vector to optimize (i.e., its power allocation across frequency bins) and the best-response function of each user (the waterfilling mapping) is not a standard function (in the sense of [9]–[11]). Hence, the classical framework of [7]–[11] cannot be successfully applied to our game theoretical formulation.

A special case of the rate-maximization game proposed in this paper was studied in [21] in the absence of spectral mask constraints, where the authors formulated the vector power control problem for a digital subscriber line (DSL) system, modeled as a Gaussian frequency-selective interference channel, as a two-person strategic noncooperative game. To reach the Nash equilibria of the game, the authors proposed the *sequential* iterative waterfilling algorithm (IWFA), which is an instance of the Gauss–Seidel scheme [3]: The users maximize their own information rates sequentially (one after the other), according to a fixed updating order. Each user performs the single-user waterfilling solution given the interference generated by the others as additive (colored) noise. The most appealing features of the sequential IWFA are its low-complexity and its distributed nature. In fact, to compute the waterfilling solution, each user only needs to measure the noise-plus-interference power spectral density (PSD), without requiring specific knowledge of the power allocations and the channel transfer functions of the other users. The convergence of the sequential IWFA has been studied in a number of works [22]–[28], each time obtaining milder conditions that guarantee convergence. However, despite its appealing properties, the sequential IWFA suffers from slow convergence if the number of users in the network is large, because of the sequential updating strategy. In addition, the algorithm requires some form of central scheduling to determine the order in which users update their strategy.

The original contributions of this paper with respect to the current literature on vector games [21]–[26] are listed next. First, to compute the Nash equilibria of both games P.1 and P.2 (introduced in Part I [1]), we generalize the sequential IWFA of [21], including the spectral mask constraint and a possible memory in the updating process. Then, to overcome the potential slow convergence rate of the sequential IWFA, we propose a new iterative algorithm, called *simultaneous* IWFA. The simultaneous IWFA is an instance of the Jacobi scheme [3]: At each iteration, all users update their own strategies *simultaneously*, still according to the single-user waterfilling solution, but using the interference generated by the others in the *previous* iteration. We provide results on the convergence speed of both algorithms, showing that the simultaneous IWFA is faster than the sequential IWFA, still keeping the desired properties of the sequential IWFA, i.e., its distributed nature and low complexity. The second important contribution of the paper is to provide a unified set of sufficient conditions ensuring the *global* convergence of both algorithms. Our conditions are proved to have broader validity than those given in [21]–[25], [26] (obtained without mask constraints) and, more recently, in [27] (obtained including mask constraints) for the sequential IWFA. Moreover, they show that the range of applicability with guaranteed convergence of both sequential and simultaneous IWFAs includes scenarios where the interfering users may be rather close to each other. Finally, exploring the link between the Nash equilibria of our game theoretical formulation and the solutions to the so-called variational inequality problems [30]–[32], we propose, as alternative to the IWFAs, two novel gradient projection based iterative algorithms, namely the *sequential* and *simultaneous* iterative gradient projection algorithms (IGPAs) and provide conditions for their global convergence.

Throughout the paper, there is a common thread relating the algorithms and the derivation of their convergence conditions: The interpretation of the waterfilling operator as the Euclidean projector of a vector onto a convex set. In the single-user case, this provides an alternative perspective of the well-known waterfilling solution, that dates back to Shannon in 1949 [33]. Interestingly, in the multiuser case, this interpretation plays a key role in proving the convergence of the proposed algorithms.

The paper is organized as follows. After briefly reviewing, in Section II, the game theoretic formulation addressed in Part I of the paper [1, Theorem 2], Section III provides the interpretation of the waterfilling operator as a projector. Section IV contains the main contribution of the paper: A variety of distributed algorithms for the computation of the Nash equilibria of the game, along with their convergence properties. Finally, in Section V, some conclusions are drawn. Preliminary versions of this paper appeared in [25], [28], and [29].

II. System Model and Problem Formulation

We consider a Gaussian vector interference channel [35], composed by Q noncooperative links. Aiming at finding distributed algorithms, we focus on transmission techniques where no interference cancellation is performed and multiuser interference is treated as additive colored noise. Moreover, we consider a block transmission without loss of generality (w.l.o.g.), as it is a capacity-lossless strategy for sufficiently large block length [36]–[38]. Then, under assumptions detailed in Part I [1], the system design consists of finding the optimal transmit/receive matrix pair for each link independently of the others, according to some performance metrics. In Part I of this paper [1], we assumed as optimality criterion the achievement

[1]Note that, even though some of these papers do not contain any explicit reference to game theory, the problems therein can be naturally reformulated as a strategic noncooperative game, where the Nash equilibria are the fixed points of proper best response correspondences.

of the NE and considered the two following strategic noncooperative games:

P.1) the maximization of mutual information on each link, given constraints on the transmit power and on the spectral radiation mask;

P.2) the maximization of the transmission rate on each link, using finite-order constellations, under the same constraints as in P.1 plus a constraint on the average (uncoded) error probability.

After showing that the solution set of both games is always nonempty, in [1, Theorem 1] we proved that the optimal transmission strategy for each link leads to Gaussian signaling plus the diagonal transmission through the channel eigenmodes (i.e., the frequency subchannels), irrespective of the channel state, power budget, spectral mask constraints and interference levels. Thanks to this result, both *matrix-valued* games P.1 and P.2 can be recast, with no performance loss, as the following simpler *vector* power control game [1, Theorem 1]:

$$(\mathcal{G}): \quad \begin{array}{ll} \underset{\mathbf{p}_q}{\text{maximize}} & R_q(\mathbf{p}_q, \mathbf{p}_{-q}) \\ \text{subject to} & \mathbf{p}_q \in \mathcal{P}_q \end{array}, \quad \forall\, q \in \Omega \qquad (1)$$

where $\Omega \triangleq \{1, 2, \ldots, Q\}$ is the set of players (i.e., active links), $\mathcal{P}_q$ is the set of admissible strategies of player q[2]

$$\mathcal{P}_q \triangleq \left\{ \mathbf{p}_q \in \mathbb{R}^N : \frac{1}{N}\sum_{k=1}^{N} p_q(k) = 1, \; 0 \le p_q(k) \le p_q^{\max}(k), \quad \forall\, k \in \{1, \ldots, N\} \right\} \qquad (2)$$

where $p_q^{\max}(k) \triangleq \bar{p}_q^{\max}(k)/P_q$, with $\bar{p}_q^{\max}(k)$ denoting the maximum power that is allowed to be allocated on the kth frequency bin from the qth user, and $R_q(\mathbf{p}_q, \mathbf{p}_{-q})$ is the payoff function of player q

$$R_q(\mathbf{p}_q, \mathbf{p}_{-q}) = \frac{1}{N}\sum_{k=1}^{N} \log\left(1 + \frac{1}{\Gamma_q}\mathsf{sinr}_q(k)\right) \qquad (3)$$

with

$$\mathsf{sinr}_q(k) = \frac{P_q |\bar{H}_{qq}(k)|^2 p_q(k)/d_{qq}^{\gamma}}{\sigma_q^2 + \sum_{r \ne q} P_r |\bar{H}_{rq}(k)|^2 p_r(k)/d_{rq}^{\gamma}} \triangleq \frac{|H_{qq}(k)|^2 p_q(k)}{1 + \sum_{r \ne q} |H_{rq}(k)|^2 p_r(k)} \qquad (4)$$

where $H_{rq}(k) \triangleq \bar{H}_{rq}(k)\sqrt{P_r/(\sigma_q^2 d_{rq}^{\gamma})}$; $\bar{H}_{rq}(k)$ denotes the frequency response on the subcarrier k of the channel between source r and destination q, d_{rq} is the distance between source r and destination q, and γ is the path loss. The signal-to-noise ratio (SNR) gap Γ_q in (3) is set equal to 1 if game in P.1 is considered, whereas $\Gamma_q = (\mathcal{Q}^{-1}(P_{e,q}^{\star}/4))^2/3$ [39], if we consider P.2, where $\mathcal{Q}(\cdot)$ denotes the $\mathcal{Q}$-function [34] and $P_{e,q}^{\star}$ is the maximum tolerable (uncoded) average symbol error probability on link q.

In [1, Theorem 2], we showed that the solution set of $\mathcal{G}$ is always nonempty and coincides with the solution set of the following nonlinear fixed-point equation:

$$\mathbf{p}_q^{\star} = \mathsf{WF}_q(\mathbf{p}_1^{\star}, \ldots, \mathbf{p}_{q-1}^{\star}, \mathbf{p}_{q+1}^{\star}, \ldots, \mathbf{p}_Q^{\star}) = \mathsf{WF}_q(\mathbf{p}_{-q}^{\star}), \quad \forall\, q \in \Omega \qquad (5)$$

with the waterfilling operator $\mathsf{WF}_q(\cdot)$ defined as

$$[\mathsf{WF}_q(\mathbf{p}_{-q})]_k \triangleq \left[\mu_q - \Gamma_q \frac{1 + \sum_{r \ne q} |H_{rq}(k)|^2 p_r(k)}{|H_{qq}(k)|^2}\right]_0^{p_q^{\max}(k)} \qquad (6)$$

with $k \in \{1, \ldots, N\}$, where $[x]_a^b$ denotes the Euclidean projection of x onto the interval $[a, b]$.[3] The water level μ_q is chosen to satisfy the power constraint $(1/N)\sum_{k=1}^{N} p_q^{\star}(k) = 1$.

Observe that system (5) contains, as special cases, the solutions to power control games already studied in the literature [21]–[26], when all the players are assumed to transmit with the same power and no spectral mask constraints are imposed (i.e., when $p_q^{\max}(k) = +\infty$, $\forall q$, $\forall k$). In this case, the Nash equilibria of game $\mathcal{G}$ are given by the classical simultaneous waterfilling solutions [21]–[26], where $\mathsf{WF}_q(\cdot)$ in (5) is still obtained from (6) simply setting $p_q^{\max}(k) = +\infty$, $\forall q$, $\forall k$. However, in the presence of spectral mask constraints, the results of [21]–[26] cannot be applied to system (5). In Part I of this paper [1, Theorem 2], we studied system (5) and provided sufficient conditions for the uniqueness of the solution. The problem we address here is how to reach solutions to (5) (leading to the Nash equilibria of $\mathcal{G}$) by means of totally distributed algorithms.

III. Waterfilling Operator as a Projector

In this section, we provide an interpretation of the waterfilling operator as a proper Euclidean projector. This interpretation will be instrumental to prove the convergence properties of some of the algorithms proposed in the subsequent sections.

A. A New Look at the Single-User Waterfilling Solution

Consider a parallel additive colored Gaussian noise channel composed of N subchannels with coefficients $\{H(k)\}$, subject to some spectral mask constraints $\{p^{\max}(k)\}_{k=1}^{N}$ and to a global average transmit power constraint across the subchannels. It is well known that the capacity-achieving solution for this channel is obtained using independent Gaussian signaling across the subchannels with the following waterfilling power allocation [40]

$$p^{\star}(k) = \left[\mu - \frac{\sigma_k^2}{|H(k)|^2}\right]_0^{p^{\max}(k)}, \quad k \in \{1, \ldots, N\} \qquad (7)$$

[2] In order to avoid the trivial solution $p_q^{\star}(k) = p_q^{\max}(k)$ for all $k \in \{1, \ldots, N\}$, $(1/N)\sum_{k=1}^{N} p_q^{\max}(k) > 1$ is assumed for all $q \in \Omega$. Furthermore, in the feasible strategy set of each player, we can replace, w.l.o.g., the original *inequality* power constraint $(1/N)\sum_{k=1}^{N} p_q(k) \le 1$, with equality, since, at the optimum, this constraint must be satisfied with equality from all users.

[3] The Euclidean projection $[x]_a^b$, with $a \le b$, is defined as follows: $[x]_a^b = a$, if $x \le a$, $[x]_a^b = x$, if $a < x < b$, and $[x]_a^b = b$, if $x \ge b$.

where σ_k^2 denotes the noise variance on the kth subchannel, and $p^\star(k)$ is the optimal power allocation over the kth subchannel. The water-level μ in (7) is chosen in order to satisfy the power constraint $(1/N)\sum_{k=1}^{N} p^\star(k) = 1$ and can be computed efficiently using, e.g., the framework proposed in [50] .

We show now that, interesting ly, the solution in (7) can be interpreted as the Euclidean projection of the vector **-insr**, defined as

$$\mathbf{insr} \triangleq \left[\sigma_1^2/|H(1)|^2, \ldots, \sigma_N^2/|H(N)|^2\right]^T \quad (8)$$

onto the simplex

$$\mathcal{S} \triangleq \left\{ \mathbf{x} \in \mathbb{R}^N : \frac{1}{N}\sum_{k=1}^{N} x_k = 1, \right. \\ \left. 0 \le x_k \le p^{\max}(k), \quad \forall\, k \in \{1, \ldots, N\} \right\}. \quad (9)$$

Lemma 1: The Euclidean projection of the N-dimensional real nonpositive vector $-\mathbf{x}_0 \triangleq -[x_{0,1}, \ldots, x_{0,N}]^T$ onto the simplex $\mathcal{S}$ defined in (9), denoted by $[-\mathbf{x}_0]_{\mathcal{S}}$, is by definition the solution to the following convex optimization problem:

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & \|\mathbf{x} - (-\mathbf{x}_0)\|_2^2 \\ \text{subject to} \quad & 0 \le x_k \le p^{\max}(k), \quad \forall\, k \in \{1, \ldots, N\} \\ & \frac{1}{N}\sum_{k=1}^{N} x_k = 1. \end{aligned} \quad (10)$$

and assumes the following form:

$$x_k^\star = [\mu - x_{0,k}]_0^{p^{\max}(k)}, \quad k \in \{1, \ldots, N\} \quad (11)$$

where $\mu > 0$ is chosen in order to satisfy the constraint $(1/N)\sum_{k=1}^{N} x_k^\star = 1$.

Proof: See Appendix A. ■

Lemma 1 is an extension of [46, Lemma 1] to the case where interval bounds $[0, p^{\max}(k)]$ are included in the optimization. But what is important to remark about Lemma 1 (and this is a contribution of this paper) is that it allows us to interpret the waterfilling operator as a projector, according to the following corollary.

Corollary 1: The waterfilling solution $\mathbf{p}^\star = [p^\star(1), \ldots, p^\star(N)]^T$ in (7) can be expressed as the projection of $-\mathbf{insr}$ given in (8) onto the simplex $\mathcal{S}$ in (9):

$$\mathbf{p}^\star = [-\mathbf{insr}]_{\mathcal{S}}. \quad (12)$$

Corollary 2: The waterfilling solution in the form

$$p^\star(k) = \left[\frac{\mu}{w_k} - \frac{\sigma_k^2}{|H(k)|^2}\right]_0^{p^{\max}(k)}, \quad k \in \{1, \ldots, N\} \quad (13)$$

where $\mathbf{w} = [w_1, \ldots, w_N]^T$ is any positive vector, can be expressed as the projection with respect to the weighted Euclidean norm[4] with weights $w_1, \ldots, w_N$, of $-\mathbf{insr}$ given in (8) onto the simplex $\mathcal{S}$ in (9):

$$\mathbf{p}^\star = [-\mathbf{insr}]_{\mathcal{S}}^{\mathbf{w}}. \quad (14)$$

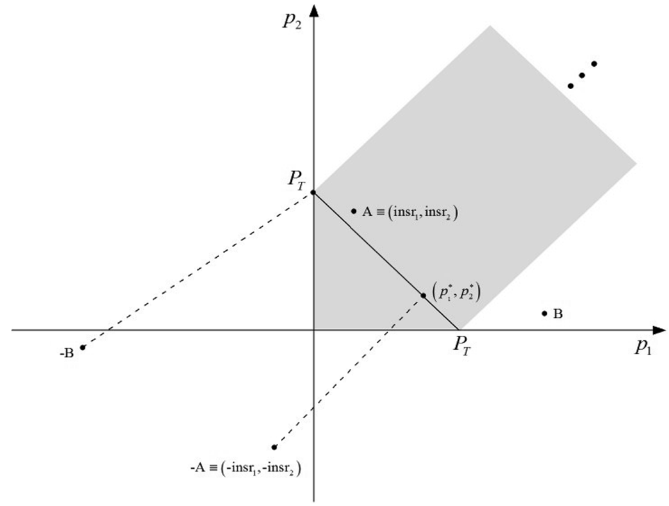


Fig. 1. Graphical interpretation of waterfilling solution (7) as a projection onto the two-dimensional simplex.

The graphical interpretation of the waterfilling solution as a Euclidean projector, for the single-carrier two-user case, is given in Fig. 1: For any $\mathbf{insr} \equiv (\mathrm{insr}_1, \mathrm{insr}_2)$ corresponding to a point in the interior of the gray region (e.g., point A), the waterfilling solution allocates power over both the channels. If, instead, the vector **insr** is outside the gray region (e.g., point B), all the power is allocated only over one channel, the one with the highest normalized gain.

B. Simultaneous Multiuser Waterfilling

In the multiuser scenario described in game $\mathcal{G}$, the optimal power allocation of each user depends on the power allocation of the other users through the received interference, according to the simultaneous multiuser waterfilling solution in (5).

As in the single-user case, introducing the vector $\mathbf{insr}_q(\mathbf{p}_{-q})$, defined as

$$[\mathbf{insr}_q(\mathbf{p}_{-q})]_k \triangleq \Gamma_q \frac{1 + \sum_{r \ne q} |H_{rq}(k)|^2 p_r(k)}{|H_{qq}(k)|^2} \quad (15)$$

with $k \in \{1, \ldots, N\}$, and invoking Lemma 1, we obtain the following corollary.

Corollary 3: The waterfilling operator $\mathsf{WF}_q(\mathbf{p}_{-q})$ in (6) can be expressed as the projection of $-\mathbf{insr}_q(\mathbf{p}_{-q})$ defined in (15) onto the simplex $\mathcal{P}_q$ given in (2):

$$\mathsf{WF}_q(\mathbf{p}_{-q}) = [-\mathbf{insr}_q(\mathbf{p}_{-q})]_{\mathcal{P}_q}. \quad (16)$$

Comparing (5) with (16), it is straightforward to see that all the Nash equilibria of game $\mathcal{G}$ can be alternatively obtained as

[4]The weighted Euclidean norm $\|\mathbf{x}\|_{2,\mathbf{w}}$ is defined as $\|\mathbf{x}\|_{2,\mathbf{w}} \triangleq (\sum_i w_i |x_i|^2)^{1/2}$ [44].

the fixed points of the mapping defined in (16), whose existence is guaranteed by [1, Theorem 2]

$$\mathbf{p}_q^{\star} = [-\mathbf{insr}_q(\mathbf{p}_{-q}^{\star})]_{\mathcal{P}_q}, \quad \forall\, q \in \Omega. \tag{17}$$

In Appendix B, we provide the key properties of the mapping in (16), that will be instrumental to obtain sufficient conditions for the convergence of the distributed iterative algorithms based on the waterfilling solution and described in Section IV-A.

IV. Distributed Algorithms

In [1, Theorem 2], we proved that, under some (sufficient) conditions on transmit powers, channels and network topology, the NE for game $\mathcal{G}$ is unique. Since there is no reason to expect a system to be initially at the equilibrium, the concept of equilibrium has a useful meaning in practice only if one is able to find a procedure that reaches such an equilibrium from nonequilibrium states. In this section, we focus on algorithms that converge to these equilibria.

Since we are interested in a decentralized implementation, where no signaling among different users is allowed, we consider only totally distributed iterative algorithms, where each user acts independently of the others to optimize its own power allocation while perceiving the other users as interference. The main issue of this approach is to guarantee the convergence of such an iterative scheme. In the following, we propose two alternative classes of totally distributed iterative algorithms along with their convergence properties, namely: iterative algorithms based on the waterfilling solution (6), and iterative algorithms based on the gradient projection mapping.

A. Distributed Algorithms Based on Waterfilling

So far, we have shown that the Nash equilibria of game $\mathcal{G}$ are fixed points [see (5)] of the waterfilling mapping defined in (6). Hence, to achieve these solutions by a distributed scheme, it is natural to employ an iterative algorithm based on the best response (6). Based on this idea, we consider two classes of iterative algorithms: *sequential* algorithms, where the users update their strategies sequentially according to a given schedule; and *simultaneous* algorithms, where all the users update their strategies at the same time. In the following sections, we provide a formal description of both algorithms and derive the conditions guaranteeing their convergence to the unique NE of the game.

Before describing the proposed algorithms, we introduce the following intermediate definitions. Given game $\mathcal{G}$, let $\mathcal{D}_q^{\min} \subseteq \{1,\ldots,N\}$ denote the set $\{1,\ldots,N\}$ deprived of the carrier indices that user q would never use as the best response set to any strategies adopted by the other users, for the given set of transmit power and propagation channels [1]:

$$\mathcal{D}_q^{\min} \triangleq \{k \in \{1,\ldots,N\} : \exists \mathbf{p}_{-q} \in \mathcal{P}_{-q} \text{ such that } [\mathsf{WF}_q(\mathbf{p}_{-q})]_k \neq 0\} \tag{18}$$

with $\mathsf{WF}_q(\cdot)$ defined in (6) and $\mathcal{P}_{-q} \triangleq \mathcal{P}_1 \times \cdots \times \mathcal{P}_{q-1} \times \mathcal{P}_{q+1} \times \cdots \times \mathcal{P}_Q$. We also introduce the nonnegative matrix $\mathbf{S}^{\max} \in \mathbb{R}_+^{Q\times Q}$, defined as

$$[\mathbf{S}^{\max}]_{qr} \triangleq \begin{cases} \Gamma_q \max\limits_{k\in\mathcal{D}_q\cap\mathcal{D}_r} \dfrac{|\bar{H}_{rq}(k)|^2}{|\bar{H}_{qq}(k)|^2}\dfrac{d_{qq}^{\gamma}}{d_{rq}^{\gamma}}\dfrac{P_r}{P_q}, & \text{if } r\neq q \\ 0, & \text{otherwise} \end{cases} \tag{19}$$

with the convention that the maximum in (19) is zero if $\mathcal{D}_q \cap \mathcal{D}_r$ is empty. In (19), each set $\mathcal{D}_q$ can be chosen as any subset of $\{1,\ldots,N\}$ such that $\mathcal{D}_q^{\min} \subseteq \mathcal{D}_q \subseteq \{1,\ldots,N\}$, with $\mathcal{D}_q^{\min}$ defined in (18). In Part I of the paper [1], we provided a simple procedure to compute such a set $\mathcal{D}_q$.

1) Sequential Iterative Waterfilling Algorithm Revisited: The sequential IWFA is an instance of the Gauss–Seidel scheme [3]: All users update their own strategies sequentially, performing the waterfilling solution (6). The algorithm is described in Algorithm 1.

Algorithm 1: Sequential IWFA

Set $\mathbf{p}_q^{(0)} =$ any feasible power allocation, $\forall q \in \Omega$;

`for` $n = 0 : \mathrm{Nit}$,

$\forall\, q \in \Omega$,

$$\mathbf{p}_q^{(n+1)} = \begin{cases} \mathsf{WF}_q\left(\mathbf{p}_{-q}^{(n)}\right), & \text{if } (n+1) \bmod Q = q \\ \mathbf{p}_q^{(n)}, & \text{otherwise} \end{cases} \tag{20}$$

`end`

The convergence of the algorithm is guaranteed under the following sufficient conditions.

Theorem 1: Assume that the following condition is satisfied:

$$\rho(\mathbf{S}^{\max}) < 1 \tag{C1}$$

where $\mathbf{S}^{\max}$ is defined in (19) and $\rho(\mathbf{S}^{\max})$ denotes the spectral radius[5] of the matrix $\mathbf{S}^{\max}$. Then, as $\mathrm{Nit} \to \infty$, the sequential IWFA described in Algorithm 1 converges linearly to the unique NE of game $\mathcal{G}$, for any set of initial conditions belonging to $\mathcal{P}$ and for any updating schedule.

Proof: See Appendix C. ■

Remark 1—Global Convergence and Uniqueness of the NE: Even though the optimization problem (1) is nonlinear, condition (C1) guarantees the *global* convergence of the sequential IWFA, irrespective of the specific users' updating order. Moreover, the global asymptotic stability of the NE implies also the uniqueness of the equilibrium. Condition (C1) indeed coincides with the uniqueness condition given in [1, Corollary 1].

To give additional insight into the physical interpretation of sufficient conditions for the convergence of the sequential IWFA, we provide the following corollaries of Theorem 1.

[5]The spectral radius $\rho(\mathbf{S})$ of the matrix $\mathbf{S}$, is defined as $\rho(\mathbf{S}) = \max\{|\lambda| : \lambda \in \mathrm{eig}(\mathbf{S})\}$, with $\mathrm{eig}(\mathbf{S})$ denoting the set of eigenvalues of $\mathbf{S}$ [44].

Corollary 4: A sufficient conditions for (C1) in Theorem 1 is given by one of the two following set of conditions:

$$\frac{\Gamma_q}{w_q}\sum_{r\neq q}\max_{k\in\mathcal{D}_r\cap\mathcal{D}_q}\left\{\frac{|\bar{H}_{rq}(k)|^2}{|\bar{H}_{qq}(k)|^2}\right\}\frac{d_{qq}^{\gamma}}{d_{rq}^{\gamma}}\frac{P_r}{P_q}w_r<1, \quad \forall q\in\Omega \quad \text{(C2)}$$

$$\frac{1}{w_r}\sum_{q\neq r}\Gamma_q\max_{k\in\mathcal{D}_r\cap\mathcal{D}_q}\left\{\frac{|\bar{H}_{rq}(k)|^2}{|\bar{H}_{qq}(k)|^2}\right\}\frac{d_{qq}^{\gamma}}{d_{rq}^{\gamma}}\frac{P_r}{P_q}w_q<1, \quad \forall r\in\Omega \quad \text{(C3)}$$

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

Corollary 5: The best vector $\mathbf{w}$ in (C2) and (C3) is given by the solution to the following geometric programming problem:

$$\begin{aligned}\underset{\mathbf{w},t}{\text{minimize}}\quad & t\\ \text{subject to}\quad & \sum_{r\neq q}G_{rq}t^{-1}w_q^{-1}w_r\leq 1,\quad \forall\, q\\ & \mathbf{w}>\mathbf{0}, t>0\end{aligned} \tag{21}$$

where G_{rq} is defined as

$$G_{rq}\triangleq\Gamma_q\max_{k\in\mathcal{D}_r\cap\mathcal{D}_q}\left\{\frac{|\bar{H}_{rq}(k)|^2}{|\bar{H}_{qq}(k)|^2}\right\}\frac{d_{qq}^{\gamma}}{d_{rq}^{\gamma}}\frac{P_r}{P_q}\frac{w_r}{w_q} \tag{22}$$

if (C2) is used, or as

$$G_{rq}\triangleq\Gamma_r\max_{k\in\mathcal{D}_r\cap\mathcal{D}_q}\left\{\frac{|\bar{H}_{qr}(k)|^2}{|\bar{H}_{rr}(k)|^2}\right\}\frac{d_{rr}^{\gamma}}{d_{qr}^{\gamma}}\frac{P_q}{P_r}\frac{w_q}{w_r} \tag{23}$$

if (C3) is used.

Note that, according to the definition of $\mathcal{D}_q$ in (19), one can always choose the full set $\mathcal{D}_q=\{1,\ldots,N\}$ in (C1) and (C2)–(C3). However, less stringent conditions are obtained by removing the unnecessary carriers, i.e., the carriers that, for the given power budget and interference levels, are never going to be used.

Remark 2—Physical Interpretation of Convergence Conditions: As already shown in Part I of the paper [1] for the uniqueness conditions of the NE, the convergence of sequential IWFA is guaranteed if the interferers are sufficiently far apart from the destinations. In fact, from (C1) or (C2)–(C3), one infers that, for any given set of channel realizations and power constraints, there exists a distance beyond which the sequential IWFA is guaranteed to converge, corresponding to the maximum level of interference that may be tolerated by each receiver [as quantified, e.g., in (C2)] or that may be generated by each transmitter [as quantified, e.g., in (C3)]. Interestingly, the presence of spectral mask constraints does not affect the convergence capability of the algorithm. Moreover, convergence condition (C1) [or (C2)–(C3)] has the same desired properties as the uniqueness conditions obtained in Part I of the paper: It is robust against the worst normalized channels $|\bar{H}_{rq}(k)|^2/|\bar{H}_{qq}(k)|^2$, since the subchannels corresponding to the highest ratios $|\bar{H}_{rq}(k)|^2/|\bar{H}_{qq}(k)|^2$ (and, in particular, the subchannels where $|\bar{H}_{qq}(k)|^2$ is vanishing) do not necessarily affect (C1) [or (C2)–(C3)], as their subcarrier indices may not belong to the set $\mathcal{D}_q$. This strongly relaxes the convergence conditions.

We can generalize the sequential IWFA given in Algorithm 1 by introducing a memory in the updating process, as given in Algorithm 2. We call this new algorithm *smoothed* sequential IWFA.

Algorithm 2: Smoothed Sequential IWFA

Set $\mathbf{p}_q^{(0)}=$ any feasible power allocation and $\alpha_q\in[0,1),\ \forall q\in\Omega$;

`for` $n=0:$ Nit.
[See (24), shown at the bottom of the page.]

`end`

Each factor $\alpha_q\in[0,1)$ in Algorithm 2 can be interpreted as a forgetting factor: The larger is α_q, the longer is the memory of the algorithm. In this paper, we are only considering constant channels. Nevertheless, in a time-varying scenario, (24) could be used to smooth the fluctuations due to the channel variations. In such a case, if the channel is fixed or highly stationary, it is convenient to take α_q close to 1; conversely, if the channel is rapidly varying, it is better to take a small α_q. Interestingly, the choice of $\{\alpha_q\}_{q\in\Omega}$ does not affect the convergence property of the algorithm (although it may affect the speed of convergence), as proved in the following.

Theorem 2: Assume that conditions of Theorem 1 are satisfied. Then, as $\text{Nit}\to\infty$, the smoothed sequential IWFA described in Algorithm 2 converges linearly to the unique NE of game $\mathcal{G}$, for any set of initial conditions in $\mathcal{P}$, updating schedule, and $\{\alpha_q\}_{q\in\Omega}$, with $\alpha_q\in[0,1),\forall q\in\Omega$.

Proof: See Appendix C. ∎

Remark 3—Comparison With Previous Results: The sequential IWFA described in Algorithm 1 generalizes the well-known sequential iterative waterfilling algorithm originally proposed by Yu *et al.* in [21] and then studied in [22]–[26], to the case in which the users have (possibly) different power budgets and there are spectral mask constraints. In fact, the algorithm in [21] can be obtained as a special case of Algorithm 1, by removing the spectral mask constraints in each set $\mathcal{P}_q$ in (2), (i.e., setting $p_q^{\max}(k)=+\infty,\forall k,q$) and replacing the waterfilling operator in (6) with the classical waterfilling solution

$$\mathsf{WF}_q(\mathbf{p}_1,\ldots,\mathbf{p}_{q-1},\mathbf{p}_{q+1},\ldots,\mathbf{p}_Q)\triangleq(\mu_q\mathbf{1}-\mathbf{insr}_q(\mathbf{p}_{-q}))^+ \tag{25}$$

$$\mathbf{p}_q^{(n+1)}=\begin{cases}\alpha_q\mathbf{p}_q^{(n)}+(1-\alpha_q)\mathsf{WF}_q\left(\mathbf{p}_{-q}^{(n)}\right), & \text{if } (n+1)\bmod Q=q,\\ \mathbf{p}_q^{(n)}, & \text{otherwise,}\end{cases}\qquad \forall\, q\in\Omega; \tag{24}$$

where $(x)^+ = \max(0, x)$ and $\mathbf{insr}_q$ is defined in (15).

The convergence of the sequential IWFA based on the mapping (25) has been studied in a number of works, each time obtaining milder convergence conditions. Specifically, in [21], the authors provided sufficient conditions for the existence of a NE and the convergence of the sequential IWFA, for a game composed by two players. This was later generalized to an arbitrary number of players in [22]–[25]. In [26], the case of flat-fading channels was considered. Interestingly, although derived under stronger constraints, incorporating for example spectral mask constraints, our convergence conditions have a broader validity than those obtained in [21]–[26], as shown in the following.[6]

Corollary 6: Sufficient conditions for (C2) are [21]–[25]

$$\Gamma_q \max_{k\in\{1,\ldots,N\}} \left\{ \frac{|\bar{H}_{rq}(k)|^2}{|\bar{H}_{qq}(k)|^2} \right\} \frac{d_{qq}^{\alpha}}{d_{rq}^{\alpha}} \frac{P_r}{P_q} < \frac{1}{Q-1}, \quad \forall\, r, q \in \Omega, r \neq q \qquad \text{(C4)}$$

or [23]

$$\Gamma_q \max_{k\in\{1,\ldots,N\}} \left\{ \frac{|\bar{H}_{rq}(k)|^2}{|\bar{H}_{qq}(k)|^2} \right\} \frac{d_{qq}^{\alpha}}{d_{rq}^{\alpha}} \frac{P_r}{P_q} < \frac{1}{2Q-3}, \quad \forall\, r, q \in \Omega, r \neq q. \qquad \text{(C5)}$$

In the case of flat-fading channels (i.e., $\bar{H}_{rq}(k) = \bar{H}_{rq}, \forall r, q$), condition (C2) becomes [26]

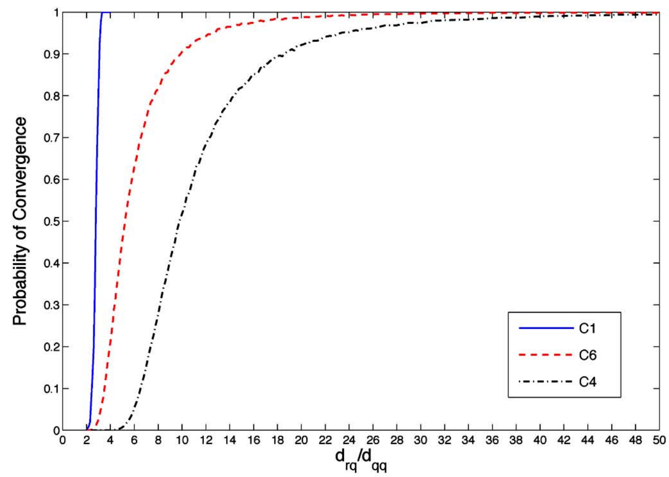
$$\Gamma_q \sum_{r\neq q} \frac{|\bar{H}_{rq}|^2}{|\bar{H}_{qq}|^2} \frac{d_{qq}^{\alpha}}{d_{rq}^{\alpha}} \frac{P_r}{P_q} < 1, \quad \forall\, q \in \Omega. \qquad (26)$$

Recently, alternative sufficient conditions for the convergence of sequential IWFA as given in Algorithm 1 were independently given in [27]. Specifically, the sequential IWFA was proved to converge to the unique NE of the game if the following condition is satisfied[7]:

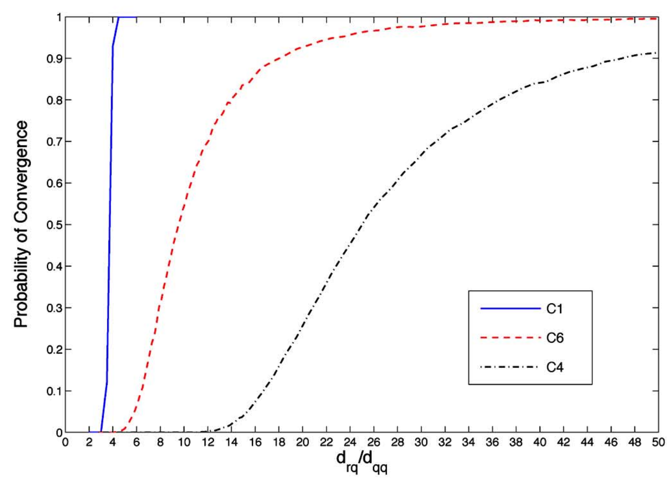
$$\rho(\mathbf{\Upsilon}) < 1 \qquad \text{(C6)}$$

where $\rho(\mathbf{\Upsilon})$ denotes the spectral radius of the matrix $\mathbf{\Upsilon} \triangleq (\mathbf{I} - \mathbf{S}_{\text{low}}^{\max})^{-1}\mathbf{S}_{\text{upp}}^{\max}$, with $\mathbf{S}_{\text{low}}^{\max}$ and $\mathbf{S}_{\text{upp}}^{\max}$ denoting the strictly lower and strictly upper triangular part of the matrix $\mathbf{S}^{\max}$, respectively, with $\mathbf{S}^{\max}$ defined, in our notation, as in (19), where each $\mathcal{D}_q$ is replaced by the full set $\{1, \ldots, N\}$.

As an example, in Fig. 2, we compare the range of validity of our convergence condition (C1) with that of (C4), and (C6), over a set of channel impulse responses generated as vectors composed of i.i.d. complex Gaussian random variables with zero mean and unit variance. In the figure, we plot the probability that conditions (C1), (C4), and (C6) are satisfied versus the ratio d_{rq}/d_{qq}, which measures how far apart are the interferers from the destination, with respect to the intended source. In Fig. 2(a), we consider a system composed by $Q = 5$ users, and in Fig. 2(b), a system with $Q = 15$ links. For the sake of simplicity, to limit the number of free parameters, we assumed $d_{rq} = d_{qr}$, $P_q = P_r \forall r, q$, and $\mathbf{w} = \mathbf{1}$. We tested our condition considering the set $\mathcal{D}_q$, obtained using the algorithm given in [1]. We can see, from Fig. 2, that the probability of guaranteeing convergence increases as the distance of the interferers, normalized to the source–destination distance, increases (i.e., the ratio d_{rq}/d_{qq} increases). Interestingly, the probability that (C1) is satisfied, differently from (C4) and (C6), exhibits a neat threshold behavior as it transits very rapidly from the nonconvergence guarantee to the almost certain convergence, as the ratio d_{rq}/d_{qq} increases by a small percentage. This shows that the convergence conditions depend, fundamentally, on the interferers distance, rather than on the channel realizations. Finally, it is worthwhile noticing that our conditions have a broader validity than (C4) and (C6). As an example, for a system with probability of guaranteeing convergence of 0.99 and $Q = 15$, conditions (C1) only require $d_{rq}/d_{qq} \simeq 4.2$, whereas conditions (C4) and (C6) require $d_{rq}/d_{qq} > 50$ and $d_{rq}/d_{qq} \simeq 40$, respectively. Furthermore, comparing Fig. 2(a) with (b), one can see that this difference increases as the number Q of links increases.



(a)



(b)

Fig. 2. Probability of (C1), (C4), and (C6) versus d_{rq}/d_{qq}; $Q = 5$ [subplot (a)], $Q = 15$ [subplot (b)], $\gamma = 2.5, d_{rq} = d_{qr}, d_{rr} = d_{qq} = 1, P_q = P_r$, $\Gamma_q = 1$, $P_q/\sigma_q^2 = 7$ dB, $P_r/(\sigma_q^2 d_{rq}^{\gamma}) = 3$ dB, $\forall r, q \in \Omega$, $\boldsymbol{w} = 1$.

[6] We summarize the main results of [21]–[25] using our notation.

[7] We write conditions of [27] using our notation.

Remark 4—Distributed Nature of the Algorithm: The sequential IWFA as described in Algorithms 1 and 2 can be implemented in a distributed way, since each user, to maximize its own rate, needs only to measure the PSD of the thermal noise plus the overall MUI [see (4)]. However, despite its appealing properties, the algorithm may suffer from slow convergence if the number of users in the network is large, as we will also show in Section IV-A-3). This drawback is due to the sequential schedule in the users' updates, wherein each user, to choose its own strategy, is forced to wait for all the other users scheduled before. Moreover, although distributed, both algorithms require that all users share a prescribed updating schedule. This requires a centralized synchronization mechanism that determines the order and the update times of the users. We show next how to remove these limitations.

2) Simultaneous Iterative Waterfilling Algorithm: To overcome the main limitation of sequential IWFAs given in Algorithms 1 and 2, we consider in this section the simultaneous version of the IWFA, called *simultaneous* IWFA. The algorithm is an instance of the Jacobi scheme [3]: At each iteration, all users update their own power allocation *simultaneously*, performing the waterfilling solution (6), given the interference generated by the other users in the *previous* iteration. Stated in mathematical terms, the proposed algorithm is described in Algorithm 3 [28], [29].

Algorithm 3: Simultaneous IWFA

Set $\mathbf{p}_q^{(0)}$ = any feasible power allocation, $\forall q \in \Omega$;

`for` $n = 0 : \mathrm{Nit}$,

$$\mathbf{p}_q^{(n+1)} = \mathsf{WF}_q\left(\mathbf{p}_{-q}^{(n)}\right), \qquad \forall\, q \in \Omega \tag{27}$$

`end`

As for the sequential IWFA, also in the simultaneous IWFA we can introduce a memory in the updating process and obtain the so-called *smoothed* simultaneous IWFA, as described in Algorithm 4 [28], [29].

Algorithm 4: Smoothed SIWFA

Set $\mathbf{p}_q^{(0)}$ = any feasible power allocation and $\alpha_q \in [0,1), \forall q \in \Omega$;

`for` $n = 0 : \mathrm{Nit}$

$\forall\, q \in \Omega$,

$$\mathbf{p}_q^{(n+1)} = \alpha_q \mathbf{p}_q^{(n)} + (1-\alpha_q)\mathsf{WF}_q\left(\mathbf{p}_{-q}^{(n)}\right), \forall q \in \Omega, \tag{28}$$

`end`

Interestingly, both Algorithm 3 and 4 are guaranteed to globally converge to the unique NE of the game, under the same sufficient conditions of the sequential IWFA, as proved in the following.

Theorem 3: Assume that conditions of Theorem 1 are satisfied. Then, as $\mathrm{Nit} \to \infty$, the simultaneous IWFAs described in Algorithm 3 and Algorithm 4 converge linearly to the unique NE of game $\mathcal{G}$, for any set of initial conditions in $\mathcal{P}$ and $\{\alpha_q\}_{q\in\Omega}$, with $\alpha_q \in [0,1), \forall q \in \Omega$.

Proof: See Appendix D. ■

Additional (weaker) convergence conditions for Algorithm 3 and 4 are given next. Introducing the matrix $\mathbf{S}(k) \in \mathbb{R}^{Q\times Q}$, defined as

$$[\mathbf{S}(k)]_{qr} \triangleq \begin{cases} \Gamma_q \dfrac{|\bar{H}_{rq}(k)|^2}{|\bar{H}_{qq}(k)|^2} \dfrac{d_{qq}^{\gamma}}{d_{rq}^{\gamma}} \dfrac{P_r}{P_q}, & \text{if } k \in \mathcal{D}_q \cap \mathcal{D}_r, \quad q \neq r \\ 0, & \text{otherwise} \end{cases} \tag{29}$$

we have the following theorem.

Theorem 4: Assume that the following conditions are satisfied:

$$\rho^{1/2}(\mathbf{S}^T(k)\mathbf{S}(k)) < 1, \quad \forall k \in \{1,\ldots,N\} \tag{30}$$

where $\mathbf{S}(k)$ is defined in (29). Then, as $\mathrm{Nit} \to \infty$, the sequential IWFA[8] described in Algorithm 3 converges linearly to the unique NE of game $\mathcal{G}$, for any set of initial conditions in $\mathcal{P}$.

Proof: See Appendix E. ■

Remark 5—Sequential Versus Simultaneous IWFA: Since both simultaneous IWFAs in Algorithms 3 and 4 are still based on the waterfilling solution (6), they keep the most appealing features of the sequential IWFA, namely its low complexity and distributed nature. In fact, as in the sequential IWFA, also in the simultaneous IWFA each user only needs to locally measure the PSD of the interference received from the other users and water-pour over this level. In addition, thanks to the Jacobi-based update, all the users are allowed to choose their optimal power allocation *simultaneously*. Hence, the simultaneous IWFA is expected to be faster than the sequential IWFA, especially if the number of active users in the network is large. We formalize this intuition in the next section.

3) Asymptotic Convergence Rate: In this section, we provide an upper bound of the convergence rate of both (smoothed) sequential and simultaneous IWFAs. The convergence rate can be either measured on the average or for the worst possible initial vector $\mathbf{p}^{(0)}$. In the following we focus on the latter approach, introducing the asymptotic convergence exponent.

Denoting by $\mathbf{p}^\star$ and $\mathbf{p}^{(n)}$ the NE of game $\mathcal{G}$ and the power allocation vector obtained by the proposed algorithm at the nth iteration, respectively, the distance between $\mathbf{p}^{(n)}$ and $\mathbf{p}^\star$ can be measured by some vector norm $\|\mathbf{p}^{(n)} - \mathbf{p}^\star\|$, which is to be compared with the initial distance $\|\mathbf{p}^{(0)} - \mathbf{p}^\star\|$. This leads to the following asymptotic convergence exponent for the worst-case convergence rate [43]:

$$d = -\sup_{\mathbf{p}^{(0)}\neq\mathbf{p}^\star} \lim_{n\to\infty} \frac{1}{n} \ln\left(\frac{\left\|\mathbf{p}^{(n)} - \mathbf{p}^\star\right\|}{\left\|\mathbf{p}^{(0)} - \mathbf{p}^\star\right\|}\right). \tag{31}$$

Since for large n

$$\left\|\mathbf{p}^{(n)} - \mathbf{p}^\star\right\| \simeq Ce^{-dn} \tag{32}$$

[8]Condition (30) is sufficient also for the convergence of the smoothed simultaneous IWFA described in Algorithm 4, provided that the second hand-side of (30) is replaced by $\epsilon = (1 - \max_{q\in\Omega}\alpha_q)/(1 - \min_{q\in\Omega}\alpha_q) \leq 1$.

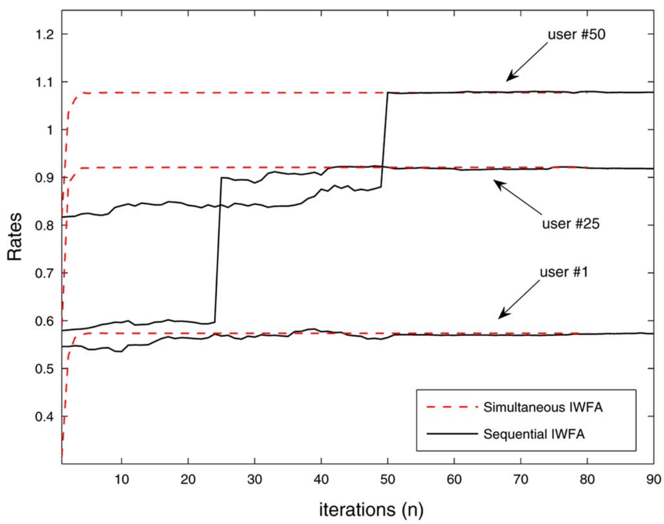


Fig. 3. Rates of the users versus iterations: sequential IWFA (solid line curves), simultaneous IWFA (dashed line curves), $Q = 50, \gamma = 2.5, d_{rq} = d_{qr}, d_{rr} = d_{qq} = 1, P_q = P_r, \Gamma_q = 1, P_q/\sigma_q^2 = 7$ dB, $P_r/(\sigma_q^2 d_{rq}^{\gamma}) = 3$ dB, $\forall r, q \in \Omega$.

where C is a constant that depends on the initial conditions, d gives the (asymptotic) number of iterations for the error to decrease by the factor $1/e$ (for the worst possible initial vector).

Since the waterfilling operator is not a monotone mapping, only (upper) bounds for the asymptotic convergence exponent can be obtained [41], [42], as given in the following.

Proposition 1: Let $d_{\text{seq}}^{\text{low}}$ and $d_{\text{sim}}^{\text{low}}$ be lower bound of d in (31) obtained using (smoothed) sequential IWFA in Algorithm 2 and (smoothed) simultaneous IWFA in Algorithm 4, respectively. Under condition (C2) of Corollary 4, we have

$$d_{\text{seq}}^{\text{low}} = -\log\left(\max_{q\in\Omega}\left\{\alpha_q + (1-\alpha_q)\frac{\Gamma_q}{w_q} \times \sum_{r\neq q}\max_{k\in\mathcal{D}_r\cap\mathcal{D}_q}\left\{\frac{|\bar{H}_{rq}(k)|^2}{|\bar{H}_{qq}(k)|^2}\right\}\frac{d_{qq}^{\gamma}}{d_{rq}^{\gamma}}\frac{P_r}{P_q}w_r\right\}\right), \quad (33)$$

$$d_{\text{sim}}^{\text{low}} = Q d_{\text{seq}}^{\text{low}}. \quad (34)$$

Proof: The proof follows directly from Proposition 2 in Appendix B. ■

Remark 6—Convergence Speed: Expression (33) shows that the convergence speed of the algorithms depends, as expected, on the memory factors $\{\alpha_q\}_{q\in\Omega}$ and on the level of interference. Given $\{\alpha_q\}_{q\in\Omega}$, the convergence speed increases as the interference level decreases. Since $d_{\text{sim}}^{\text{low}}$ and $d_{\text{seq}}^{\text{low}}$ are only bounds of the asymptotic convergence exponent, a comparison between the sequential IWFA and the simultaneous IWFA by $d_{\text{sim}}^{\text{low}}$ and $d_{\text{seq}}^{\text{low}}$ might not be fair. These bound becomes meaningful if $d_{\text{sim}}^{\text{low}}$ and $d_{\text{seq}}^{\text{low}}$ approximate with equality d_{sim} and d_{seq}, respectively, for some initial conditions (cf. [41]).

In Fig. 3, we compare the performance of the sequential and simultaneous IWFA, in terms of convergence speed. We consider a network composed of 50 links, and we show the rate evolution of three of the links corresponding to the sequential IWFA and simultaneous IWFA as a function of the iteration index n as defined in Algorithms 1 and 3. To make the figure not excessively overcrowded, we report only the curves of 3 out of 50 links. As expected, the sequential IWFA is slower than the simultaneous IWFA, especially if the number of active links Q is large, since each user is forced to wait for all the users scheduled in advance, before updating its own power allocation.

B. Distributed Algorithms Based on Gradient Projection

In this section, we propose two alternative distributed algorithms based on the gradient projection mapping. The first algorithm is an instance of the Jacobi scheme, whereas the second one is based on the Gauss–Seidel procedure. Both algorithms come out from an interesting interpretation of the Nash equilibria in (5) as solutions to a proper Nonlinear Variational Inequality (NVI) problem [30, Sec. 1.4.2], as we show next. The NVI problem is defined as follows. Given a subset $\mathcal{X} \subset \mathbb{R}^n$ of the Euclidean nth-dimensional space $\mathbb{R}^n$ and a mapping $\mathbf{f}: \mathbb{R}^n \mapsto \mathbb{R}^n$, the (nonlinear) *variational inequality* is to find a vector $\mathbf{x}^\star \in \mathcal{X}$ such that [30, Def. 1.1.1]

$$\text{(NVI)} \quad (\mathbf{x}-\mathbf{x}^\star)^T\mathbf{f}(\mathbf{x}^\star) \geq 0, \quad \forall\, \mathbf{x}\in\mathcal{X}. \quad (35)$$

All the Nash equilibria in (5) of game $\mathcal{G}$, can be written as solutions to a NVI problem. In fact, a feasible strategy profile $\mathbf{p}^\star = [\mathbf{p}_1^{\star T}, \ldots, \mathbf{p}_Q^{\star T}]^T$ satisfies (5) if and only if the following necessary and sufficient optimality conditions hold true [45][9] :

$$(\mathbf{p}_q - \mathbf{p}_q^\star)^T(-\nabla_q R_q(\mathbf{p}_q^\star, \mathbf{p}_{-q}^\star)) \geq 0, \quad \forall \mathbf{p}_q \in \mathcal{P}_q, \quad \forall q \in \Omega \quad (36)$$

where $\nabla_q R_q(\mathbf{p}_q^\star, \mathbf{p}_{-q}^\star)$ denotes the gradient vector of R_q with respect to $\mathbf{p}_q$, evaluated in $(\mathbf{p}_q^\star, \mathbf{p}_{-q}^\star)$, and $\mathcal{P}_q$ is defined in (2). Comparing (35) with (36), it is straightforward to see that a strategy profile $\mathbf{p}^\star$ is a NE of $\mathcal{G}$ if and only if it is a solution to the NVI problem (35), with the following identifications:

$$\mathcal{X} = \mathcal{X}_1 \times \ldots \times \mathcal{X}_Q \longleftrightarrow \mathcal{P} = \mathcal{P}_1 \times \ldots \times \mathcal{P}_Q \quad (37)$$

$$\mathbf{x} = \left[\mathbf{x}_1^T, \ldots, \mathbf{x}_Q^T\right]^T \longleftrightarrow \mathbf{p} = \left[\mathbf{p}_1^T, \ldots, \mathbf{p}_Q^T\right]^T \quad (38)$$

$$\mathbf{f}_q(\mathbf{x}_q, \mathbf{x}_{-q}) \longleftrightarrow -\nabla_q R_q(\mathbf{p}_q, \mathbf{p}_{-q}) \quad (39)$$

where $\mathbf{f}_q(\mathbf{x})$ denotes the qth component of $\mathbf{f}(\mathbf{x}) = [\mathbf{f}_1(\mathbf{x}), \ldots, \mathbf{f}_Q(\mathbf{x})]^T$. In fact, if (36) is satisfied for each q, then summing over q, (35) follows. Conversely, assume that (35) holds true for some $\mathbf{p}^\star$. Then, for any given q, choosing $\mathbf{p}_{-q} = \mathbf{p}_{-q}^\star$ and $\mathbf{p}_q \in \mathcal{P}_q$, we have $\mathbf{p} \triangleq (\mathbf{p}_q, \mathbf{p}_{-q}^\star) \in \mathcal{P}$ and $(\mathbf{p} - \mathbf{p}^\star)^T[-\nabla_1^T R_1(\mathbf{p}^\star), \ldots, -\nabla_Q^T R_Q(\mathbf{p}^\star)]^T = (\mathbf{p}_q - \mathbf{p}_q^\star)^T(-\nabla_q R_q(\mathbf{p}_q^\star, \mathbf{p}_{-q}^\star)) \geq 0, \forall \mathbf{p}_q \in \mathcal{P}_q$.

Building on the equivalence between (35) and (36), [30], [32], we can obtain distributed algorithms that reach the Nash equilibria of game $\mathcal{G}$ by looking for algorithms that solve the NVI problem in (35). A similar approach was already followed in [23], where the equivalence between the Nash equilibria of a DSL game that is a special case of $\mathcal{G}$ and the solutions to a proper nonlinear complementary problem [30] was shown. However,

[9]Observe that, given the strategy profiles of the other players, the optimal strategy of each player is a solution to the convex optimization problem defined in (1), whose optimality conditions, for any given q and $\mathbf{p}_{-q}^\star$, can be written as in (36) [45, Sec. 4.2.3].

the algorithms proposed in [23] to compute the NE solutions, in general, cannot be implemented in a distributed way, since they need a centralized control having access to all channel state information and to the PSD of all users. Differently from [23], we exploit the equivalence between (35) and (36) and propose two alternative totally distributed algorithms that do not require any centralized control to be implemented and have the same computational complexity as the IWFAs. To this end, we need the following intermediate result that comes directly from the NVI formulation in (36) [3, Prop. 5.1].

Lemma 2: Let β be a positive scalar and $\{\mathbf{G}_q\}_{q\in\Omega}$ be a set of symmetric positive definite matrices. A vector $\mathbf{p}^\star = [\mathbf{p}_1^{\star T},\ldots,\mathbf{p}_Q^{\star T}]^T$ is a NE of game $\mathcal{G}$ if and only if it is a fixed point of the following mapping[10]:

$$\mathbf{p}_q^\star = \mathbf{T}_{\mathbf{G}q}(\mathbf{p}^\star) \triangleq \left[\mathbf{p}_q^\star + \beta\mathbf{G}_q^{-1}\nabla_q R_q(\mathbf{p}_q^\star,\mathbf{p}_{-q}^\star)\right]_{\mathcal{P}_q}^{\mathbf{G}_q}, \quad \forall\, q\in\Omega \qquad (40)$$

where $[\,\cdot\,]_{\mathcal{P}_q}^{\mathbf{G}_q}$ is the Euclidean projection on $\mathcal{P}_q$ with respect to the vector norm $=\|\mathbf{x}\|_{\mathbf{G}_q} \triangleq (\mathbf{x}^T\mathbf{G}_q\mathbf{x})^{1/2}$.

Given Lemma 2, to reach the Nash equilibria of $\mathcal{G}$, it is natural to employ an iterative algorithm, based either on Jacobi or Gauss–Seidel schemes, using as best response for each user the mapping in (40). Specifically, if the mapping in (40) is used in the Jacobi scheme, we obtain the *simultaneous* iterative gradient projection algorithm (IGPA), as described in Algorithm 5.

Algorithm 5: Simultaneous IGPA

Set $\mathbf{p}_q^{(0)} =$ any feasible power allocation, $\forall q\in\Omega$, and $\beta>0$;

`for` $n=0:\mathrm{Nit}$

$$\mathbf{p}_q^{(n+1)} = \left[\mathbf{p}_q^{(n)} + \beta\mathbf{G}_q^{-1}\nabla_q R_q(\mathbf{p}^{(n)})\right]_{\mathcal{P}_q}^{\mathbf{G}_q}, \quad \forall\, q\in\Omega \quad (41)$$

`end`

The sequential update of the strategies from the players can be easily obtained from (40) by using the Gauss–Seidel scheme, and provides the *sequential* IGPA, as given in Algorithm 6.

[10]The mapping in (40) always admits at least one fixed point, since it satisfies Brouwer's fixed-point theorem [48, Theorem 4.2.5]. In fact, each set $\mathcal{P}_q$ is compact and convex, and the mapping $\mathbf{T}_{\mathbf{G}}(\mathbf{p}) = (\mathbf{T}_{\mathbf{G}_1}(\mathbf{p}))_{q\in\Omega}$ in (40) is continuous on $\mathcal{P} = \mathcal{P}_1\times\cdots\times\mathcal{P}_Q$, since each $R_q(\mathbf{p})$ is a continuously differentiable function of $\mathbf{p}$ and the projector operator is continuous as well [3, Prop. 3.2c].

Algorithm 6: Sequential IGPA

Set $\mathbf{p}_q^{(0)} =$ any feasible power allocation, $\forall q\in\Omega$, and $\beta>0$;

`for` $n=0:\mathrm{Nit}$

[see (42), shown at the bottom of the page].

`end`

The positive constant β and the set of (positive definite) matrices $\{\mathbf{G}_q\}_{q\in\Omega}$ are free parameters that affect the convergence property of the algorithms [3]. Sufficient conditions for the convergence of both sequential and simultaneous IGPAs are given in Appendix F.

Remark 7—Computation of the Projection: According to the best response mapping defined in (41) and (42), both Algorithms 5 and 6 require, at each iteration, the computation of the Euclidean projection $[\,\cdot\,]_{\mathcal{P}_q}^{\mathbf{G}_q}$ on the feasible strategy set $\mathcal{P}_q$ given in (2), with respect to the norm $\|\cdot\|_{\mathbf{G}_q}$. For any given β, $\{\mathbf{G}_q\}_{q\in\Omega}$, and $\mathbf{p}=[\mathbf{p}_1^T,\ldots,\mathbf{p}_Q^T]^T\in\mathcal{P}$, the projections in (41) and (42), written as

$$\mathbf{T}_{\mathbf{G}q}(\mathbf{p}) = \left[\mathbf{p}_q - \beta\mathbf{G}_q^{-1}\mathbf{f}_q(\mathbf{p})\right]_{\mathcal{P}_q}^{\mathbf{G}_q}, \quad \text{with} \quad \mathbf{f}_q(\mathbf{p}) = -\nabla_q R_q(\mathbf{p}) \qquad (43)$$

can be computed solving the following convex quadratic programming:

$$\begin{array}{ll}\underset{\tilde{\mathbf{p}}_q}{\text{minimize}} & \frac{1}{2}\tilde{\mathbf{p}}_q^T\mathbf{M}_q\tilde{\mathbf{p}}_q + \tilde{\mathbf{p}}_q^T\mathbf{q}_q \\ \text{subject to} & \tilde{\mathbf{p}}_q\in\mathcal{P}_q\end{array} \qquad (44)$$

where

$$\mathbf{M}_q \triangleq (1/\beta)\mathbf{G}_q$$

and

$$\mathbf{q}_q = \mathbf{q}_q(\mathbf{p}) \triangleq \mathbf{f}_q(\mathbf{p}) - (1/\beta)\mathbf{G}_q\mathbf{p}_q \qquad (45)$$

with $\mathbf{f}_q(\mathbf{p})$ given in (43). Observe that in the special case of $\mathbf{G}_q=\mathbf{I}$, the mapping $\mathbf{T}_{\mathbf{G}q}(\mathbf{p})$ in (43) becomes the classical Euclidean projection on the set $\mathcal{P}_q$, that can be efficiently computed as a waterfilling solution, as shown in Section III-A (cf. Lemma 1).

Interestingly, to compute the projection in (43), a variety of alternative algorithms can be obtained, interpreting $\mathbf{T}_{\mathbf{G}q}(\mathbf{p})$ in

$$\mathbf{p}_q^{(n+1)} = \begin{cases}\left[\mathbf{p}_q^{(n)} + \beta\mathbf{G}_q^{-1}\nabla_q R_q\left(\mathbf{p}^{(n)}\right)\right]_{\mathcal{P}_q}^{\mathbf{G}_q}, & \text{if } (n+1) \bmod Q = q, \\ \mathbf{p}_q^{(n)}, & \text{otherwise,}\end{cases} \quad \forall q\in\Omega. \qquad (42)$$

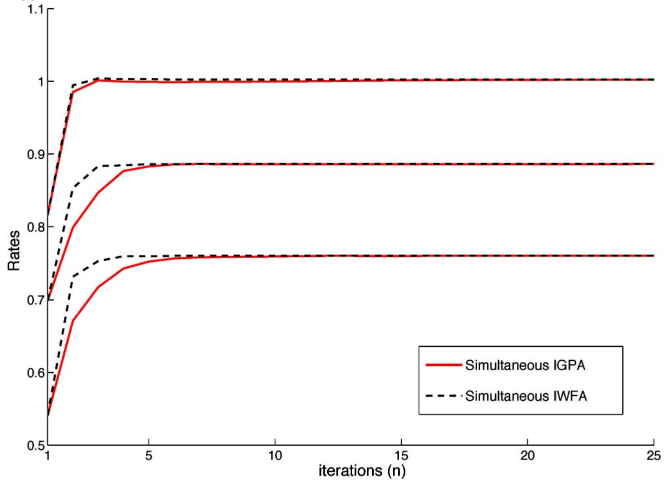


Fig. 4. Users' rates versus iterations; simultaneous IGPA (solid line curves) and simultaneous IWFA (dashed-line curves), $Q = 35, \gamma = 2.5, d_{rq} = d_{qr}, d_{rr} = d_{qq} = 1, P_q = P_r, \Gamma_q = 1, P_q/\sigma_q^2 = 7$ dB, $P_r/(\sigma_q^2 d_{rq}^{\gamma}) = 3$ dB, $\forall r, q \in \Omega$.

(43) as the unique solution to a proper linear variational inequality (LVI) problem [3], as we show next. Using the scaled projection theorem [3, Prop. 3.7(b)],[11] one can find that the projection $\mathbf{T}_{\mathbf{G}_q}(\mathbf{p})$ in (43) can be equivalently defined as the unique vector $\mathbf{p}_q^\star \in \mathcal{P}_q$ such that

$$(\mathbf{y}_q - \mathbf{p}_q^\star)^T \mathbf{G}_q \left(\mathbf{p}_q - \beta \mathbf{G}_q^{-1} \mathbf{f}_q(\mathbf{p}) - \mathbf{p}_q^\star\right) \leq 0, \quad \forall\, \mathbf{y}_q \in \mathcal{P}_q$$

which, since $\beta > 0$, can be rewritten as

$$\text{(LVI)} \quad (\mathbf{y}_q - \mathbf{p}_q^\star)^T (\mathbf{M}_q \mathbf{p}_q^\star + \mathbf{q}_q) \geq 0, \quad \forall \mathbf{y}_q \in \mathcal{P}_q \qquad (46)$$

where $\mathbf{M}_q$ and $\mathbf{q}_q$ are defined in (45). Inequality in (46) still defines a variational inequality problem [see (35)], but computationally simpler than the original one given in (36), as the function $\mathbf{M}_q \mathbf{p}_q^\star + \mathbf{q}_q(\mathbf{p})$ in (46), for any given $\mathbf{p}$, is linear in $\mathbf{p}_q^\star$. Observe that, since $\mathbf{G}_q$ (and thus $\mathbf{M}_q$) is positive definite, the (unique) solution $\mathbf{p}_q^\star$ in (46) is well defined, as LVI in (46) admits a unique solution [3, Prop. 5.5]. A variety of algorithms, known in the literature as linearized algorithms can be used to efficiently solve the LVI in (46). The interested reader may refer to [30]–[32] for a broad overview of these algorithms.

Remark 8—Distributed Nature of the Algorithms: Interestingly, both IGPAs keep the most appealing features of IWFAs, namely their low-complexity distributed nature. In fact, as in IWFAs, also in IGPAs each user needs only to locally measure the PSD of the overall interference received from the other users and project a vector that depends on this interference (i.e., $\nabla_q R_q(\mathbf{p})$) onto its own feasible set.

Numerical Example: As an example, in Fig. 4, we compare the performance of the simultaneous IGPA with the simultaneous IWFA, in terms of convergence speed. We consider a network composed of $Q = 35$ active users and compare the rate evolution of 3 out of 35 links as a function of the iteration index n, as defined in Algorithms 3 and 5. Interestingly, the simultaneous IGPA shows similar convergence speed than simultaneous IWFA. Thus, it can be used as a valid alternative to the simultaneous IWFA.

[11]The scaled projection theorem says that, given some $\mathbf{x} \in \mathbb{R}^n$ and a convex set $\mathcal{X} \subseteq \mathbb{R}^n$, a vector $\mathbf{z} \in \mathcal{X}$ is equal to $[\mathbf{x}]_{\mathcal{X}}^{\mathbf{G}}$ if and only if $(\mathbf{y}-\mathbf{z})^T \mathbf{G}(\mathbf{y}-\mathbf{z}) \leq 0$, $\forall \mathbf{y} \in \mathcal{X}$, where $[\mathbf{x}]_{\mathcal{X}}^{\mathbf{G}}$ denotes the Euclidean projection of $\mathbf{x}$ on $\mathcal{X}$ with respect to the norm $\|\cdot\|_{\mathbf{G}}$.

V. Conclusion

In this two-part paper, we have formulated the problem of finding the optimal linear transceivers in a multipoint-to-multipoint wideband network, as a strategic noncooperative game. We first considered the theoretical problem of maximizing mutual information on each link, given constraints on the spectral mask and transmit power. Then, to accommodate for practical implementation aspects, we focused on the competitive maximization of the transmission rate on each link, using finite-order constellations, under the same constraints as above plus a constraint on the average error probability. In Part I of the paper, we fully characterized both games by providing a unified expression for the optimal structure of the linear transceivers and deriving conditions for the uniqueness of the NE. In this Part II, we have focused on how to reach these equilibria using totally decentralized algorithms. We have proposed and analyzed alternative distributed iterative algorithms along with their convergence conditions, namely: 1) the sequential IWFA, which is a generalization of the well-known (sequential) iterative water-filling algorithm proposed by Yu *et al.* to the case where spectral mask constraints are incorporated in the optimization; 2) the simultaneous IWFA, which has been shown to converge faster than the sequential IWFA; and 3) the sequential and simultaneous IGPAs, which are based on the gradient projection best response, and establish an interesting link between the Nash equilibria of the game and the solutions to the corresponding variational inequality problem. Interestingly, the simultaneous IGPA has been shown to have approximately the same convergence speed and computational complexity of the simultaneous IWFA, and thus it can be a valid alternative to the algorithms based on the waterfilling solutions. We have derived the sufficient conditions for the *global* convergence of all the proposed algorithms that, although proved under stronger constraints (e.g., the additional spectral mask constraint), have broader validity than the convergence conditions known in the current literature for the sequential IWFA proposed by Yu *et al.*

We are currently investigating the extension of the proposed algorithms to the case in which the updating strategies are performed in a totally asynchronous way [49]. The other major extension that needs to be addressed is the situation where the channels and interference covariance matrices are known only within an inevitable estimation error.

Appendix

A. *Proof of Lemma 1*

First of all, observe that the objective function of the convex problem (10) is coercive on the feasible set [47]. Hence, a solution for the problem (10) exists [47]. Since problem (10) satisfies Slater's condition [45], [47], the Karush–Kuhn–Tucker (KKT) conditions are both necessary and sufficient for the optimality. The Lagrangian corresponding to the constrained convex problem (10) is

$$\mathcal{L} = \frac{1}{2}\sum_{k=1}^{N}(x_k + x_{0,k})^2 - \sum_{k=1}^{N}\nu_k x_k - \tilde{\mu}\left(\frac{1}{N}\sum_{k=1}^{N}x_k - 1\right) + \sum_{k=1}^{N}\lambda_k(x_k - p^{\max}(k)) \quad (47)$$

and the KKT conditions are

$$\begin{array}{ll} x_k + x_{0,k} - \nu_k + \lambda_k - \mu = 0, & \\ \nu_k \geq 0, \quad x_k \geq 0, \quad \nu_k x_k = 0, & \\ \lambda_k \geq 0, \quad x_k \leq p^{\max}(k), & k \in \{1,\ldots,N\} \\ \lambda_k(x_k - p^{\max}(k)) = 0, & \\ \frac{1}{N}\sum_{k=1}^{N} x_k = 1, & \end{array} \quad (48)$$

with $x_{0,k} \geq 0, \forall k \in \{1,\ldots,N\}$ and $\mu \triangleq \tilde{\mu}/N$.

Observe that, if $(1/N)\sum_{k=1}^{N} p^{\max}(k) < 1$ or $p^{\max}(k) < 0$ for some k, then the problem is infeasible; if $(1/N)\sum_{k=1}^{N} p^{\max}(k) = 1$, then the problem admits the trivial solution $x_k = p^{\max}(k), \forall k$; if $p^{\max}(k) = 0$ for some k, then $x_k = 0$. Here after, we thus assume that all the subcarrier indices corresponding to the zero-valued $p^{\max}(k)$'s have been removed and $(1/N)\sum_{k=1}^{N} p^{\max}(k) > 1$ (to avoid the trivial solution).

First of all, observe that $\mu > 0$ (in fact, $\mu \leq 0$ is not admissible, since the constraint $(1/N)\sum_{k=1}^{N} x_k = 1$ necessarily implies $x_k = -x_{0,k} - \lambda_k + \mu > 0$, for at least one k). If $x_k = 0$, since $\nu_k \geq 0$ and (by the complementary slackness condition) $\lambda_k = 0$, then $\mu - x_{0,k} = -\nu_k \leq 0$. If $0 < x_k < p^{\max}(k)$, then $\nu_k = 0$ and $\lambda_k = 0$; which provides $x_k = \mu - x_{0,k}$, (observe that $0 < \mu - x_{0,k} < p^{\max}(k)$). Finally, if $x_k = p^{\max}(k)$, then $\nu_k = 0$ and $\lambda_k \geq 0$, which implies $\mu - x_{0,k} \geq p^{\max}(k)$. Since, for each k, the values of the admissible solution induce a partition on the set of the μ values, the solution can be written as

$$x_k = [\mu - x_{0,k}]_0^{p^{\max}(k)} = \begin{cases} 0, & \text{if } \mu - x_{0,k} \leq 0, \\ \mu - x_{0,k}, & \text{if } 0 < \mu - x_{0,k} < p^{\max}(k), \\ p^{\max}(k), & \text{if } \mu - x_{0,k} \geq p^{\max}(k), \end{cases}$$

with $k \in \{1,\ldots,N\}$, where μ is chosen so that $(1/N)\sum_{k=1}^{N}[\mu - x_{0,k}]_0^{p^{\max}(k)} = 1$. ■

B. Properties of Waterfilling Projection

First of all, it is convenient to rewrite the waterfilling operator in (16) as

$$\mathsf{WF}_q(\mathbf{p}_{-q}) = \left[-\boldsymbol{\sigma}_q - \sum_{r\neq q}\mathbf{H}_{rq}\mathbf{p}_r\right]_{\mathcal{P}_q} \quad (49)$$

where $\mathcal{P}_q$ is defined in (2), and

$$\begin{aligned} \mathbf{H}_{rq} &\triangleq \Gamma_q \text{diag}\left(\frac{|H_{rq}(1)|^2}{|H_{qq}(1)|^2},\ldots,\frac{|H_{rq}(N)|^2}{|H_{qq}(N)|^2}\right), \\ \boldsymbol{\sigma}_q &\triangleq \Gamma_q[1/|H_{qq}(1)|^2,\ldots,1/|H_{qq}(N)|^2]^T. \end{aligned} \quad (50)$$

Building on (49), we derive now a key property of the waterfilling operator that will be fundamental in proving Theorems 1 and 3. To this end, we introduce the following mapping. For technical reasons, we first define

$$\tilde{p}_q^{\max}(k) \triangleq \begin{cases} p_q^{\max}(k), & \text{if } k \in \mathcal{D}_q \\ 0, & \text{otherwise} \end{cases} \quad (51)$$

where $\mathcal{D}_q^{\min} \subseteq \mathcal{D}_q \subseteq \{1,\ldots,N\}$ [see (19)], and introduce the admissible set $\mathcal{P}^{\text{eff}} = \mathcal{P}_1^{\text{eff}} \times \cdots \times \mathcal{P}_Q^{\text{eff}} \subseteq \mathcal{P}$, where $\mathcal{P}_q^{\text{eff}}$ is the subset of $\mathcal{P}_q$ containing all the feasible power allocations of user q, with zero power over the carriers that user q would never use, for the given power budget and interference level, in any of its waterfilling solutions (6), against any admissible strategy of the others[12]:

$$\begin{aligned} \mathcal{P}_q^{\text{eff}} &\triangleq \{\mathbf{p}_q \in \mathcal{P}_q \text{ with } p_q(k) = 0,\ \forall k \notin \mathcal{D}_q\} \\ &= \left\{\mathbf{p}_q \in \mathbb{R}^N : \frac{1}{N}\sum_{k=1}^{N} p_q(k) = 1, \right. \\ &\qquad \left. 0 \leq p_q(k) \leq \tilde{p}_q^{\max}(k), \quad \forall k \in \{1,\ldots,N\}\right\} \end{aligned} \quad (52)$$

where the second equality in (52) follows from the properties of the waterfilling solution (6) (cf. Appendix A). Observe that, because of (52), the game does not change if we use $\mathcal{P}^{\text{eff}}$ instead of the original $\mathcal{P}$. For any given $\{\alpha_q\}_{q\in\Omega}$with $\alpha_q \in [0,1)$, let $\mathbf{T}(\mathbf{p}) = (\mathbf{T}_q(\mathbf{p}))_{q\in\Omega} : \mathcal{P}^{\text{eff}} \mapsto \mathcal{P}^{\text{eff}}$ be the mapping defined, for each q, as

$$\begin{aligned} \mathbf{T}_q(\mathbf{p}) &\triangleq \alpha_q\, \mathbf{p}_q + (1-\alpha_q)\left[-\boldsymbol{\sigma}_q - \sum_{r\neq q}\mathbf{H}_{rq}\mathbf{p}_r\right]_{\mathcal{P}_q^{\text{eff}}} \\ &= \alpha_q\, \mathbf{p}_q + (1-\alpha_q)\left[-\boldsymbol{\sigma}_q - \sum_{r\neq q}\mathbf{H}_{rq}\mathbf{p}_r\right]_{\mathcal{P}_q}, \\ &\qquad \mathbf{p} \in \mathcal{P}^{\text{eff}}, \alpha_q \in [0,1) \end{aligned} \quad (53)$$

where the second equality follows from (52). Observe that the operator in (53) is indeed a mapping from $\mathcal{P}^{\text{eff}}$ to $\mathcal{P}^{\text{eff}}$, due to the convexity of $\mathcal{P}^{\text{eff}}$. Moreover, it follows from (49) that all the Nash equilibria $\mathbf{p}^\star \triangleq (\mathbf{p}_q^\star)_{q\in\Omega}$ of game $\mathcal{G}$ in (1) [see (17)] satisfy the following set of equations:

$$\mathbf{p}_q^\star = \left[-\boldsymbol{\sigma}_q - \sum_{r\neq q}\mathbf{H}_{rq}\mathbf{p}_r^\star\right]_{\mathcal{P}_q^{\text{eff}}}, \quad \forall\, q \in \Omega \quad (54)$$

[12]Observe that *all* the subcarriers that user q will never use are considered if $\mathcal{D}_q$ is chosen so that $\mathcal{D}_q = \mathcal{D}_q^{\min}$.

which correspond to the fixed points in $\mathcal{P}^{\text{eff}}$ of the mapping $\mathbf{T}$ defined in (53). Hence, the existence of at least one fixed point for $\mathbf{T}$ is guaranteed by the existence of a NE for game $\mathcal{G}$ [1, Theorem 2].

Before proving the main property of the mapping $\mathbf{T}$, we need the following intermediated definitions. Given $\mathbf{T}$ in (53) and some $\mathbf{w} \triangleq [w_q, \ldots, w_Q]^T > \mathbf{0}$, let $\|\cdot\|_{2,\text{block}}^{\mathbf{w}}$ denote the (vector) block-maximum norm, defined as [3]

$$\|\mathbf{T}(\mathbf{p})\|_{2,\text{block}}^{\mathbf{w}} \triangleq \max_{q\in\Omega} \frac{\|\mathbf{T}_q(\mathbf{p})\|_2}{w_q} \tag{55}$$

where $\|\cdot\|_2$ is the Euclidean norm. Let $\|\cdot\|_{\infty,\text{vec}}^{\mathbf{w}}$ be the *vector* weighted maximum norm, defined as [44]

$$\|\mathbf{x}\|_{\infty,\text{vec}}^{\mathbf{w}} \triangleq \max_{q\in\Omega} \frac{|x_q|}{w_q}, \quad \mathbf{w} > \mathbf{0}, \quad \mathbf{x} \in \mathbb{R}^Q \tag{56}$$

and let $\|\cdot\|_{\infty,\text{mat}}^{\mathbf{w}}$ denote the *matrix* norm induced by $\|\cdot\|_{\infty,\text{vec}}^{\mathbf{w}}$, defined as [44]

$$\|\mathbf{A}\|_{\infty,\text{mat}}^{\mathbf{w}} \triangleq \max_{q} \frac{1}{w_q}\sum_{r=1}^{Q}[\mathbf{A}]_{qr}w_r, \quad \mathbf{A} \in \mathbb{R}^{Q\times Q}. \tag{57}$$

Finally, define $\|\cdot\|_{2,\mathcal{D}_q}$ as

$$\|\mathbf{x}\|_{2,\mathcal{D}_q} \triangleq \left(\sum_{k\in\mathcal{D}_q}(x_k)^2\right)^{1/2}, \quad \mathbf{x} \in \mathbb{R}^N \tag{58}$$

with $\mathcal{D}_q$ defined in (18). Observe that $\|\cdot\|_{2,\mathcal{D}_q}$ is not a vector norm (as does not satisfy the positivity property), but it is a vector seminorm [44].

The mapping $\mathbf{T}$ in (53) is said to be a *block-contraction* with modulus β, with respect to the norm $\|\cdot\|_{2,\text{block}}^{\mathbf{w}}$ in (55), if there exists $\beta \in [0,1)$ such that [3, Sec. 3.1.2]

$$\left\|\mathbf{T}(\mathbf{p}^{(1)}) - \mathbf{T}(\mathbf{p}^{(2)})\right\|_{2,\text{block}}^{\mathbf{w}} \le \beta\left\|\mathbf{p}^{(1)} - \mathbf{p}^{(2)}\right\|_{2,\text{block}}^{\mathbf{w}}, \quad \forall\, \mathbf{p}^{(1)},\mathbf{p}^{(2)} \in \mathcal{P}_q^{\text{eff}}. \tag{59}$$

We provide now some interesting properties for the mapping $\mathbf{T}$ in (53), that will be instrumental to study the convergence of both sequential and simultaneous IWFAs.

Lemma 3 (Nonexpansive Property of the Waterfilling Mapping): Given $\mathcal{D}_q$, $\mathcal{P}_q^{\text{eff}}$, and $\|\cdot\|_{2,\mathcal{D}_q}$ defined in (19), (52), and (58), respectively, let $[\,\cdot\,]_{\mathcal{P}_q^{\text{eff}}}^{\|\cdot\|}$ denote the projector operator onto the convex set $\mathcal{P}_q^{\text{eff}}$ with respect to the vector norm $\|\cdot\|$. Then, $[\,\cdot\,]_{\mathcal{P}_q^{\text{eff}}}^{\|\cdot\|_2}$ satisfies the following nonexpansive property:

$$\left\|[\mathbf{x}]_{\mathcal{P}_q^{\text{eff}}}^{\|\cdot\|_2} - [\mathbf{y}]_{\mathcal{P}_q^{\text{eff}}}^{\|\cdot\|_2}\right\|_{2,\mathcal{D}_q} \le \|\mathbf{x}-\mathbf{y}\|_{2,\mathcal{D}_q}, \qquad \forall \mathbf{x},\mathbf{y} \in \mathbb{R}_+^N. \tag{60}$$

Proof: For any given $\varepsilon > 0$, let $\|\cdot\|_{2,\mathcal{D}_q}^{\varepsilon}$ denote the weighted vector norm (derived from an inner product [44]), defined as

$$\|\mathbf{x}\|_{2,\mathcal{D}_q}^{\varepsilon} \triangleq \left(\sum_{k\in\mathcal{D}_q}(x_k)^2 + \varepsilon\sum_{k\notin\mathcal{D}_q}(x_k)^2\right)^{1/2}, \quad \mathbf{x} \in \mathbb{R}^N, \quad \varepsilon > 0. \tag{61}$$

Then the projector $[\,\cdot\,]_{\mathcal{P}_q^{\text{eff}}}^{\|\cdot\|_{2,\mathcal{D}_q}^{\varepsilon}}$ satisfies the following inequality:

$$\begin{aligned}\left\|[\mathbf{x}]_{\mathcal{P}_q^{\text{eff}}}^{\|\cdot\|_2} - [\mathbf{y}]_{\mathcal{P}_q^{\text{eff}}}^{\|\cdot\|_2}\right\|_{2,\mathcal{D}_q}^{\varepsilon} &= \left\|[\mathbf{x}]_{\mathcal{P}_q^{\text{eff}}}^{\|\cdot\|_{2,\mathcal{D}_q}^{\varepsilon}} - [\mathbf{y}]_{\mathcal{P}_q^{\text{eff}}}^{\|\cdot\|_{2,\mathcal{D}_q}^{\varepsilon}}\right\|_{2,\mathcal{D}_q}^{\varepsilon} \\ &\le \|\mathbf{x}-\mathbf{y}\|_{2,\mathcal{D}_q}^{\varepsilon}, \\ &\forall\, \mathbf{x},\mathbf{y} \in \mathbb{R}_+^N, \quad \forall \varepsilon > 0 \end{aligned} \tag{62}$$

where the equality in (62) follows from $[\mathbf{x}_0]_{\mathcal{P}_q^{\text{eff}}}^{\|\cdot\|_2} = [\mathbf{x}_0]_{\mathcal{P}_q^{\text{eff}}}^{\|\cdot\|_{2,\mathcal{D}_q}^{\varepsilon}}$, due to the equivalence between the optimization problem (10) and the same problem where the original objective function is replaced by $(\|\mathbf{x} - (-\mathbf{x}_0)\|_{2,\mathcal{D}_q}^{\varepsilon})^2 = \sum_{k\in\mathcal{D}_q}(x_k + x_{0,k})^2 + \varepsilon\sum_{k\notin\mathcal{D}_q}(-x_{0,k})^2$ (since any $\mathbf{x} \in \mathcal{P}_q^{\text{eff}}$ is such that $x_k = 0$, $\forall k \notin \mathcal{D}_q$); and the inequality in (62) represents the nonexpansion property of the projector $[\,\cdot\,]_{\mathcal{P}_q^{\text{eff}}}^{\|\cdot\|_{2,\mathcal{D}_q}^{\varepsilon}}$ in the norm $\|\cdot\|_{2,\mathcal{D}_q}^{\varepsilon}$ [3, Prop. 3.2(c)].[13]

Since $\|\cdot\|_{2,\mathcal{D}_q}^{\varepsilon}$ is a continuous function of $\varepsilon > 0$, taking in (62) the limit as $\varepsilon \to 0$, and using

$$\lim_{\varepsilon\to 0}\|\cdot\|_{2,\mathcal{D}_q}^{\varepsilon} = \|\cdot\|_{2,\mathcal{D}_q}$$

we obtain the desired inequality, as stated in (60). ■

Proposition 2 (Contraction Property of Mapping $\mathbf{T}$*):* Given $\mathbf{w} \triangleq [w_1, \ldots, w_Q]^T > \mathbf{0}$, assume that the following condition is satisfied:

$$\|\mathbf{S}^{\max}\|_{\infty,\text{mat}}^{\mathbf{w}} < 1 \tag{63}$$

where $\mathbf{S}^{\max}$ and $\|\cdot\|_{\infty,\text{mat}}^{\mathbf{w}}$ are defined in (19) and (57), respectively. Then, the mapping $\mathbf{T}$ defined in (53) is a block-contraction with modulus $\beta = \|\mathbf{S}^{\max}\|_{\infty,\text{mat}}^{\mathbf{w}}$, with respect to the block-maximum norm $\|\cdot\|_{2,\text{block}}^{\mathbf{w}}$ defined in (55).

Proof: The proof consists of showing that, under (63), the mapping $\mathbf{T}$ satisfies (59), with $\beta = \|\mathbf{S}^{\max}\|_{\infty}^{\mathbf{w}}$. Given $\mathbf{p}^{(1)} = (\mathbf{p}_q^{(1)}, \ldots, \mathbf{p}_Q^{(1)}) \in \mathcal{P}^{\text{eff}}$ and $\mathbf{p}^{(2)} = (\mathbf{p}_1^{(2)}, \ldots, \mathbf{p}_Q^{(2)}) \in \mathcal{P}^{\text{eff}}$, define, for each $q \in \Omega$,

$$e_{\mathbf{T}q} \triangleq \left\|\mathbf{T}_q(\mathbf{p}^{(1)}) - \mathbf{T}_q(\mathbf{p}^{(2)})\right\|_2 \quad \text{and } e_q \triangleq \left\|\mathbf{p}_q^{(1)} - \mathbf{p}_q^{(2)}\right\|_2. \tag{64}$$

Then, we have, $\forall \mathbf{p}^{(1)}, \mathbf{p}^{(2)} \in \mathcal{P}^{\text{eff}}, \quad \forall q \in \Omega$,

$$e_{\mathbf{T}q} = \left\|\mathbf{T}_q(\mathbf{p}^{(1)}) - \mathbf{T}_q(\mathbf{p}^{(2)})\right\|_2 \tag{65}$$

[13]Observe that the nonexpansive property of the projector, usually given in the Euclidean norm, is preserved in any vector norm (derived from an inner product) used to define the projection.

$$\leq \alpha_q \left\| \mathbf{p}_q^{(1)} - \mathbf{p}_q^{(2)} \right\|_2 + (1-\alpha_q) \left\| \left[-\boldsymbol{\sigma}_q - \sum_{r\neq q} \mathbf{H}_{rq}\mathbf{p}_r^{(1)} \right]_{\mathcal{P}_q^{\text{eff}}} - \left[-\boldsymbol{\sigma}_q - \sum_{r\neq q} \mathbf{H}_{rq}\mathbf{p}_r^{(2)} \right]_{\mathcal{P}_q^{\text{eff}}} \right\|_2 \quad (66)$$

$$= \alpha_q e_q + (1-\alpha_q) \left\| \left[-\boldsymbol{\sigma}_q - \sum_{r\neq q} \mathbf{H}_{rq}\mathbf{p}_r^{(1)} \right]_{\mathcal{P}_q^{\text{eff}}} - \left[-\boldsymbol{\sigma}_q - \sum_{r\neq q} \mathbf{H}_{rq}\mathbf{p}_r^{(2)} \right]_{\mathcal{P}_q^{\text{eff}}} \right\|_{2,\mathcal{D}_q} \quad (67)$$

$$\leq \alpha_q e_q + (1-\alpha_q) \left\| \sum_{r\neq q} \mathbf{H}_{rq}\mathbf{p}_r^{(1)} - \sum_{r\neq q} \mathbf{H}_{rq}\mathbf{p}_r^{(2)} \right\|_{2,\mathcal{D}_q} \quad (68)$$

$$= \alpha_q e_q + (1-\alpha_q) \left\| \sum_{r\neq q} \bar{\mathbf{S}}_{rq} \left(\mathbf{p}_r^{(1)} - \mathbf{p}_r^{(2)} \right) \right\|_2 \quad (69)$$

$$\leq \alpha_q e_q + (1-\alpha_q) \sum_{r\neq q} \left(\max_k \left[\bar{\mathbf{S}}_{rq} \right]_{kk} \right) \left\| \mathbf{p}_r^{(1)} - \mathbf{p}_r^{(2)} \right\|_2 \quad (70)$$

$$= \alpha_q e_q + (1-\alpha_q) \sum_{r\neq q} \left(\max_{k\in\mathcal{D}_r\cap\mathcal{D}_q} [\mathbf{H}_{rq}]_{kk} \right) e_r \quad (71)$$

where (66) follows from (53) and the triangle inequality [44]; in (67) $\|\cdot\|_{2,\mathcal{D}_q}$ is defined in (58) and the equality follows from the fact that $\mathcal{P}_q^{\text{eff}}$ sets to zero the elements not in $\mathcal{D}_q$; (68) follows from Lemma 3 [see (60)]; in (69) $\bar{\mathbf{S}}_{rq}$ is a diagonal matrix defined as

$$[\bar{\mathbf{S}}_{rq}]_{kk} \triangleq \begin{cases} [\mathbf{S}_{rq}]_{kk}, & \text{if } k\in\mathcal{D}_r\cap\mathcal{D}_q, \\ 0, & \text{otherwise.} \end{cases} \quad (72)$$

Introducing the vectors

$$\mathbf{e}_{\mathbf{T}} \triangleq [\mathbf{e}_{\mathbf{T}_1},\ldots,\mathbf{e}_{\mathbf{T}_Q}]^T \quad \text{and} \quad \mathbf{e} \triangleq [e_1,\ldots,e_Q]^T \quad (73)$$

with $\mathbf{e}_{\mathbf{T}_q}$ and e_q defined in (64), and the matrix

$$\mathbf{S}_{\boldsymbol{\alpha}}^{\max} \triangleq \mathbf{D}_{\boldsymbol{\alpha}} + (\mathbf{I}-\mathbf{D}_{\boldsymbol{\alpha}})\mathbf{S}^{\max} \quad (74)$$

where $\mathbf{D}_{\boldsymbol{\alpha}} \triangleq \text{diag}(\alpha_q \ldots \alpha_Q)$ and $\mathbf{S}^{\max}$ is defined in (19). Then, the set of inequalities in (71) for all q, can be rewritten in vectorial form as

$$\mathbf{0} \leq \mathbf{e}_{\mathbf{T}} \leq \mathbf{S}_{\boldsymbol{\alpha}}^{\max}\mathbf{e}, \quad \forall\, \mathbf{p}^{(1)},\mathbf{p}^{(2)}\in\mathcal{P}^{\text{eff}}. \quad (75)$$

Using the weighted maximum norm $\|\cdot\|_{\infty,\text{vec}}^{\mathbf{w}}$ defined in (56) in combination with (75), we have, $\forall\mathbf{p}^{(1)},\, \mathbf{p}^{(2)}\in\mathcal{P}^{\text{eff}}$ and $\forall\mathbf{w}>\mathbf{0}$,

$$\|\mathbf{e}_{\mathbf{T}}\|_{\infty,\text{vec}}^{\mathbf{w}} \leq \|\mathbf{S}_{\boldsymbol{\alpha}}^{\max}\mathbf{e}\|_{\infty,\text{vec}}^{\mathbf{w}} \leq \|\mathbf{S}_{\boldsymbol{\alpha}}^{\max}\|_{\infty,\text{mat}}^{\mathbf{w}}\, \|\mathbf{e}\|_{\infty,\text{vec}}^{\mathbf{w}} \quad (76)$$

where $\|\cdot\|_{\infty,\text{mat}}^{\mathbf{w}}$ is the matrix norm induced by the vector norm $\|\cdot\|_{\infty,\text{vec}}^{\mathbf{w}}$ in (56) and defined in (57), [44]. Finally, using (76) and (55), we obtain, $\forall\mathbf{p}^{(1)},\ \mathbf{p}^{(2)}\in\mathcal{P}^{\text{eff}}$ and $\forall\mathbf{w}>\mathbf{0}$,

$$\begin{aligned} &\left\|\mathbf{T}(\mathbf{p}^{(1)}) - \mathbf{T}(\mathbf{p}^{(2)})\right\|_{2,\text{block}}^{\mathbf{w}} \\ &\quad = \max_q \frac{\left\|\mathbf{T}_q\left(\mathbf{p}^{(1)}\right) - \mathbf{T}_q\left(\mathbf{p}^{(2)}\right)\right\|_2}{w_q} = \|\mathbf{e}_{\mathbf{T}}\|_{\infty,\text{vec}}^{\mathbf{w}} \\ &\quad \leq \|\mathbf{S}_{\boldsymbol{\alpha}}^{\max}\|_{\infty,\text{mat}}^{\mathbf{w}}\, \|\mathbf{e}\|_{\infty,\text{vec}}^{\mathbf{w}} \\ &\quad = \|\mathbf{S}_{\boldsymbol{\alpha}}^{\max}\|_{\infty,\text{mat}}^{\mathbf{w}} \left\|\mathbf{p}^{(1)} - \mathbf{p}^{(2)}\right\|_{2,\text{block}}^{\mathbf{w}} \end{aligned} \quad (77)$$

which leads to a block-contraction for the mapping $\mathbf{T}$, if $\|\mathbf{S}_{\boldsymbol{\alpha}}^{\max}\|_{\infty,\text{mat}}^{\mathbf{w}} < 1$, implying condition (63) [since each $\alpha_q \in [0,1)$].

C. Proof of Theorem 1 and Theorem 2

Since the sequential IWFA described in Algorithm 1 is an instance of the smoothed sequential IWFA given in Algorithm 1 when $\alpha_q = 0$ for all $q\in\Omega$, to prove convergence of both algorithms, it is sufficient to show that Algorithm 2, under condition (C1), *globally* converges to the NE of game $\mathcal{G}$, for *any* given set $\{\alpha_q\}_{q\in\Omega}$, provided that each $\alpha_q\in[0,1)$. We thus focus in the following only on Algorithm 2, w.l.o.g.

It follows from Corollary 3 and (53) that Algorithm 2 is just an instance of the Gauss–Seidel scheme based on the mapping $\mathbf{T}$, defined in (53). Observe that, to study the convergence of Algorithm 2, there is no loss of generality in considering the mapping $\mathbf{T}$ defined in $\mathcal{P}^{\text{eff}} \subset \mathcal{P}$ instead of $\mathcal{P}$, since all the points produced by the algorithm (except possibly the initial point, which does not affect the convergence of the algorithm in the subsequent iterations) as well as the Nash equilibria of the game are confined, by definition, in $\mathcal{P}^{\text{eff}}$ [see (17) in Appendix B]. Convergence of the Gauss–Seidel scheme based on the mapping $\mathbf{T}$ is given by the following result that comes from [3, Prop. 1.4][14] and [3, Prop. 1.1a].

Proposition 3: If the mapping $\mathbf{T}:\mathcal{P}^{\text{eff}}\mapsto\mathcal{P}^{\text{eff}}$ defined in (53) is a block-contraction with respect to some vector norm, then 1) the mapping $\mathbf{T}$ has a unique fixed point in $\mathcal{P}^{\text{eff}}$ and 2) the sequence of vectors starting from any arbitrary point in $\mathcal{P}^{\text{eff}}$ and generated by the Gauss–Seidel algorithm based on the mapping $\mathbf{T}$, converges linearly to the fixed point of $\mathbf{T}$.

It follows from Proposition 2 and Proposition 3 that the global convergence of Algorithm 2 is guaranteed under the sufficient condition given in (63). Moreover, since (63) does not depend on $\{\alpha\}_{q\in\Omega}$, the convergence of the algorithm is not affected by the particular choice of α_q's as well [provided that each $\alpha_q\in[0,1)$].

To complete the proof, we just need to show that (63) is equivalent to (C1). Since $\mathbf{S}^{\max}$ is a nonnegative matrix, there exists a positive vector $\bar{\mathbf{w}}$ such that [3, Corollary 6.1]

$$\|\mathbf{S}^{\max}\|_{\infty,\text{mat}}^{\bar{\mathbf{w}}} < 1 \Leftrightarrow \rho(\mathbf{S}^{\max}) < 1. \quad (78)$$

Since the convergence of Algorithm 2 is guaranteed under (63), for any given $\mathbf{w}>\mathbf{0}$, we can choose $\mathbf{w}=\bar{\mathbf{w}}$ and use (78), which proves the desired result.

[14]Observe that the set $\mathcal{P}^{\text{eff}}$ defined in (52) is closed, as required in [3, Prop. 1.4].

Conditions (C2) and (C3) in Corollary 4 can be obtained as follows. Using [3, Prop. 6.2e]

$$\rho(\mathbf{S}^{\max}) \leq \|\mathbf{S}^{\max}\|_{\infty,\mathrm{mat}}^{\mathbf{w}}, \quad \forall\, \mathbf{w} > \mathbf{0} \tag{79}$$

a sufficient condition for the $\Rightarrow$ direction in (78) is

$$\|\mathbf{S}^{\max}\|_{\infty,\mathrm{mat}}^{\mathbf{w}} < 1 \tag{80}$$

for some given $\mathbf{w} > \mathbf{0}$; which provides (C2). The optimal vector $\mathbf{w}$ is given by the following geometric programming [45]

$$\begin{array}{ll} \underset{\mathbf{w}}{\text{minimize}} \max\limits_{q} & \sum\limits_{r\neq q} [\mathbf{S}^{\max}]_{qr} w_q^{-1} w_r \\ \text{subject to} & \mathbf{w} > \mathbf{0} \end{array}$$

which provides (21).

Condition (C3) is obtained similarly, still using (78) and $\rho(\mathbf{S}^{\max}) = \rho(\mathbf{S}^{\max\, T})$. ■

D. Proof of Theorem 3

The proof is based on the same steps as in Appendix C. It follows from [3, Prop. 1.1] that both Algorithm 3 and Algorithm 4 linearly converge to the unique NE of game $\mathcal{G}$, starting from any arbitrary point in $\mathcal{P}$, if the mapping $\mathbf{T}$ defined in (53) is a contraction [see (59)] in some vector norm. Using the block-maximum norm as defined in (55), invoking Proposition 2, and following the same approach as in Appendix C we obtain the desired sufficient condition (C1) for the global convergence of both Algorithm 3 and Algorithm 4, for any given set $\{\alpha\}_{q\in\Omega}$, [provided that each $\alpha_q \in [0,1)$]. ■

E. Proof of Theorem 4

Since Algorithm 3 is an instance of Algorithm 4 when each $\alpha_q = 0$, we focus only on the latter, w.l.o.g. The proof of the theorem is based on (pseudo) contraction arguments, similarly to what we already shown in Appendix C. The main difference with respect to the approach proposed in Appendix C is due to the alternative definition of the error vector generated by Algorithm 4, as detailed next.

Denoting by $\mathbf{p}^{(n)} \triangleq (\mathbf{p}_q^{(n)})_{q\in\Omega}$ the power allocation vector generated by Algorithm 4 at iteration $n \geq 1$, with arbitrary starting point $\mathbf{p}^{(0)} \in \mathcal{P}$, and using the mapping $\mathbf{T}$ defined in (53), we have

$$\mathbf{p}^{(n+1)} = \mathbf{T}\left(\mathbf{p}^{(n)}\right), \quad \forall\, n \geq 1. \tag{81}$$

Let $\mathbf{p}^{\star} \triangleq (\mathbf{p}_{\mathbf{q}}^{\star})_{q\in\Omega}$ be a NE of game $\mathcal{G}$ (and thus a fixed point of the mapping $\mathbf{T}$), whose existence is guaranteed by [1, Theorem 2]. Define the vector $\mathbf{e}^{(n)} \triangleq [\mathbf{e}_1^{(n)T}, \ldots, \mathbf{e}_Q^{(n)T}]^T$, with

$$\mathbf{e}_q^{(n)} \triangleq \mathbf{p}_q^{(n)} - \mathbf{p}_q^{\star}, \quad n \geq 1 \quad \text{and} \quad q \in \Omega \tag{82}$$

and, given $\{\alpha_q\}_{q\in\Omega}$ with $\alpha_q \in [0,1)$, define the $Q \times Q$ matrix $\mathbf{D}_{\boldsymbol{\alpha}} \triangleq \mathrm{diag}(\alpha_1 \ldots \alpha_Q)$. Then, for each $n \geq 1$, we have

$$\begin{aligned} &\left\|\mathbf{e}^{(n+1)}\right\|_2 \\ &\leq \left\|(\mathbf{D}_{\boldsymbol{\alpha}} \otimes \mathbf{I}_N)\mathbf{e}^{(n)}\right\|_2 \\ &\quad + \left\| \left[\begin{array}{c} -(1-\alpha_1)\left(\boldsymbol{\sigma}_1 + \sum_{r\neq 1} \mathbf{H}_{r1}\mathbf{p}_r^{(n)}\right) \\ \vdots \\ -(1-\alpha_Q)\left(\boldsymbol{\sigma}_Q + \sum_{r\neq Q} \mathbf{H}_{rQ}\mathbf{p}_r^{(n)}\right) \end{array}\right]_{\mathcal{P}^{\mathrm{eff}}} \right. \\ &\quad \left. - \left[\begin{array}{c} -(1-\alpha_1)\left(\boldsymbol{\sigma}_1 + \sum_{r\neq 1} \mathbf{H}_{r1}\mathbf{p}_r^{\star}\right) \\ \vdots \\ -(1-\alpha_Q)\left(\boldsymbol{\sigma}_Q + \sum_{r\neq Q} \mathbf{H}_{rQ}\mathbf{p}_r^{\star}\right) \end{array}\right]_{\mathcal{P}^{\mathrm{eff}}} \right\|_2 &(83) \\ &\leq \|(\mathbf{D}_{\boldsymbol{\alpha}} \otimes \mathbf{I}_N)\|_{2,\mathrm{mat}} \left\|\mathbf{e}^{(n)}\right\|_2 \\ &\quad + \left\| \left[\begin{array}{c} (1-\alpha_1)\sum_{r\neq 1} \bar{\mathbf{S}}_{r1}\left(\mathbf{p}_r^{(n)} - \mathbf{p}_r^{\star}\right) \\ \vdots \\ (1-\alpha_Q)\sum_{r\neq Q} \bar{\mathbf{S}}_{rQ}\left(\mathbf{p}_r^{(n)} - \mathbf{p}_r^{\star}\right) \end{array}\right] \right\|_2 &(84) \\ &= \|(\mathbf{D}_{\boldsymbol{\alpha}} \otimes \mathbf{I}_N)\|_{2,\mathrm{mat}} \left\|\mathbf{e}^{(n)}\right\|_2 \\ &\quad + \left\|(\mathbf{I}_N \otimes (\mathbf{I}_Q - \mathbf{D}_{\boldsymbol{\alpha}}))\mathbf{S}\mathbf{P}\mathbf{e}^{(n)}\right\|_2 &(85) \\ &\leq (\|(\mathbf{D}_{\boldsymbol{\alpha}} \otimes \mathbf{I}_N)\|_{2,\mathrm{mat}} \\ &\quad + \|(\mathbf{I}_N \otimes (\mathbf{I}_Q - \mathbf{D}_{\boldsymbol{\alpha}})\|_{2,\mathrm{mat}} \|\mathbf{S}\|_{2,\mathrm{mat}})^n \left\|\mathbf{e}^{(1)}\right\|_2 &(86) \end{aligned}$$

where (83) follows from the triangle inequality and $\mathcal{P}^{\mathrm{eff}} = \mathcal{P}_1^{\mathrm{eff}} \times \cdots \times \mathcal{P}_Q^{\mathrm{eff}}$, and "$\otimes$" denotes the Kronecker product; (84) follows from the nonexpansive property of the Euclidean projector, the definition of $\mathcal{P}^{\mathrm{eff}}$ [see (52)] and the definition of the diagonal matrices $\bar{\mathbf{S}}_{rq}$, as given in (72); and in (85) we have used a permutation matrix $\mathbf{P}$ so that the vector $\mathbf{e}^{(n)}$ given in (82), is replaced by $\tilde{\mathbf{e}}^{(n)} \triangleq \mathbf{P}\mathbf{e}^{(n)} = [\tilde{\mathbf{e}}_1^{(n)T}, \ldots, \tilde{\mathbf{e}}_N^{(n)T}]^T$, with $\tilde{\mathbf{e}}_k^{(n)} \triangleq [p_1^{(n)}(k), \ldots, p_Q^{(n)}(k)]^T - [p_1^{\star}(k), \ldots, p_Q^{\star}(k)]^T$, and the matrix $\mathbf{S}$ is defined as

$$\mathbf{S} \triangleq \mathrm{diag}(\mathbf{S}(1), \ldots, \mathbf{S}(N)) \tag{87}$$

with $\mathbf{S}(k)$ given in (29). The matrix norm $\|\mathbf{S}\|_{2,\mathrm{mat}}$ in (84) is the spectral norm (induced by the vector Euclidean norm [44]), defined as $\|\mathbf{S}\|_{2,\mathrm{mat}} \triangleq \rho^{1/2}(\mathbf{S}^T\mathbf{S})$.

From (86) it follows that Algorithm 4 converges to the NE $\mathbf{p}^{\star}$, from any starting point $\mathbf{p}^{(0)} \in \mathcal{P}$, if $(\|(\mathbf{D}_{\boldsymbol{\alpha}} \otimes \mathbf{I}_N)\|_{2,\mathrm{mat}} + \|(\mathbf{I}_N \otimes (\mathbf{I}_Q - \mathbf{D}_{\boldsymbol{\alpha}})\|_{2,\mathrm{mat}} \|\mathbf{S}\|_{2,\mathrm{mat}})^n$ in (86) approaches to zero as $n \to \infty$, which is guaranteed if the following conditions are satisfied:

$$\rho^{1/2}(\mathbf{S}^T(k)\mathbf{S}(k)) < \frac{1 - \max_{q\in\Omega}\alpha_q}{1 - \min_{q\in\Omega}\alpha_q}, \qquad \forall k \in \{1, \ldots, N\}$$

which provides the desired result. Given (86), the linear convergence of the algorithm follows directly from [3, Sec. 1.3.1]. ■

F. Proof of Convergence of Algorithm 5 and Algorithm 6

The global convergence of both sequential and simultaneous IGPAs, described in Algorithms 5 and 6, is guaranteed if Algorithms 5 and 6 satisfy [3, Prop. 1.1] and [3, Prop. 1.4], respectively. To this end, since each $\mathcal{P}_q$ is compact (and thus also $\mathcal{P} = \mathcal{P}_1 \times \cdots \times \mathcal{P}_Q$), it is sufficient that the mapping $\mathbf{T}_{\mathbf{G}}(\mathbf{p}) = (\mathbf{T}_{\mathbf{G}_q}^T(\mathbf{p}))_{q\in\Omega} : \mathcal{P} \mapsto \mathcal{P}$ is a block-contraction [see (59)] with respect to the norm $\|\cdot\|_{\mathbf{G},\text{block}}$, defined as

$$\|\mathbf{T}_{\mathbf{G}}(\mathbf{p})\|_{\mathbf{G},\text{block}} \triangleq \max_{q\in\Omega} \|\mathbf{T}_{\mathbf{G}_q}(\mathbf{p})\|_{\mathbf{G}_q,2} \quad (88)$$

where $\|\mathbf{T}_{\mathbf{G}_q}(\mathbf{p})\|_{\mathbf{G}_q,2} \triangleq (\mathbf{T}_{\mathbf{G}_q}^T(\mathbf{p})\mathbf{G}_q\mathbf{T}_{\mathbf{G}_q}(\mathbf{p}))^{1/2}$ and $\mathbf{T}_{\mathbf{G}_q}(\mathbf{p})$ is defined as

$$\mathbf{T}_{\mathbf{G}q}(\mathbf{p}) \triangleq \left[\mathbf{p}_q - \beta\mathbf{G}_q^{-1}\mathbf{f}_q(\mathbf{p}_q,\mathbf{p}_{-q})\right]_{\mathcal{P}_q}^{\mathbf{G}_q} \quad (89)$$

with $\mathbf{f}_q(\mathbf{p}_q,\mathbf{p}_{-q}) \triangleq -\nabla_q R_q(\mathbf{p}_q,\mathbf{p}_{-q})$. Rewriting $\mathbf{T}_{\mathbf{G}_q}(\mathbf{p})$ in (89) as $\mathbf{T}_{\mathbf{G}q}(\mathbf{p}) \triangleq [\mathbf{R}_{\mathbf{G}q}(\mathbf{p})]_{\mathcal{P}_q}^{\mathbf{G}_q}$, with

$$\mathbf{R}_{\mathbf{G}_q}(\mathbf{p}) \triangleq \mathbf{p}_q - \beta\mathbf{G}_q^{-1}\mathbf{f}_q(\mathbf{p}_q,\mathbf{p}_{-q}) \quad (90)$$

and using (88), we obtain $\forall \mathbf{p}^{(1)},\mathbf{p}^{(2)} \in \mathcal{P}$,

$$\begin{aligned}\left\|\mathbf{T}_{\mathbf{G}}(\mathbf{p}^{(1)}) - \mathbf{T}_{\mathbf{G}}(\mathbf{p}^{(2)})\right\|_{\mathbf{G},\text{block}} &= \max_q \left\|\left[\mathbf{R}_{\mathbf{G}q}(\mathbf{p}^{(1)})\right]_{\mathcal{P}_q}^{\mathbf{G}_q} - \left[\mathbf{R}_{\mathbf{G}q}(\mathbf{p}^{(2)})\right]_{\mathcal{P}_q}^{\mathbf{G}_q}\right\|_{\mathbf{G}_q,2} \\ &\le \max_q \left\|\mathbf{R}_{\mathbf{G}q}(\mathbf{p}^{(1)}) - \mathbf{R}_{\mathbf{G}q}(\mathbf{p}^{(2)})\right\|_{\mathbf{G}_q,2} \\ &= \left\|\mathbf{R}_{\mathbf{G}}(\mathbf{p}^{(1)}) - \mathbf{R}_{\mathbf{G}}(\mathbf{p}^{(2)})\right\|_{\mathbf{G},\text{block}} \end{aligned} \quad (91)$$

where the inequality follows from the nonexpansive property of the projection $[\cdot]_{\mathcal{P}_q}^{\mathbf{G}_q}$ in the norm $\|\cdot\|_{\mathbf{G}_q,2}$. From (91), it follows that a sufficient condition for $\mathbf{T}_{\mathbf{G}}$ being a contraction with respect to the norm $\|\cdot\|_{\mathbf{G},\text{block}}$ in (88) is that the mapping $\mathbf{R}_{\mathbf{G}}(\mathbf{p}) \triangleq (\mathbf{R}_{\mathbf{G}_1}(\mathbf{p}))_{q\in\Omega} : \mathcal{P} \mapsto \mathbb{R}^{QN}$ defined in (90) be a contraction with respect to the same norm.

We derive now sufficient conditions for $\mathbf{R}_{\mathbf{G}}(\mathbf{p})$ being a contraction with respect to $\|\cdot\|_{\mathbf{G},\text{block}}$ defined in (88). For the sake of simplicity, we will consider only the case in which $\mathbf{G}_q = \mathbf{I}, \forall q \in \Omega$.

We introduce the following notation: For any $\mathbf{f}_q(\mathbf{p}) \triangleq -\nabla_q R_q(\mathbf{p}_q,\mathbf{p}_{-q})$, let $\nabla_r \mathbf{f}_q(\mathbf{p})$ denote the $N \times N$ matrix, whose jth column is the gradient vector of the jth component of $\mathbf{f}_q(\mathbf{p})$, when viewed as function of $\mathbf{p}_r$. Then, we have the following result that comes directly from [3, Prop. 1.10].

Proposition 4: As $\text{Nit} \longrightarrow \infty$, the IGPAs described in Algorithms 5 and 6 converge to the unique NE of game $\mathcal{G}$ from any set of initial conditions in $\mathcal{P}$, if there exists a scalar $\delta \in [0,1)$ such that

$$\|\mathbf{I} - \beta\nabla_q \mathbf{f}_q(\mathbf{p})\|_{2,\text{mat}} + \sum_{r\neq q} \|\beta\nabla_r \mathbf{f}_q(\mathbf{p})\|_{2,\text{mat}} \le \delta, \quad \forall\, \mathbf{p} \in \mathcal{P}, \forall q \in \Omega \quad (92)$$

where $\|\mathbf{A}\|_{2,\text{mat}}$ denotes the spectral norm of the matrix $\mathbf{A}$.

We derive now a sufficient condition for (92). Using $\mathbf{f}_q(\mathbf{p}) \triangleq -\nabla_q R_q(\mathbf{p}_q,\mathbf{p}_{-q})$ and (3) we have

$$\nabla_r \mathbf{f}_q(\mathbf{p}) = \mathbf{D}_q(\mathbf{p})\mathbf{H}_{rq} \quad (93)$$

where

$$\begin{aligned}\mathbf{H}_{rq} &\triangleq \text{diag}\left(\left\{\frac{|H_{rq}(k)|^2}{|H_{qq}(k)|^2}\right\}_k\right) \\ \mathbf{D}_q(\mathbf{p}) &\triangleq \text{diag}\left(\left\{\frac{1}{\left(\frac{1}{|H_{qq}(k)|^2} + \sum_{r=1}^{Q} \Gamma_q^{-\delta_{rq}} \frac{|H_{rq}(k)|^2}{|H_{qq}(k)|^2} p_r(k)\right)^2}\right\}_k\right).\end{aligned} \quad (94)$$

Using (94), condition (92) becomes

$$\max_k |1 - \beta[\mathbf{D}_q(\mathbf{p})]_{kk}| + \beta \sum_{r\neq q} \max_k [\mathbf{D}_q(\mathbf{p})\mathbf{H}_{rq}]_{kk} \le \delta, \quad \forall\, \mathbf{p} \in \mathcal{P}, \forall q \in \Omega. \quad (95)$$

A sufficient condition for (95) is

$$\sum_{r\neq q} \max_k [\mathbf{H}_{rq}]_{kk} \le \frac{\delta - \max_k |1 - \beta[\mathbf{D}_q(\mathbf{p})]_{kk}|}{\beta \max_k [\mathbf{D}_q(\mathbf{p})]_{kk}}, \quad \forall\, \mathbf{p} \in \mathcal{P}, \quad \forall q \in \Omega. \quad (96)$$

It is straightforward to see that it is always possible to find proper (sufficiently small) $\beta > 0$ and $\delta \in [0,1)$ (close to one) such that (96) is satisfied, provided that $\forall\, q \in \Omega$

$$\sum_{r\neq q} \max_k [\mathbf{H}_{rq}]_{kk} = \Gamma_q \sum_{r\neq q} \max_k \frac{|\bar{H}_{rq}(k)|^2}{|\bar{H}_{qq}(k)|^2} \frac{d_{qq}^{\gamma}}{d_{rq}^{\gamma}} \frac{P_r}{P_q} < \varepsilon_q \quad (97)$$

where

$$\varepsilon_q = \min_{\mathbf{p}\in\mathcal{P}} \frac{\min_k [\mathbf{D}_q(\mathbf{p})]_{kk}}{\max_k [\mathbf{D}_q(\mathbf{p})]_{kk}} \le 1 \quad (98)$$

and $\mathbf{D}_q(\mathbf{p})$ is defined in (94). ■

ACKNOWLEDGMENT

The authors would like to thank Prof. F. Facchinei, who kindly brought to their attention [27], after this paper was completed.

REFERENCES

[1] G. Scutari, D. P. Palomar, and S. Barbarossa, “Optimal linear precoding strategies for non-cooperative systems based on game theory-Part I: Nash equilibria,” *IEEE Trans. Signal Process.*, vol. 56, no. 3, Mar. 2008.

[2] M. J. Osborne and A. Rubinstein, *A Course in Game Theory*. Cambridge, MA: MIT Press, 1994.

[3] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*. Belmont, MA: Athena Scientific, 1989.

[4] A. Scaglione, G. B. Giannakis, and S. Barbarossa, “Redundant filterbank precoders and equalizers. I. Unification and optimal designs,” *IEEE Trans. Signal Process.*, vol. 47, no. 7, pp. 1988–2006, Jul. 1999.

[5] D. P. Palomar, “Convex primal decomposition for multicarrier linear MIMO transceivers,” *IEEE Trans. Signal Process.*, vol. 53, no. 12, pp. 4661–4674, Dec. 2005.

[6] D. P. Palomar and S. Barbarossa, “Design MIMO communication system: Constellation choice and linear transceiver design,” *IEEE Trans. Signal Process.*, vol. 53, no. 10, pp. 3804–3818, Oct. 2005.
[7] D. Mitra, “An asynchronous distributed algorithm for power control in celular radio systems,” in *Proc. 4th Winlab Workshop Third Generation Wireless Information Network*, Oct. 1993, pp. 249–257.
[8] G. J. Foschini and Z. Miljanic, “A simple distributed autonomous power control algorithm and its convergence,” *IEEE Trans. Veh. Technol.*, vol. 42, no. 4, pp. 641–646, Nov. 1993.
[9] R. D. Yates, “A framework for uplink power control in cellular radio systems,” *IEEE J. Sel. Areas Commun.*, vol. 13, no. 7, pp. 1341–1347, Sep. 1995.
[10] K. K. Leung, C. W. Sung, W. S. Wong, and T. M. Lok, “Convergence theorem for a general class of power control algorithms,” *IEEE Trans. Commun.*, vol. 52, no. 9, pp. 1566–1574, Sep. 2004.
[11] C. W. Sung and K. K. Leung, “A generalized framework for distributed power control in wireless networks,” *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2625–2635, Jul. 2005.
[12] H. Ji and C. Huang, “Non cooperative uplink power control in cellular radio system,” *Wireless Netw.*, vol. 4, no. 3, pp. 233–240, Apr. 1998.
[13] N. Bambos, “Toward power-sensitive network architectures in wireless communications: Concepts, issues, and design aspects,” *IEEE Pers. Commun.*, vol. 5, no. 3, pp. 50–59, June 1998.
[14] T. Alpcan, T. Basar, R. Srikant, and E. Altman, “CDMA uplink power control as a non-cooperative game,” in *Proc. 4th Int. Symp. Personal, Indoor, Mobile Radio Communications (PIMRC)*, Sep. 8–11, 1998, vol. 3, pp. 1427–1432.
[15] C. U. Saraydar, N. Mandayam, and D. J. Goodman, “Efficient power control via pricing in wireless-data networks,” *IEEE Trans. Commun.*, vol. 50, no. 2, pp. 291–303, Feb. 2002.
[16] T. Alpcan, X. Fan, T. Basar, M. Arcak, and J. T. Wen, “Power control for multicell CDMA wireless networks: A team optimization approach,” presented at the 3rd Int. Symp. Modeling Optimization in Mobile, Ad Hoc, Wireless Networks, (WIOPT), Riva del Garda, Trentino, Italy, Apr. 3–7, 2005.
[17] M. Xiao, N. B. Shroff, and E. K. P. Chong, “Utility-based power-control scheme in cellular wireless systems,” *IEEE Trans. Netw.*, vol. 11, no. 2, pp. 210–221, Apr. 2003.
[18] M. Chiang, C. W. Tan, D. P. Palomar, D. O’Neill, and D. Julian, , W. Li and Y. Pan, Eds., “Power control by geometric programming,” in *Resource Allocation in Next Generation Wireless Networks*. Commack, NY: Nova, 2005, vol. 5, ch. 13, pp. 289–313.
[19] S. Boyd, S. J. Kim, L. Vandenberghe, and A. Hassibi, “A tutorial on geometric programming,” Elect. Eng. Dept., Stanford Univ., Stanford, CA, Tech. Rep., 2005 [Online]. Available: http://www.stanford.edu/~boyd/reports/gp_tutorial.pdf
[20] 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. 2254–2265, Dec. 2007.
[21] 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.
[22] S. T. Chung, S. J. Kim, J. Lee, and J. M. Cioffi, “A game-theoretic approach to power allocation in frequency-selective Gaussian interference channels,” in *Proc. 2003 IEEE Int. Symp. Information Theory (ISIT)*, Jun. 2003, p. 316.
[23] N. Yamashita and Z. Q. Luo, “A nonlinear complementarity approach to multiuser power control for digital subscriber lines,” *Optimiz. Methods Softw.*, vol. 19, no. 5, pp. 633–652, Oct. 2004.
[24] G. Scutari, S. Barbarossa, and D. Ludovici, “On the maximum achievable rates in wireless meshed networks: Centralized versus decentralized solutions,” in *Proc. 2004 IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, Quebec, Canada, May 17–21, 2004, vol. IV, pp. 573–576.
[25] G. Scutari, S. Barbarossa, and D. Ludovici, “Cooperation diversity in multihop wireless networks using opportunistic driven multiple access,” in *Proc. 2003 IEEE Workshop Signal Processing Advances in Wireless Communication (SPAWC)*, Jun. 2003, pp. 170–174.
[26] R. Etkin, A. Parekh, and D. Tse, “Spectrum sharing for unlicensed bands,” presented at the Allerton Conf. Communication, Control, Computing, Monticello, IL, Sep. 28–30, 2005.
[27] 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.* May 2006 [Online]. Available: http://www.ece.umn.edu/users/luozq/recent_work.html.
[28] G. Scutari, “Competition and cooperation in wireless communication networks,” Ph.D. dissertation, INFOCOM Dept., Univ. of Rome, La Sapienza, Rome,, Italy, Nov. 2004.
[29] G. Scutari, D. P. Palomar, and S. Barbarossa, “Simultaneous iterative water-filling for gaussian frequency-selective interference channels,” presented at the 2006 IEEE Int. Symp. Information Theory (ISIT), Seattle, WA, Jul. 9–14, 2006.
[30] F. Facchinei and J.-S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*. New York: Springer, 2000.
[31] R. W. Cottle, J.-S. Pang, and R. E. Stone, *The Linear Complementarity Problem*. Boston, MA: Academic, 1992.
[32] A. Nagurney, *Network Economics, A Variational Inequality Approach*, 2nd Rev. ed. Norwell, MA: Kluwer, 1999.
[33] C. E. Shannon, “Communication in the presence of noise,” *Proc. IRE*, vol. 37, no. 1, pp. 10–21, Jan. 1949.
[34] J. G. Proakis, *Digital Communications*, 3rd ed. New York: McGraw-Hill, 1995.
[35] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.
[36] W. Hirt and J. L. Massey, “Capacity of the discrete-time Gaussian channel with intersymbol interference,” *IEEE Trans. Inf. Theory*, vol. 34, no. 3, pp. 380–388, May 1988.
[37] G. G. Raleigh and J. M. Cioffi, “Spatio-temporal coding for wireless communication,” *IEEE Trans. Commun.*, vol. 46, no. 3, pp. 357–366, Mar. 1998.
[38] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
[39] A. J. Goldsmith and S.-G. Chua, “Variable-rate variable-power MQAM for fading channels,” *IEEE Trans. Commun.*, vol. 45, no. 10, pp. 1218–1230, Oct. 1997.
[40] E. Baccarelli, A. Fasano, and M. Biagi, “Novel efficient bit-loading algorithms for peak-energy-limited ADSL-type multicarrier systems,” *IEEE Trans. Signal Process.*, vol. 50, no. 5, pp. 1237–1247, May 2002.
[41] D. P. Bertsekas and J. N. Tsitsiklis, “Convergence rate and termination of asynchronous iterative algorithms,” in *Proc. 1989 Int. Conf. Supercomputing*, Irakleion, Greece, Jun. 1989, pp. 461–470.
[42] J. N. Tsitsiklis, “A comparison of Jacobi and Gauss-Seidel parallel iterations,” *Appl. Math. Lett.*, vol. 2, no. 2, pp. 167–170, 1989.
[43] J. M. Ortega and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*. Philadelphia, PA: SIAM, 2000, Ed..
[44] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1985.
[45] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2003.
[46] D. P. Palomar, “Convex primal decomposition for multicarrier linear MIMO transceivers,” *IEEE Trans. Signal Process.*, vol. 53, no. 12, pp. 4661–4674, Dec. 2005.
[47] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Nonlinear Programming—Theory and Algorithms*, 2nd ed. New York: Wiley, 1993.
[48] V. I. Istratescu, *Fixed Point Theory, an Introduction*. : Reider Publishing, 1981.
[49] G. Scutari, D. P. Palomar, and S. Barbarossa, “Asynchronous iterative water-filling for Gaussian frequency-selective interference channels: A unified framework,” *IEEE Trans. Inf. Theory* 2008 [Online]. Available: http://arxiv.org/abs/0801.
[50] D. P. Palomar and J. R. Fomollosa, “Practical algorithms for a family of waterfilling solutions,” *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 686–695, Feb. 2005.

Gesualdo Scutari (S’01–M’06) received the Electrical Engineering and Ph.D. degrees from 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, and during 2007, he was a visiting researcher at the Department of Electronic and Computer Engineering at the Hong Kong University of Science and Technology, Hong Kong. He currently holds a Postdoctoral position at INFOCOM Department, University of Rome “La Sapienza.” 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 communication aspects of wireless multiple-input multiple-output (MIMO) channels, with special emphasis on convex optimization theory and game theory applied to communications systems, and sensor networks.

Daniel P. Palomar (S'99–M'03) 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.

He is an Assistant Professor in the Department of Electronic and Computer Engineering at the Hong Kong University of Science and Technology, Hong Kong. He has held several research appointments, namely, at King's College London (KCL), London, U.K., during 1998; the Technical University of Catalonia (UPC), Barcelona, from January 1999 to December 2003; Stanford University, Stanford, CA, from April to November 2001; the Telecommunications Technological Center of Catalonia (CTTC), Barcelona, Spain, from January to December 2002; the Royal Institute of Technology (KTH), Stockholm, Sweden, from August to November 2003; the University of Rome "La Sapienza," Rome, Italy, from November 2003 to February 2004; and Princeton University, Princeton, NJ, from March 2004 to July 2006. His primary research interests include information-theoretic and signal processing aspects of MIMO channels, with special emphasis on convex optimization theory and majorization theory applied to communication systems.

Dr. Palomar received a 2004–2006 Fulbright Research Fellowship; the 2004 Young Author Best Paper Award by the IEEE Signal Processing Society; the 2006 Best Student Paper Award (corecipient) at the International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2006; 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. He is an Associate Editor of IEEE TRANSACTIONS ON SIGNAL PROCESSING, a Guest Editor of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS (JSAC) 2008 Special Issue on Game Theory in Communication Systems, and the lead Guest Editor of the IEEE JSAC 2007 Special Issue on Optimization of MIMO Transceivers for Realistic Communication Networks.

Sergio Barbarossa (S'82–M'88) received the M.Sc. degree and the Ph.D. degree in electrical engineering from the University of Rome "La Sapienza," Rome, Italy, in 1984 and 1988, respectively.

He joined the Radar System Division of Selenia in 1985 as a Radar System Designer. In 1987, he was a Research Engineer at the Environmental Research Institute of Michigan (ERIM), Ann Arbor, MI. From 1988 until 1991, he was an Assistant Professor at the University of Perugia, Italy. In November 1991, he joined the University of Rome "La Sapienza," where he is now a Full Professor and Director of Graduate Studies. He has held positions as Visiting Scientist and Visiting Professor at the University of Virginia (1995 and 1997), the University of Minnesota (1999), and the Polytechnic University of Catalonia (2001). He has participated in international projects on synthetic aperture radar systems, flown on the space shuttle, and he has been the primary investigator in four European projects on multiantenna systems and multihop systems. He is currently the scientist responsible of the European project WINSOC (on wireless sensor networks), and he is one of the primary investigators of the European project SURFACE (on reconfigurable air interfaces for wideband multiantenna communication systems). His current research interests lie in the areas of sensor networks, cooperative communications, and distributed decision.

Dr. Barbarossa was a member of the IEEE Signal Processing for Communications Technical Committee from 1998 to 2004. He served as an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 1998 to 2001 and from 2004 to 2006. He has co-edited a special issue of the *EURASIP Journal of Applied Signal Processing on MIMO Communications and Signal Processing* and the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS Special Issue on Optimization of MIMO Transceivers for Realistic Communication Networks: Challenges and Opportunities. He received the 2000 IEEE Best Paper Award from the IEEE Signal Processing Society for a coauthored paper in the area of signal processing for communications. He is the author of a research monograph titled *Multiantenna Wireless Communication Systems*."